CAMBRIDGE LAW REVIEW

Volume VI
Issue ii
Autumn 2021

Editor-In-Chief
Despoina Georgiou

Proudly Supported By
Cambridge University Law Society

Editor-in-Chief's Introduction to the Autumn Issue of Volume VI of the Cambridge Law Review

It is with great pleasure that I present the Autumn Issue of Volume VI of the Cambridge Law Review. The journal has flourished. This semester we strengthened our partnerships with the Oxford Undergraduate Law Journal, the London School of Economics Law Review, the Bristol Law Review, the Exeter Law Review, the Durham Law Review, and the Harvard Undergraduate Law Review. We also established a new partnership with the Warwick Undergraduate Law Journal and the vLex database.

As with the Spring Issue, for the Autumn Issue we received a record number of high-quality submissions. The articles published in this Issue deal with a wide range of contemporary legal matters and jurisdictions. In her article, "Judicial Activism and the Constitutional Imperative: Addressing the Issue of Spousal Privilege Under the Nigerian Evidence Act", Doctor Ayodele Morocco-Clarke examines the issue of spousal privilege under the Nigerian Evidence Act 2011. The concept of marriage in Nigeria is addressed in depth and juxtaposed with 'marriage' under common law. Issues regarding judicial activism and the enforcement of fundamental rights are also considered and evaluated in order to determine the best approach that will bring greater equality and fairness to criminal trials in Nigeria.

Assistant Professor Pranav Verma writes on the contentious topic of the death penalty in the Indian jurisdiction. His article, "The Inevitable Inconsistency of the Death Penalty in India", highlights new and robust empirical research on the administration of the death penalty and shows how it deviates from the sentencing framework developed by the Indian Supreme Court. To establish 'inevitable inconsistency', the article ventures into hypothesizing a 'best-case scenario' that removes such deviations by infusing consistency and fairness into death sentencing, to the maximum extent possible. It then highlights how, even the 'best-case scenario', fails to prevent inconsistencies or arbitrariness at a magnitude

not acceptable in a rule-based system. The article concludes that the abolition of the death penalty is the only viable end to the search of consistency in the Indian jurisprudence.

Oways Kinsara's article, "Clash of Dilemmas: How Should UK Copyright Law Approach the Advent of Autonomous AI Creations?", revisits the different manners in which today's AI creations encounter copyright law and explains why the current UK approach fails to address the issue. It examines ways forward by carefully inspecting various proposed approaches to the question of copyright ownership for AI-generated works in view of the UK regime. Upon examining different models, the article highlights numerous dilemmas in each and thus argues in favour of entrance into the public domain as the least dilemmatic and most appropriate solution for AI-generated works, with promising economic and social benefits.

Ana Rosenthal writes on the topical issue of technology surveillance. In her article, "Individuals Under Observation: The Law Responds to (Live) Facial Recognition Technology", Rosenthal engages critically with the recent case of R (Bridges) v Chief Constable of the South Wales Police in which the Appellate Court found that the use of facial recognition technology by the South Wales Police had been unlawful. In her article, Rosenthal explores the theoretical and legal implications behind facial recognition, particularly at a time when individual and fundamental rights have been brought into even sharper focus as a result of the global pandemic.

In his article, "Factortame-like Judicial Statute Disapplication and Dicey's Constitutional Orthodoxy: A Case for their Mutual Compatibility", Vincent Lafortune criticises Wade's analysis of the Factortame case. The author contends that Factortame-like judicial statute disapplication in virtue of an earlier statute is well within the boundaries of an orthodox Diceyan conception of Parliamentary Sovereignty. To reach this conclusion, Lafortune formulates a new definition of 'constitutional statute' and argues for a reconceptualization of Parliament's temporality. These two arguments, which the author names the 'technical' and 'constitutional' arguments respectively, fuse together to show that a pristine Diceyan conception of Parliamentary sovereignty enjoys more expansive bounds than previously thought, so as to even encompass judicial disapplication of an Act of Parliament or part thereof in virtue of an earlier statute, when a particular set of conditions are present..

Overall, the five articles included in the Autumn Issue constitute exceptional pieces of academic work that enrich the literature in their respective fields. They provide valuable insights into the selected areas of research, constituting enjoyable reads that would be of interest to British and international, academic and

professional audiences alike. I owe heartfelt thanks to the Managing Board and to our team of Associate, Senior, and International Editors for their dedication and work during these challenging times. Despite the difficulties caused by the COVID-19 pandemic and the subsequent lockdowns, the Editorial Board worked tirelessly to ensure the highest standards of quality for this Issue. I would also like to express my gratitude to the Honorary Board for their invaluable guidance and to the Cambridge University Law Society for their continued support, without which this Issue would not have been possible. I wish the incoming Editorial Board every success with the seventh volume and I look forward to the future growth of the Cambridge Law Review.

Despoina Georgiou
Editor-in-Chief

Cambridge Law Review

Table of Contents

Cambridge Law Review (2021) Vol VI, Issue ii, 1–23

Judicial Activism and the Constitutional Imperative: Addressing the Issue of Spousal Privilege Under the Nigerian Evidence Act

Ayodele Morocco-Clarke*

Abstract

Rules of evidence are applied in all trial proceedings in Nigeria and the Evidence Act 2011 is the principal statute that regulates the procedure of trials in the country. Under the Evidence Act, spouses are granted a form of privilege with regards to communications carried out during their marriage and, in a criminal trial, one spouse cannot be compelled to testify against the other spouse in most circumstances. This paper analyses this privilege and the non-compellability under the Evidence Act and examines why these are only extended to a select class of spouses instead of all spouses. The concept of marriage in Nigeria is addressed in-depth and juxtaposed against marriage under common law. The issue of judicial activism and the enforcement of fundamental rights are also considered and evaluated to determine the best approach which might bring greater equality and fairness to criminal trials in Nigeria.

Keywords: compellability, discrimination, judicial activism, marriage, privilege.

I. Introduction

The twin concerns of competence and compellability have been integral concepts that have arisen and dominated trials and procedures of evidence almost since the

* Legal Practitioner and Senior Lecturer at the Faculty of Law of Nile University of Nigeria, Abuja, Nigeria. LL.B. (Hons), B.L., LL.M. (Dundee). Ph.D. (Aberdeen). The author can be contacted via e-mail at ayomorocco@hotmail.com.

inception of trials. In jurisdictions all over the world, requirements and conditions have been laid down with regard to the conduct of trials by the prosecution or plaintiffs on the one hand and the defence on the other. There are rules of evidence which dictate the privilege that is applicable to various parties to a suit and the shield that might be available to an accused person in criminal proceedings. The yardstick placed by the Nigerian Evidence Act 2011[1] for the trigger of the protection of spousal privilege will be addressed.

Marriage is an institution that is generally recognised as bestowing some duties and benefits on the parties who are legally married. The status granted to spouses is one which various jurisdictions recognise and this often forms the bedrock of trust and dependency between spouses. Many jurisdictions acknowledge the need and desirability of spouses being able to communicate freely and frankly with each other and have extended a form of privilege to communications made between spouses. According to Mr. Justice McLean in *Stein v Bowman*,[2] "to break down or impair the great principles which protect the sanctities of husband and wife, would be to destroy the best solace of human existence".[3]

By virtue of the provisions of the Evidence Act, it is clear that Nigeria recognises the importance of privilege being attached to the communications between husband and wife. However, the same Act has routinely restricted the benefits of this privilege to only a select class of spouses. Thus, a colossal hurdle appears to have been routinely placed in the path of other classes of married persons not recognised by the Evidence Act. These other classes of spouses have been left open by the Evidence Act to having spousal communications which have been carried out during the subsistence of their marriage exposed to the public. This article shall seek to excavate the intention for what appears to be a blockade of an unacknowledged group of married individuals. The ultimate intention of this article is to shed more light on certain provisions of the Evidence Act and proffer solutions and recommendations to ensure that, where possible, there is consistency, uniformity, equality and fairness in the provisions of the Evidence Act as well as in the manner litigants are treated in proceedings before requisite courts of law which are bound to follow the provisions of the Evidence Act in Nigeria.

II. Competence and Compellability of Spouses Under the Nigerian Evidence Act 2011 and Under Common Law

When analysing and examining trials, it is clear that litigation often involves a lot of legal and factual manoeuvrings on the part of the parties, their counsels,

1 Hereinafter referred to as 'The Evidence Act'.

2 38 US 209.

3 ibid 223.

and the trial judge. Judicial proceedings in courts in Nigeria are governed by the Evidence Act 2011.[4] The Evidence Act contains rules and procedures governing facts, witnesses and the taking and giving of evidence in court/judicial proceedings in Nigeria.

Under Nigerian law, the baseline and default position are that anyone can be a competent witness. In the law of evidence, competence has been defined as "the presence of those characteristics, or the absence of those disabilities, which render a witness legally fit and qualified to give testimony in a court of justice".[5] The presumption of competence may be displaced by a court on various grounds provided by the governing statute. Thus, according to Section 175(1) of the Evidence Act:

> "All persons shall be competent to testify, unless the court considers that they are prevented from understanding the questions put to them, or from giving rational answers to those questions, by reason of tender years, extreme old age, disease, whether of body or mind, or any other cause of the same kind".

From the foregoing, it is clear that it is the court which is adjudicating over each trial that is empowered to determine whether any witness is not competent to testify before it. In the absence of the court declaring a witness as not competent to testify, such witness is deemed competent to testify before the court. Furthermore, Section 178[6] states that in all civil proceedings the parties to the suit and the husband or wife of any party to the suit shall be competent witnesses and Section 179[7] goes on to provide that in criminal cases, the defendant, his wife or her husband, or any person jointly charged with such defendant and tried at the same time, and the wife or husband of the person so jointly charged, is competent to testify. Therefore, it is clear that spouses of accused persons are competent witnesses at all times.

It is imperative to understand, however, that the fact of a witness being competent is quite different to the determination of whether such a witness is compellable to testify before a court. A compellable witness is a "witness who may lawfully be required to give evidence and who may be punished for contempt of court for refusal".[8] Compellability can be defined as the capability to coerce an

4 The Evidence Act 2011 governs both civil and criminal proceedings in all courts in Nigeria apart from civil proceedings in Area Courts, Customary Courts, Sharia Courts of Appeal and Customary Courts of Appeal of all States and the Federal Capital Territory. Section 256 of the Evidence Act 2011.

5 Henry Campbell Black, *Black's Law Dictionary* (4th edn., West Publishing 1968) 355.

6 Evidence Act 2011.

7 ibid.

8 Oxford Reference, 'Overview: Compellable Witness' <www.oxfordreference.com/view/10.1093/oi/authority.20110803095628763> accessed on 17/05/2021.

individual to do what he ought to or obliged to do.[9] This work shall focus on the issue of the compellability of spouses in judicial proceedings and any privilege that might accrue to spouses when it comes to giving testimonies in court proceedings.

Section 182 of the Evidence Act[10] sets out the provisions relating to the compellability of spouses and states as follows,

> "(1) When a person is charged –
>
> (a) with an offence under sections 217, 218, 219, 221, 222, 223, 224, 225, 226, 231, 300, 301, 340, 341, 357 to 362, 369, 370, or 371 of the Criminal Code;
>
> (b) subject to section 36 of the Criminal Code with an offence against the property of his wife or her husband; or
>
> (c) with inflicting violence on his wife or her husband, the wife or husband of the person charged shall be a competent and compellable witness for the prosecution or defence without the consent of the person charged.
>
> (2) When a person is charged with an offence other than one of those mentioned in subsection (1) of this section. the husband or wife of such person is a competent and compellable witness but only upon the application of the person charged.
>
> (3) Nothing in this section shall make a husband compellable to disclose any communication made to him by his wife during the marriage or a wife compellable to disclose any communication made to her by her husband during the marriage".

By virtue of the provisions of Section 182(2) a husband or wife of an accused person is not a competent or compellable witness except in instances where the accused person makes an application for such spouse to testify.[11] This is also the position under common law. In *R v Mount*,[12] three men were convicted of breaking into a shop and one of the witnesses for the prosecution had been the wife of one of the convicted men. The conviction of the men was quashed on appeal

[9] B.E. Ewulum and Obinna Mbanugo, 'Competence and Compellability Under the Evidence Act of Nigeria' (2017) 2(1) *SAJMS* 1. Available online at <www.nigerianlawguru.com/articles/practice%20and%20procedure/COMPETENCE%20AND%20COMPELLABILITY%20UNDER%20THE%20NIGERIAN%20EVIDENCE%20ACT.pdf> accessed 17/05/2021.

[10] Evidence Act (n 6).

[11] This applies to all offences not mentioned in Section 182(1) of the Evidence Act 2011.

[12] [1934] 24 Cr. App. R. 135.

as it was held that the wife of an accused was not a competent or compellable witness against the accused persons. The wife would have been a compellable and competent witness for the prosecution and against the co-accused of her husband if the husband had entered his plea and been convicted (or acquitted) prior to his wife testifying as a witness. In *Leach v R*,[13] Lord Atkinson said that the principle of not compelling a spouse to testify against their partner was 'deep seated' in the common law. Furthermore, in *Hoskyn v Metropilitan Police Commissioner*,[14] Lord Wilberforce stated,

> "a wife is in principle not a competent witness on a criminal charge against her husband. This is because of the identity of interest between husband and wife and because to allow her to give evidence would give rise to discord and to perjury and would be, to ordinary people, repugnant".[15]

Sections 182(1)(b) and (c) and 182(3) of the Evidence Act is a codification of the common law rule as set out in the case of *Stein v Bowman*[16] which is to the effect that, "the wife is not competent, except in cases of violence upon her person, directly to criminate her husband, or to disclose that which she has learned from him in their confidential intercourse".[17] Mr. Justice McLean gave the reason for the rule when he stated:

> "This rule is founded upon the deepest and soundest principles of our nature. Principles which have grown out of those domestic relations, that constitute the basis of civil society; and which are essential to the enjoyment of that confidence which should subsist between those who are connected by the nearest and dearest relations of life. To break down or impair the great principles which protect the sanctities of husband and wife, would be to

[13] [1912] AC 305 at 311.
[14] [1979] AC 474.
[15] ibid.
[16] 38 US 209.
[17] Per McLean J at 222. Also Section 187 of the Evidence Act 2011, which provides, "No husband or wife shall be compelled to disclose any communication made to him or her during marriage by any person to whom he or she is or has been married nor shall he or she be permitted to disclose any such communication, unless the person who made it or that person's representative in interest, consents, except in suits between married persons or proceedings in which one married person is prosecuted for an offence specified in section 182 (1) of this Act".

destroy the best solace of human existence".[18]

Thus, it is clear that the sanctity of marriage and the unity of spouses and the family is one which the law recognises, respects, and seeks to protect. Within the sacred institution of marriage, the law and courts are loathe to foist a situation of discord between a married couple.

The issue of the compellability or otherwise of the spouse of an accused person is one which has the potential of adversely affecting the defence of such an accused person, as where the spouse of the accused can be compelled to give evidence, hitherto confidential information that might have been divulged by the accused to his or her spouse might become admissible in court to the detriment of the accused. In the case of *Ayo v The State*,[19] the court held that the spouse of an accused person is not a competent or compellable witness for the prosecution unless upon the application of the accused person. The court further made it clear that the section of the Evidence Act under consideration was not talking about the consent of the accused person being obtained but that the spouse of the accused person can only testify for the prosecution upon the *application* of the accused. Thus, the accused person must apply to the court to have his/her wife or husband testify as a witness on behalf of the prosecution.

III. "Marriage": What Does the Evidence Act 2011 Mean by This Term?

Having addressed the issue of the competence and compellability of spouses in Section II of this work and due to the high stakes that might be involved, it is imperative to understand the concept of marriage and the twin terms of 'husband' and 'wife' within the ambit of the Evidence Act. Ordinarily, it would be easy to presume that the terms 'husband' and 'wife' mean a man and a woman who have gone through a valid marriage ceremony and union. According to the Legal Dictionary, the terms 'husband' and 'wife' mean "a man and woman who are legally married to one another and are thereby given by law specific rights and duties resulting from that relationship".[20] The status of a husband or a wife only becomes activated by virtue of the man and woman entering into a marriage. Black's Law Dictionary defines 'husband and wife' as "one of the great domestic relationships; being that of a man and woman lawfully joined in marriage, by

[18] *Stein v Bowman* (n 2) 223.

[19] [2010] All FWLR (Pt. 530) 1377.

[20] The Free Dictionary, 'Legal Dictionary' <https://legal-dictionary.thefreedictionary.com/Husband+and+wife> accessed on 19/05/2021.

which, at common law, the legal existence of a wife is incorporated with that of her husband".[21]

It therefore is obvious that marriage is the key institution of reference when seeking to understand the twin concepts of 'husband and wife' and Black's Law Dictionary states that marriage is the

> "[…] civil status, condition, or relation of one man and one woman united in law for life, for the discharge to each other and the community of the duties legally incumbent on those whose association is founded on the distinction of sex".[22]

In the case of *Davis v Davis*,[23] marriage was denoted as "[…] the act, ceremony, or formal proceeding by which persons take each other for husband and wife". The Nigerian Interpretation Act 1964 defines a "Monogamous marriage" as meaning "a marriage which is recognised by the law of the place where it is contracted as a voluntary union of one man and one woman to the exclusion of all others during the continuance of the marriage".[24] Whilst the Interpretation Act restricts it definition of a marriage to only monogamous marriages, there are other types of marriages applicable and recognised in Nigeria.[25] The Evidence Act acknowledges the existence of customary and Islamic marriages.[26] Under Section 166, it states:

> "When in any proceeding whether civil or criminal, there is a question as to whether a man or woman is the husband or wife under Islamic or Customary law, of a party to the proceeding, the court shall, unless the contrary is proved, presume the existence of a valid and subsisting marriage between the two persons where evidence is given to the satisfaction of the court of cohabitation as husband and wife by such man and woman".

The Supreme Court also acknowledged the existence of customary and Islamic marriages in the case of *Jadesimi v Okotie-Eboh*.[27] Furthermore, the Evidence Act explicitly states its yardstick for determining who a husband and a wife is and

[21] Black (n 5) 875.
[22] ibid 1123.
[23] (1934) 119 Conn. 194, 175 A. 574-575.
[24] Section 18 of the Interpretation Act 1964.
[25] These are customary and Islamic marriages.
[26] Section 258(1) of the Evidence Act.
[27] (1996) 2 NWLR Part 128 at 142-148.

gives its legislative nod and acknowledgement only to marriages prescribed in its interpretation section Section 258(1), which states:

> "In this Act 'wife' and 'husband' mean respectively the wife and husband of *a marriage validly contracted under the Marriage Act, or under Islamic law or a Customary law applicable in Nigeria, and includes any marriage recognised as valid under the Marriage Act* (emphasis added)".

By the foregoing provision of the Evidence Act, it is incontrovertible that any other provisions within the Act which refers to the terms 'husband' and 'wife' intend to cover any husband and/or wife of a marriage which has been validly conducted under Nigerian Islamic or Customary Law or a marriage conducted under the Marriage Act. There are some schools of thought which hold that the Evidence Act when referring to 'husband' and 'wife' does not mean or apply to husbands and wives of customary or Islamic marriages.[28] However, whilst this opinion might have been applicable and held true under the old repealed Evidence Act,[29] it does not hold true today under the current Evidence Act 2011 as is seen by Section 258(1) of the Act.

Having established that the provisions of the Evidence Act apply to spouses of both Nigerian customary and Islamic marriages, marriages under the Marriage Act[30] have to be examined in other to determine the spouses who can take refuge under the provisions of Section 182(2) and (3) of the Evidence Act 2011. The Marriage Act sets out all marriages considered as valid under the Act. Consequently, Sections 34 and 35 of the Marriage Act state as follows:

> "34. All marriages celebrated under this Act shall be good and valid in law to all intents and purposes.
>
> 35. Any person who is married under this Act, or whose marriage is declared by this Act to be valid, shall be incapable, during the continuance of such marriage, of contracting a valid marriage under customary law, but, save as aforesaid, nothing in this Act contained shall affect the validity of any marriage contracted

[28] Ewulum and Mbanugo (n 9) 6.

[29] Chapter E14 Laws of the Federation of Nigeria 2004. It should be clarified and understood that under the old Evidence Act, Section 162 stipulated that communications between spouses of polygamous marriages were not privileged and spouses of such marriages were competent and compellable for both the prosecution and defendant. However, communications of spouses under an Islamic marriage (even though polygamous in nature) were privileged by virtue of the proviso to Section 162 of that same Act.

[30] Chapter M6 Laws of the Federation of Nigeria 2004.

> under or in accordance with any customary law, or in any manner apply to marriages so contracted".

The Marriage Act by virtue of the foregoing provisions explicitly acknowledges the validity of all marriages which have been contracted under the Act and does not seek to invalidate any marriage contacted under native law and custom/customary law. The provisions of Section 35 of the Act show that the Act deduces that there might be people who will purport to contract a marriage under customary law whilst already being to another party married under the Act. Section 33(1) of the Marriage Act expressly prohibits the aforesaid scenario when it states that "no marriage in Nigeria shall be valid where either of the parties thereto at the time of the celebration of such marriage is married under customary law to any person other than the person with whom such marriage is had". Therefore, while there is an acknowledgment of the existence of customary marriages[31] by the Marriage Act, the Act does not seek to invalidate or seek to dispute the validation of any marriage celebrated under customary law. In addition to the foregoing, the Marriage Act, recognises the validity of foreign marriages. It must however be pointed out that it is not all foreign marriages that the Marriage Act recognises as valid marriages. By virtue of the provisions of Section 49 of the Marriage Act, the foreign marriages the Act recognises as valid a specific niche of marriages as it states, marriages

> "[…] between parties one of whom is a citizen of Nigeria, if it is contracted in a country outside Nigeria before a marriage officer in his office, shall be as valid in law as if it had been contracted in Nigeria before a registrar in the registrar's office".

Boosting and complementing the provisions of Section 49 of the Act, Sections 50-52 of the Marriage Act provide as follows:

> "50. For the purposes of this Act, every Nigerian diplomatic or consular officer of the rank of Secretary or above shall be regarded as a marriage officer in the country to which he is accredited.
>
> 51. The office used by a marriage officer for the performance of his diplomatic or consular duties shall be regarded as the marriage officer's office for the purposes of this Act.

[31] Traditionally, Islamic marriages were deemed a form of customary marriage or marriage under customary law. S. M. Olokooba, 'Analysis of Legal Issues Involved in the Termination of "Double-Decker" Marriage Under Nigeria Law' (2007-2010) Nigerian Current Law Review 196. Available at: <www.nials-nigeria.org/pub/NCLR7.pdf> accessed on 20/05/2021.

> 52. Subject to the modifications specified in section 53 this Act shall apply in relation to a marriage contracted before a marriage officer as nearly as may be as it applies in relation to a marriage contracted before a registrar".

Thus, the Marriage Act expressly recognises the validity of only two types of statutory marriages, *viz*, marriages conducted under the Act in Nigeria and only foreign marriages conducted by a diplomatic or consular officer in his office as dictated by Sections 49-52 of the Act.

With regard to this paper and the subject-matter under consideration, the significance of the recognition of validity of marriages under the Marriage Act is magnified by the weight the Evidence Act has placed on the Marriage Act by partly tying its definition of 'husband' and 'wife' to the marriages contracted under the Marriage Act or marriages recognised as valid under the Marriage Act. As stated above, the Evidence Act also recognises marriages conducted under Nigerian Customary law and Nigerian Islamic law. Thus, going by the combined readings of the requisite provisions of the Evidence Act and the Marriage Act, the only types of husbands and wives recognised under the Evidence Act are the husbands and wives of marriages contracted under one of four headings/categories:

i. Marriage contracted in Nigeria under the provisions of the Marriage Act;

ii. Foreign marriages carried out pursuant to Sections 49-52 of the Marriage Act;

iii. Marriages contracted under Customary Law applicable in Nigeria; and

iv. Marriages contracted under Islamic Law applicable in Nigeria.

Therefore, if a husband or wife does not fall under any of the four categories of marriages mentioned above, such a spouse cannot take refuge under the provisions of Section 182(2) and (3) of the Evidence Act with regards to the privilege afforded on the compellability of persons as witnesses. The implication of the foregoing is that anyone who has validly and legally contracted a marriage in any jurisdiction outside Nigeria (and who did not do so in line with the provisions of Sections 49-52 of the Nigerian Marriage Act) will not have their marriage recognised as valid under the Evidence Act in Nigeria. These include spouses who have contracted valid monogamous marriages (whether Christian or statutory) in

foreign countries and couples who have contracted valid polygamous marriages (via the local/customary and/or Islamic law) in foreign jurisdictions. It must be emphasised that the only type of foreign marriage which the Marriage Act affords any validation is one in which at least one of the parties is a Nigerian citizen.[32]

IV. What Next?

The manifest absurdity brought on by the provisions of the Evidence Act with regard to the compellability of a spouse coupled by the classes of spouses whom those provisions cover, as examined in-depth in Sections II and III above, is heightened by the fact that we live in times wherein the world/planet is considered as a global village and the routine migration which was common in the 20th century is even more so in the 21st century, with millions of Nigerian citizens scattered around the globe.[33] These provisions mean that a plethora of Nigerian citizens (and other foreigners) who have validly and legally got married in other jurisdictions are left out in the cold and can have their spouses compelled to give evidence against them before Nigerian courts on the sole ground that their marriages are not recognised as valid under the Nigerian Evidence Act and correspondingly their communications with their husbands or wives [as the case may be], during the marriage, is not recognised as privileged under the Evidence Act (because they are not the category of husband or wife prescribed under the Act).

There are many Nigerian citizens domiciled in Nigeria who elect to celebrate or contract their marriages in other jurisdictions. Also, as specified above, there are numerous Nigerian citizens who are resident in other jurisdictions and who choose to celebrate their marriages in those foreign jurisdictions. Furthermore, there are foreigners who have celebrated their marriages in other foreign jurisdictions but for one reason or the other either choose to reside in Nigeria or to conduct their business in Nigeria. These classes of people have the validity of their marriage negated under the Evidence Act. The sheer legislative sophistry wrought by the provisions of the Evidence Act which implicitly deny recognition to the validity of foreign marriages (this includes foreign Islamic and foreign customary marriages which are legal, valid and binding in the jurisdiction where those marriages were

[32] Section 49 of the Marriage Act.

[33] Sharkdam Wapmuk, Oluwatooni Akinkuotu and Vincent Ibonye, 'The Nigerian Diaspora and National Development: Contributions, Challenges, and Lessons from Other Countries' (2014) Kritika Kultura 292-342, 295. Available at: <www.researchgate.net/publication/265569685_THE_NIGERIAN_DIASPORA_AND_NATIONAL_DEVELOPMENT_CONTRIBUTIONS_CHALLENGES_AND_LESSONS_FROM_OTHER_COUNTRIES> accessed on 21/05/2021. See also, Andrew S. Nevin et. al., 'Strength from Abroad: The Economic Power of Nigeria's Diaspora' (2019) PricewaterhouseCoopers 2, 5. Available at: <www.pwc.com/ng/en/pdf/the-economic-power-of-nigerias-diaspora.pdf> accessed on 21/05/2021.

contracted or celebrated as well as foreign monogamous/statutory marriages) shows an unbridled bias which cannot be defended on any logical ground and this stance adopted by the Act is one deserving of robust condemnation.

It should be highlighted that though the Marriage Act restricts its recognition to the validity certain marriages celebrated pursuant to the Act, the Matrimonial Causes Act[34] is not so encumbered. It implicitly acknowledges the validity of marriages contracted or celebrated within foreign countries so long as the said marriage complied with the requirements for the celebration of marriages in that jurisdiction. Thus, Section 3(1)[35] of the Matrimonial Causes Act states:

> "Subject to the provisions of this section, a marriage that takes place after the commencement of this Act is void in any of the following cases but not otherwise, that is to say, where –
>
> (a) either of the parties is, at the time of the marriage, lawfully married to some other person;
>
> (b) the parties are within the prohibited degrees of consanguinity or, subject to section 4 of this Act, of affinity;
>
> (c) *the marriage is not a valid marriage under the law of the place where the marriage takes place, by reason of a failure to comply with the requirements of the law of that place with respect to the form of solemnisation of marriages* (emphasis added)".

Traditionally under common law, recognition of the validity of marriages were made using conflict of laws rules, and the principal applicable rule was the *lex loci celebrationis*.[36] This principle regards the law of the place of celebration of a marriage. In private international law, this law governs such questions as the

[34] Chapter M7 Laws of the Federation of Nigeria 2004.

[35] Of interest here are the provisions of Section 3(1)(c) of the Matrimonial Causes Act.

[36] There were other conflict principles applied as well, when necessary, e.g. the *Lex Domicili* and the *Lex Fori* rules. For the purpose of the issue under determination here, the *Lex Loci Celebrationis* principle is endorsed simply for the determination of the validity of the marriage with regards to whether it has met the requirements of the governing law of the country wherein it is celebrated or contracted. It is of course acknowledged that with regards to marriages of same sex couples/spouses, Nigerian law expressly prohibits such marriages and explicitly bars any recognition of such marriages as valid within Nigeria. This is in line with the provisions of the Same Sex Marriage (Prohibition) Act 2013 (Particularly of reference here are the provisions of Sections 1-3 of the Act). This is a different issue than that of the acknowledgment only of some classes of marriages under the Evidence Act.

formalities required for a marriage.[37] In *Cmvchund v Barker*,[38] Lord Hardwicke stated, "[…] so in matrimonial cases, they are to be determined according to the ceremonies of marriage in the country where it was solemnised". In *Scrimshire v Scrimshire*,[39] the *lex loci celebrationis* rule was extensively examined and discussed.[40] In that case, two English minors got married in France, but did not strictly observe the requirements of the laws of France. When the issue of the validity of the marriage arose, it was argued that the marriage should be determined according to English law on the basis that the parties were English subjects and domiciled in England. This position was rejected by the court which took the position that under English law, "[...] English marriages are to be deemed good or bad according to the law of the place where they are made".[41] Sir Edward Simpson, who was the judge in the case affirmed that it was the French law which was to be applied in that case when he stated,

> "It is the law of this country [England] to take notice of the laws of France, or of any foreign country, in determining upon marriages of this kind. The question being in substance this – whether, by the law of this country, marriage contracts are not to be deemed good or bad according to the laws of the country in which they are formed, and whether they are not to be construed by that law".[42]

The Matrimonial Causes Act through its provisions of Section 3(1)(c) has adopted the *lex loci celebrationis* principle in a plain attempt to ensure that the courts are vested with powers and the jurisdiction to adjudicate over marital matters and determine marriages where necessary, especially when a party or both parties are domiciled in Nigeria and are seeking a resolution of some marital issues before Nigerian courts. And historically, Nigeria being a common law country has applied rules of common law in its judicial determination at various levels.[43] Numerous

[37] Oxford Reference, 'Overview: Lex Loci Celebrationis' <www.oxfordreference.com/view/10.1093/oi/authority.20110803100103412> accessed on 21/05/2021.

[38] [1744] 1 Atk. 22, 50.

[39] [1752] 2 Hag. Con. 393.

[40] Hamid Tahenni, *Conflict of Law Rules in Marriage: An Approach Based on the Co-ordination of the Relevant Policy Considerations*, (Ph.D. Thesis, University of Glasgow 1995) 14. Available at: <http://theses.gla.ac.uk/5009/1/1995TahenniPhd.pdf> accessed on 22/05/2021.

[41] *Scrimshire v Scrimshire* (n 40) 402.

[42] ibid 407-408.

[43] Efe Etomi and Elvis Asia, 'Family Law in Nigeria: Overview – Marriages' (01 October 2020) Thomson Reuters Practical Law. Available at: <https://uk.practicallaw.thomsonreuters.com/6-613-4665?transitionType=Default&contextData=(sc.Default)&firstPage=true> assessed on 22/05/2021.

conflict of law principles have been therefore applied by Nigerian courts through the ages. The *lex loci celebrationis* rule is thus applicable in Nigeria.

The Evidence Act 2011 is predominantly a modification of the old Evidence Act (Cap M14 LFN 2004), however there are significant differences in some of the provisions of the new Act, with the introduction of new provisions and concepts and the jettisoning of other provisions from the old Act. One of the new provisions within the current Evidence Act is the recognition of the validity of marriages contracted under Nigerian customary law. A significant reason for overhauling the old Act and embarking on the enactment of a new Evidence Act was because the old Act had become inadequate to deal with emerging and new procedures of evidence which were integral for the proper dispensation of justice in the 21st Century.

Many jurisdictions have overhauled their legislations and procedures as exigencies demand and require. For example, in the United Kingdom, the question of the validity of marriages is one that has been germane to the determination of numerous immigration applications. The case of *Kareem v Secretary of State for the Home Department*[44] is one of the cases arising from an immigration application which turned on the recognition of the validity of a foreign (proxy) marriage by a different sovereign State. This case concerned a Nigerian national whose Dutch citizen 'wife' was working in the UK. He contended that they married by proxy in Nigeria but neither of them attended the ceremony. It was said that their marriage was held in accordance with customary law. A marriage certificate was issued when the local customary court subsequently registered the ceremony. The claim was supported by numerous documents but the decision-maker was unconvinced that Kareem was married as claimed.[45]

In the circumstances, the appellate court determined that "[…] the legal system of the nationality of the Union citizen must itself govern whether a marriage has been contracted".[46] The decision turned on the fact that the decision-maker did not believe that there was a marriage contracted despite being provided with a marriage certificate, an order and an affidavit. Thus, the decision maker decided to ignore the *lex loci celebrationis* approach and instead apply the rationale of the recognition of the marriage by Kareem's spouse's home country of the Netherlands. A few years later, a different track was taken in the case of *Awuku v*

[44] This case is also referred to as *Kareem (Proxy Marriages – EU Law)* [2014] UKUT 24 (IAC).

[45] Asad Ali Khan, 'Overruling Kareem: Proxy Marriages and Recognition under European Union Law (02 January 2017) *United Kingdom Immigration Law Blog*. Available at: <https://asadakhan.wordpress.com/2017/01/02/overruling-kareem-proxy-marriages-and-recognition-under-europe-an-union-law/> accessed on 22/05/2021.

[46] *Kareem* (n 45), para 18. Available at: <www.bailii.org/uk/cases/UKUT/IAC/2014/%5B2014%5D_UKUT_24_iac.html> accessed on 22/05/2021.

Secretary of State for the Home Department,[47] which had facts which were similar to those in the *Kareem* case and wherein the appellate court held that marriage by proxy would be treated as valid in England if recognised by the local law of the place it is contracted. The adjudicating judge, Lloyd Jones, LJ., held that,

> "CMG Ockelton, Judge McKee, Deputy Judge McCarthy had been wrong in Kareem to create a new rule of private international law requiring reference to the law of the member state of the EU national in order to determine the marital status of a spouse or partner for the purposes of Directive 2004/38/EC".[48]

As has been stated above, it is routine for new rules and procedures be adopted from time to time as virtually every jurisdiction deems it expedient to adapt and amend their laws, regulations, rules etc. to address current times and needs and to provide more effective regulation and justice. There is no rational reason why the Evidence Act 2011, failed to address the archaic recognition of marriage provisions it inherited from its predecessors and instead adopt the stance taken by the Matrimonial Causes Act, which is a legislation that dates back to 1970. There was more than ample opportunity to adjust its validity of marriage stance as was done with the new recognition of the validity of marriages conducted under Nigerian Customary Law.

Obviously, under the rules of statutory interpretation, it could be successfully argued that it was never the intention of the framers of the Evidence Act to extend recognition to all foreign marriages whether or not they are valid in the jurisdiction they were celebrated in and neither was it their intention to bestow any privileges under the Act upon foreign spouses. This undoubtedly would be an argument which employs in the first instance, the literal interpretation of statutory construction to give the interpretation of the provisions of the Evidence Act their literal meaning. In addition, there will be the use of the *Expressio Unius Est Exclusio Alterius* rule of statutory interpretation which translates as, "the expression of one thing is the exclusion of the other".[49] This rule advocates that "when one or more things of a class are expressly mentioned others of the same class are excluded".[50]

[47] [2017] EWCA Civ 178.

[48] Asad Ali Khan, 'Lex Loci Celebrationis and Proxy Marriage in English Law' (15 April 2017) *United Kingdom Immigration Law Blog*. Available at: <https://asadakhan.wordpress.com/2017/04/15/lex-loci-celebrationis-and-proxy-marriage-in-english-law/> accessed on 22/05/2021.

[49] Duhaime.org, 'Duhaime's Law Dictionary: Expressio Unius Est Exclusio Alterius Definition' <www.duhaime.org/LegalDictionary/E/ExpressioUniusEstExclusioAlterius.aspx> accessed on 22/05/2021.

[50] Merriam-Webster, 'Dictionary: Expressio Unius Est Exclusio Alterius <www.merriam-webster.com/legal/expressio%20unius%20est%20exclusio%20alterius> accessed on 22/05/2021.

The case of *R v Inhabitants of Sedgley*,[51] illustrates the judicial application of the *Expressio Unius Est Exclusio Alterius* rule. In this case, the court decided that a levy which was applicable and issued on the occupiers of "lands, houses and coal mines" pursuant to the 1601 Poor Relief Act could not be applied on the owners or occupiers of other types of mine. Nigerian courts have also employed the *Expressio Unius Est Exclusio Alterius* rule of interpretation in a raft of cases. These cases include *Bello Musa Magaji v Alhaji Ishola Are Ogele*[52] and *Ehuwa v O.S.I.E.C.*[53] Thus, an argument along these lines could be validly made with regards to the exclusion of other types of marriages by the Evidence Act restricting its recognition only to marriages contracted under Nigerian Islamic law or Nigerian customary law and statutory marriages under the provisions of the Marriage Act referred to in Section 258(1) of the Evidence Act. It must however be said that while the foregoing might be a tenable argument, it does not come across as rational because the Evidence Act and the provisions thereunder relating to husbands and wives are not in danger of losing their veracity and effect if those provisions and clauses are extended to all spouses, whether local to Nigeria or foreign, so long as the fact and existence of their marriage can be validly and adequately proven.

V. A Way Forward?

From the discussion in Section IV above, it is clear that courts bear a duty to adjudicate matters in line with extant statutes and to primarily interpret statutes using their plain unambiguous meanings. Whilst this is the base position taken in the interpretation of statutes, courts are not only courts of law, but of justice and they do have the inherent power to adjudicate over matters and interpret statutory provisions in a manner that will negate injustice and ensure that justice is served.

Courts are often faced with situations in which they have to decide between applying the strict letter of the law or refusing to apply a specific piece of legislation because it is null and void on account of its contravention with the Constitution or because it will be unjust to uphold and apply such legislation. When a court finds itself at such a crossroad, it is within the power of such court to – where justice calls for it – engage in some judicial activism.

Judicial activism is advocated here because a strict reliance on the literal interpretation of the Evidence Act if applied by the courts in instances of spousal privilege will not only be discriminatory, but might result in substantial injustice. Judicial activism can be stated to be the exercise of judicial review or decision by a court or judge(es) in which the court/judge is unafraid of striking down and

[51] [1831] 2 B & Ald 65.
[52] 3PLR/2012/16 (CA); or [2012] LPELR-9476 (CA).
[53] [2006] 18 NWLR (Pt. 1012) 544 at 568-569.

invalidating legislative or executive actions or departing from judicial precedent[54] especially when dealing with constitutional or statutory interpretation. Black's Law Dictionary's definition of judicial activism is, "judicial philosophy which motives judges to depart from strict adherence to judicial precedent in favour of progressive and new social policies which are not always consistent with the restraint expected of appellate Judges".[55] In addition, the Merriam-Webster Dictionary defines judicial activism as, "the practice in the judiciary of protecting or expanding individual rights through decisions that depart from established precedent or are independent of or in opposition to supposed constitutional or legislative intent".

Some people perceive judicial activism in a pejorative manner. Such views consider any action by the judiciary which is not in strict compliance of traditional roles as overreaching and an encroachment by the judiciary of the rightful roles of the legislative or executive arms of government. It is sometimes argued that judicial activism goes much further than blurring the lines between the separation of powers of the three arms of government, but indeed involves a usurpation of some of the powers of either the legislative or executive arms by the judiciary.[56] However, it must be pointed out that there is a difference between judicial activism and judicial overreach. Judicial overreach comes into play when the judiciary has overextended itself past its judicial mandate and trespasses into the jurisdiction or mandate of another arm of government. It has been stated that there is a thin line of division between judicial activism and judicial overreach. Whilst judicial activism is said to imply the exertion by the judiciary of its powers to herald justice and is usually utilised for the benefit of the society, judicial overreach is when the judiciary exercises its power in a manner that is blatantly beyond the remit of its scope of authority and boundary. There is no disputing that there are different views and different interpretations of what judicial activism is. However, for the purpose advanced here, it is viewed as a catalyst for change from the slavish monotonous dogged adherence to skewed or lopsided legislation and a reinforcement of the fundamental rights of people in Nigeria.

According to Lord Macmillan in *Donoghue v Stevenson*,[57] "the criterion of judgment must adjust and adapt itself to the changing circumstances of life".[58]

[54] Kermit Roosevelt, 'Judicial Activism' (*Britannica Encyclopedia: Law*) <www.britannica.com/topic/judicial-activism> accessed on 25/05/2021.

[55] Henry Campbell Black et al., *Black's Law Dictionary* (6th edn., Centennial Edition 1891-1991, West Publishing 1998) 984.

[56] Ibrahim Imam et al., 'Judicial Activism and Intervention in the Doctrine of Political Questions in Nigeria: An Analytical Exposition' (2011) 1(2) African Journal of Law and Criminology 50-51.

[57] [1932] AC 562 (HL).

[58] ibid 619.

This opinion was cited by Lord Denning with aplomb in his dissenting opinion in the case of *Candler v Crane, Christmas & Co.*,[59] when he stated:

> "Let me first be destructive and destroy the submissions put forward by Mr. Foster. His first submission was that a duty to be careful in making statements arose only out of a contractual duty to the plaintiff or a fiduciary relationship to him. Apart from such cases, no action, he said, had ever been allowed for negligent statements, and he urged that this want of authority was a reason against it being allowed now. *This argument about the novelty of the action does not appeal to me in the least. It has been put forward in all the great cases which have been milestones of progress in our law, and it has always, or nearly always, been rejected. If you read the great cases of Ashby v White,*[60] *Pasley v Freeman,*[61] *and Donoghue v Stevenson you will find that in each of them the judges were divided in opinion. On the one side there were the timorous souls who were fearful of allowing a new cause of action. On the other side there were the bold spirits who were ready to allow it if justice so required. It was fortunate for the common law that the progressive view prevailed* (emphasis added)".

Whilst the *Candler* case was with regards to an action for economic loss resulting from negligent misstatement, it shows the trajectory taken by judges audacious enough to deviate from the stringent and narrow path of legislative interpretation or the application of rules and procedure. Furthermore, approximately thirteen years later, the English House of Lords in the case of *Hedley Byrne & Co. Ltd. v Heller & Partners Ltd.*,[62] gave their judicial nod to the dissenting judgement of Lord Denning in the *Candler* case and thus a new course was charted. That was a sublime manifestation of the power of judicial activism in action and the change it can usher in with regard to ensuring that the hands of the courts are not fettered inordinately by the requirement of strict adherence to *stare decisis*, rules or procedure.

In Nigeria, the courts have been known to engage in judicial activism and when they have done so (particularly when the court's decision goes against the grain of legislative or executive action), they have on occasion faced a backlash or a barrage of criticisms.[63] However, judicial activism by Nigerian courts have on many occasions provided a panacea for beleaguered individuals and bodies

[59] [1951] 2 KB 164 (CA).
[60] [1703] 2 Ld. Raym. 938.
[61] [1789] 3 Term Rep. 51.
[62] [1964] AC 465.
[63] Ibrahim Imam (n 57) 51.

who have struggled under the yoke of unconstitutional or ambiguous legislation or even oppressive administrative orders. In the case of *National Union of Electricity Enterprises v Bureau of Public Enterprises*,[64] the Nigerian Supreme Court dismissed the notion that the National Industrial Court of Nigeria had been elevated to the rank of a superior court of record by virtue of the Trade Disputes (Amendment) Act of 1992 and the National Industrial Court Act of 2006. The court unequivocally held that without an amendment of the provisions of Section 6(3) and (5) of the Constitution of the Federal Republic of Nigeria 1999, the National Industrial Court was not a superior court and was inferior in status to the High Courts. The Supreme Court stated,

> "It means therefore, that by Decree 47 of 1992 arrogating to the National Industrial Court a Superior Court of record as has been contended by the appellants does not by that token make the said National Industrial Court a Court of Superior record without due regard to amendment of the provisions of Sections 6(3) and (5) of the 1999 Constitution which has listed the only Superior Courts of record recognized and known to the 1999 Constitution and the list does not include the National Industrial Court; until the Constitution is amended, it remains a subordinate court to the High Court".[65]

The intervention of several courts by their refusal to bow to the legislative arm of government and correspondingly repudiating the purported exclusive jurisdiction of the National Industrial Court or its promotion to the rank of superior court by three different legislations over three decades,[66] was a herculean task compared to simply bestowing recognition of the new elevated status on the National Industrial Court and acknowledging its new exclusive jurisdiction in trade disputes and employment matters. But difficult or otherwise, the courts stuck to their guns and invalidated any provisions that were in conflict with the Constitution. The National Assembly had to step up and finally amend the Nigerian Constitution to ensure that the National Industrial Court was elevated to a superior court and given exclusive jurisdiction over labour matters via the Third Alteration Act 2010, instead of continuing to attempt to do so through another

[64] [2010] 3 SCM. 165 - 167.

[65] Per Chukwuma-Eneh, JSC.

[66] The first legislation was in 1976 with the Trade Disputes Act, then came the Trade Disputes (Amendment) Act in 1992 and then the National Industrial Court Act of 2006.

regular Act or legislation.[67] The Nigerian courts have engaged in judicial activism almost since the independence of Nigeria as a sovereign State and have given decisions that have bucked the *status quo* in a plethora of cases, including, *Western Steel Workers Ltd v Iron and Steel Workers Union of Nigeria (No. 2)*,[68] *Akintola v Adegbenro*[69] and *Arthur Nwankwo v The State*.[70]

The powers vested in the courts in Nigeria mean that they can *suo motu* raise (if neither party in a case does) the issue of the constitutionality of the provisions of the Nigerian Evidence Act that restrict the spousal privilege availed by the Act to a select married section of the Nigerian populace. In Nigeria, a court has the powers to raise an issue *suo motu* where the justice of the case demands it. In *Angadi v PDP & Others*,[71] the Supreme Court stated:

> "The issue of whether the trial Court below was right in considering processes which they had not been addressed on processes filed before it. This Court has held particularly in *Gbagbarigha v Toruemi* (2013) 6 NWLR (Pt.1350) 289 at 310, paragraphs C-G as follows: 'When a Judge raises an issue on his own motion, or raises an issue not in contemplation of the parties; or an issue not before the Court, the Court is said to have raised the issue *suo motu*".[72]

The court in *Angadi* stated further that when an issue is raised *suo motu* the parties should be heard before a decision is reached on the issue. The court however provided the exceptions to the rule requiring parties to be heard in instances where the court had raised an issue *suo motu*. The exceptions are: "(1) when the issue relates to the Court's own jurisdiction; (2) when both parties are not aware or ignored a statute which may have bearing on the case; or (3) when on the face of the record serious questions of the fairness of the proceedings is evident".[73]

The raising of the constitutional issue might be done by the court because the provisions of the Evidence Act (via its restriction of marriage to Nigerian Islamic or customary law and the adoption of the validity of marriage yardstick

[67] Alero E. Akeredolu and David Eyongndi, 'The Exclusivity of the Jurisdiction of the National Industrial Court Under the Nigerian Constitution Third Alteration Act and Selected Statutes: Any Usurpation?' (2019). Available at: <www.researchgate.net/publication/339335455_THE_EXCLUSIVITY_OF_THE_JURISDICTION_OF_THE_NATIONAL_INDUSTRIAL_COURT_UNDER_THE_NIGERIAN_CONSTITUTION_THIRD_ALTERATION_ACT_AND_SELECTED_STATUTES_ANY_USURPATION> accessed on 26/05/2021.

[68] [1987] 1 NWLR (Part 49) 284-303.

[69] [1963] 3 WLR 63 (PC).

[70] [1985] 6 NCLR. 228.

[71] [2018] LPELR-44375 (SC).

[72] Per Bage JSC, 30-31, paras D-E.

[73] ibid.

set out by the Nigerian Marriage Act) is discriminatory. In constitutional law, discrimination is

> "[t]he effect of a statute which confers particular privileges on a class arbitrarily selected from a large number of persons, all of whom stand in the same relation to the privileges granted and between whom and those not favored no reasonable distinction can be found".[74]

Chapter IV of the Nigerian Constitution enshrines the fundamental rights guaranteed by the Constitution and under that chapter, Section 42 deals with a person's right to freedom from discrimination and provides that a Nigerian citizen belonging to any community, ethnicity, sex, religion, place of origin or political opinion shall not be subjected either expressly by or in the practical application of any law in force in Nigeria, or any executive or administrative action of the government to disabilities or restrictions to which other citizens of Nigeria are not made subject to or be accorded any privilege or advantage which is not accorded to other Nigerian citizens.

Upon going through the provisions of Section 42 of the Constitution, it is evident to any reader that intendment and application of Section 42 of the Constitution are exact and unambiguous. The constitution enshrines the right of every Nigerian citizen to be free from discrimination whether it be in the form of *any law* in force in Nigeria or as a result of any *executive* or *administrative action*. Any law or executive/administrative action that places impediments which discriminate against any citizen or group of citizens which other citizens or groups are not subject to is illegal and unconstitutional.

Section 1(1) of the Constitution proclaims the supremacy of the Constitution over all persons and authorities in Nigeria. Section 1(3) of the Constitution goes further and states, "If any other Law is inconsistent with the provisions of this Constitution, this Constitution shall prevail, and that other Law shall to the extent of the inconsistency be void". The supremacy of the Constitution over all other laws in Nigeria is well settled and the courts have upheld this constitutional supremacy by striking down provisions of laws which have contravened the Constitution and this can be observed from the decisions in the cases of *Inspector General of Police v All Nigeria Peoples Party*[75] and *Inakoju & Ors. v Adeleke & Ors*.[76] Furthermore, the case of *Attorney-General of Lagos State v Attorney-General of the Federation*,[77] which deals with the constitutionality of an executive/administrative action carried out in contravention

[74] Black (n 5) 553.
[75] (2007) 18 NWLR (Pt.1066) 457.
[76] (2007) 4 NWLR (Pt. 1035) 403.
[77] (2004) 18 NWLR (PT.904) 1.

of explicit constitutional provisions, shows how the courts have ruled with regard to the issue of the supremacy of the constitution over such actions. In this case, the Supreme Court of Nigeria sitting in its original jurisdictional capacity on a dispute instituted before it by Lagos State against the Federal Government of Nigeria, ruled that the actions of the president in which he withheld the federal allocation payable to Lagos state were in contravention of the provisions of Section 162(5) of the Constitution and were therefore unconstitutional, null and void.

From the foregoing, it is clear that any Nigerian court, when presented with a case in which the spouse of an accused person on trial is being compelled to testify because they are not shielded by the provisions of Section 182(2) or (3) of the Evidence Act, can raise the issue of the constitutionality of those provisions as they discriminate against some Nigerian citizens which is prohibited by the Constitution, the apex statutory instrument and *grundnorm* of the land, or raise the issue of a serious question of the fairness of the proceeding before it.

In addition to a court *suo motu* raising the fairness or constitutionality of the provisions of Section 182(2) and (3) of the Evidence Act, any person who feels he/she been discriminated against can raise that issue or apply to a high court[78] to enforce his/her constitutionally guaranteed fundamental right.[79] There have been a plethora of cases in which actions have been filed by people seeking for the enforcement of their fundamental rights which have been contravened.[80] It is imperative to point out here that there is a snag that could be pointed out. This snag is the fact that the prohibition of discrimination provision under Section 42(1) of the Constitution as a fundamental right is only guaranteed to Nigerian citizens and thus, persons who are non-Nigerian citizens – even if they are long-term residents – are not within the contemplation of that provision from the wordings of that section. This means that with the high level of movement and migration in the 21st Century, a considerable amount of people are in jeopardy of being negatively impacted by the provisions of the Evidence Act referred to above. However, such persons might succeed in raising, before the court, the issue of the fairness of the provisions of Section 182(2) and (3) of the Evidence Act and getting the court to invalidate those provisions on that ground.

Another option which could do away with the problem posed by the aforementioned discriminatory provision of the Evidence Act is for the Nigerian National Assembly to exercise its powers under Section 4 of the Constitution and amend the provision of Section 258(1) of the Act which restricts the definition of a husband and wife to a husband and wife of a marriage contracted under the

[78] Either a Federal High Court or the High Court of a state in Nigeria.

[79] Section 46(1) & (2) of the Constitution of the Federal Republic of Nigeria 1999 (as amended).

[80] E.g. *Emmanuel Ejiogu Onuhikemi v Smridu Nigeria Limited* (Suit NICN/LA/265/2015), wherein the Plaintiff's action was for discrimination (among other claims).

Marriage Act or under Nigerian Islamic or customary law. The amendment by the National Assembly should define the terms husband and wife as the husband and wife of a marriage validly contracted under the laws of the place where the marriage was celebrated. This amendment will cure the mischief and injustice worked under the Evidence Act and will negate any need for judicial activism on the part of any court or even do away with the necessity for any person seeking to enforce his/her fundamental right against discrimination or even challenging the justice and fairness of offending provisions. This seems the progressive and less disruptive track to follow, as it will equiponderate the rights of all married persons who have validly entered into their marriage, irrespective of the place their marriage was contracted and thus discriminate against no spouse.

VI. Conclusion

The importance of the Evidence Act in Nigerian litigation cannot be overemphasised as it is the primary legislation that governs the taking of evidence in criminal proceedings in all the courts in Nigeria and civil proceedings in magistrate courts and most superior courts of record in the country.[81] It is therefore crucial that the legislation that dictates the rules to be followed in judicial proceedings must be logical, fair and unbiased.

The provisions of Section 182(2) and (3) of the Evidence Act, which deny some married people the benefit of a privileged protection on their marital communications and having their spouses cocooned by non-compellability, are discriminatory as they currently stand whilst linked to the interpretation section[82] of the Act (which is also tied to the Marriage Act). The status quo under the Evidence Act 2011 has revealed that there is a need to overhaul the provisions of the Act which have been shown to be discriminatory against some married people and usher in an era of fairness and equality. As has been recommended above, the easiest and fastest way to ensure equality would be for the National Assembly to amend the provisions of the Evidence Act which defines what a husband and a wife is. Until that is done, judicial activism, raising a serious question of the fairness of the trial and the enforcement of their fundamental right from discrimination might be the available options open to those persons who stand accused of a crime and who are adversely affected by the privilege and compellability provisions of the Evidence Act.

[81] By virtue of Section 256(1) of the Evidence Act.
[82] Section 258(1) of the Evidence Act.

Cambridge Law Review (2021) Vol VI, Issue ii, 24–61

The Inevitable Inconsistency of the Death Penalty in India

Pranav Verma *

Abstract

This article is an attempt to develop the theoretical framework of 'inevitable inconsistency' to re-evaluate the desirability of the death penalty in India, and propose it as a *sui generis* ground for revisiting its constitutionality. It highlights new and robust empirical research on the administration of the death penalty including: (a) sentencing trends followed by the trial courts; (b) the judicial attitudes of judges in the Supreme Court and; (c) the disparate impact of the death sentence on the socially and economically marginalized sections of the society. These research studies dictate significant deviations from the sentencing framework evolved by the Indian Supreme Court to administer the death sentence. To establish 'inevitable inconsistency', the article ventures into hypothesizing a 'best-case scenario' that removes such deviations by infusing consistency and fairness into death sentencing, to the maximum extent possible. It then highlights how even such 'best-case scenario' fails to prevent inconsistencies or arbitrariness at a magnitude not acceptable in a rule-based system. This failure of the 'best-case scenario' leaves our search for consistency with only two options – either a mandatory death sentence, or no death sentence at all. Given, that the former already stands declared unconstitutional in India, abolition emerges as the only viable end to the search of consistency that this article embarks upon. Accordingly, the article argues that the 'inevitable

* Assistant Professor, NALSAR University of Law, Hyderabad, India. B.A., L.L.B. (Hons.) (NALSAR), LLM (Cantab). I gratefully acknowledge the helpful comments received from Prof. Peter Cane, University of Cambridge on an earlier draft of this article. Any errors that remain are my own. pranavverma@nalsar.ac.in

inconsistency' framework, when coupled with the unique, irrevocable nature of the death sentence, must inform the constitutional inquiry into its retention in law.

Keywords: death penalty, inevitable inconsistency, best-case scenario, Bachan Singh, arbitrariness.

I. Introduction

The year 2021 marks forty-one years since the Supreme Court of India, by a majority of four to one (with Bhagwati J dissenting), upheld the constitutionality of the death penalty in *Bachan Singh v State of Punjab* (or 'in *Bachan Singh*').[1] The decision in *Bachan Singh* was the first attempt by the Indian Supreme Court to limit the application of the death penalty to only the most extreme cases, and to that end, the Court laid down a comprehensive sentencing framework for future courts administering the death sentence. The Court's decision in *Bachan Singh* was premised upon the perceived utility of the death penalty, but at the same time acknowledged the need to limit its application and infuse consistency into its sentencing through a sentencing framework.

I aim to highlight how the actual practice of death sentencing complies with the *Bachan Singh* sentencing framework. Instead of making an argument on abolition from the very start, I first highlight the existing gaps between the theory and the actual practice of death sentencing in India. I then evaluate how the identified gaps between the theory and practice can be possibly removed. This would involve an attempt at discovering a system where death sentence is carried out in as non-arbitrary a manner as possible. This would be followed by a critical evaluation of this supposed 'best-case scenario', and I subsequently propose to see if such scenario too leaves rooms for errors which would be unacceptable in any rule-based system. This search for non-arbitrariness (which I take to mean consistency in this context), then, would lead to only two possible alternatives, theoretically speaking – a mandatory death sentence (already abolished in India) or no death sentence at all.

This trajectory becomes worth exploring because *Bachan Singh* upheld the death penalty as it was convinced of it being administered consistently under the sentencing framework evolved by it. To not just highlight the failure of that framework but to also better upon it and still conclude the practical impossibility of a consistent application of the death penalty, even in such a 'best case scenario', would provide a stronger case for abolition and a more persuasive argument for retentionists to engage with. This shows promise to lead to a conclusion

[1] (1980) 2 SCC 684.

of 'inevitable inconsistency' which can be constructed as a larger theoretical framework to evaluate the desirability of the death penalty.

II. Legislative And Judicial History

The legislative and judicial history of death sentencing in India can be traced back to the Code of Criminal Procedure, 1898 (or 'the 1898 Code') applicable to British India. Section 367(5) thereof provided for death sentence as the default punishment for murder and required the sentencing court to provide written reasons if it were to choose a sentence of life over death.[2] This was considered to mean that the legislature regarded the death penalty to be the norm and life imprisonment to be an exception.[3] This was followed by an amendment to the 1898 Code in the year 1955 which deleted Section 367(5) altogether, and a new sub-section on an entirely different subject-area was substituted in its place.[4] It has been argued that this move reflected a shift towards the death penalty no longer being the 'norm' or the default punishment for murder,[5] or at the very least, a lack of any legislative preference between the two sentences of life and death.[6]

It was under this legislative framework that the first constitutional challenge to the death penalty was repelled by the Indian Supreme Court in the case of *Jagmohan*.[7] The Court recognised that the 1955 Amendment left it to the discretion of a sentencing judge to choose between life and death, but felt that any erroneous exercise of such discretion could be remedied within the appellate and procedural frameworks available under the existing law.[8] However, this decision soon came under reconsideration because of a significant legislative change.

In 1973, a new Code of Criminal Procedure (or 'the 1973 Code') was enacted. It repealed the 1898 Code and currently governs criminal procedure in India. Section 354(3) of the 1973 Code introduced the requirement of "special reasons" to be recorded by a sentencing court for choosing the death sentence over life imprisonment or imprisonment for a term of years, if the offence provides for

2 Code of Criminal Procedure, 1898 [Repealed], section 367(5).

3 *Bachan Singh v State of Punjab* (1982) 3 SCC 24 [20] (Bachan Singh Dissenting Opinion); Law Commission of India, 'The Death Penalty' (Government of India 2015) 262 para 2.2.5 <https://lawcommissionofindia.nic.in/reports/Report262.pdf> accessed 9 February 2021; Project 39A, 'Death Penalty Sentencing in Trial Courts' (National Law University, Delhi 2020) <https://static1.squarespace.com/static/5a843a9a9f07f5ccd61685f3/t/5ebc3dc0879c75754ab23f78/1589399023 71/Death+Penalty+Sentencing+in+Trial+Courts.pdf> accessed 10 February 2021, page 13;.

4 Code of Criminal Procedure (Amendment) Act, 1955 [Repealed], s 66.

5 Law Commission of India (n 3) para 2.2.3.

6 Project 39A (n 3) 13.

7 *Jagmohan v State of UP* (1973) 1 SCC 20.

8 ibid [26].

these alternative sentences.[9] This requirement of "special reasons" for choosing the death sentence was considered to reflect a new legislative policy: life imprisonment was now the default penalty for murder, and death sentence an exception.[10] The 1973 Code, therefore, ushered in a complete reversal of the position under the 1898 Code.

The scope of the expression "special reasons" under Section 354(3) first came to be considered by the Supreme Court in *Rajendra Prasad*.[11] The Court held that "special reasons" meant that the courts must consider only the circumstances of the criminal and not the crime (i.e., only mitigating circumstances, and not the aggravating ones) in choosing whether to apply the death sentence.[12] This was overruled when a new challenge was mounted to the constitutionality of the death penalty in *Bachan Singh*. While upholding the death sentence, the Court held that both the circumstances particular to the crime as well as the criminal must be weighed against each other;[13] and this set the groundwork for weighing both aggravating and mitigating circumstances.[14] More importantly, the Court acknowledged that the import of requiring "special reasons" was that it was the legislative policy to limit the application of the death sentence only to "exceptionally grave" or "extreme" cases.[15] To guide the sentencing courts in so administering the death penalty, the Court made the first serious attempt in the Indian death sentencing jurisprudence at laying down a comprehensive sentencing framework for future courts. This framework is discussed in some detail in Section III of the article.

However, before proceeding to what the framework lays down, a quick mention may be made of the third and final challenge to the constitutionality of the death penalty attempted in the case of *Shashi Nayar*,[16] where the petitioner requested a reconsideration of *Bachan Singh*. The Supreme Court, however, remained unmoved. It took 'judicial notice' of the worsening law and order situation in the country and simply declared that, "[t]he death penalty has a deterrent effect and it does serve a social purpose".[17] The final constitutional challenge to the death

9 Code of Criminal Procedure, 1973 (CrPC 1973), s 354 (3).
10 *Bachan Singh* (n 1) [209].
11 *Rajendra Prasad v State of Uttar Pradesh* (1979) 3 SCC 646.
12 ibid [88(9)].
13 *Bachan Singh* (n 1) [161].
14 ibid.
15 ibid.
16 *Shashi Nayar v Union of India* (1992) 1 SCC 96.
17 ibid [99].

penalty having thus been repelled, the sentencing framework laid down in *Bachan Singh* achieved finality.

III. The Bachan Singh Sentencing Framework

The Court in *Bachan Singh* provided a detailed sentencing framework to guide judicial discretion in choosing the death sentence, consistent with the legislative policy indicated in Section 354(3) of the 1973 Code.[18] Integral to this sentencing framework is the weighing of aggravating factors (relating to the crime) against the mitigating factors (relating to the criminal). In fact, the Court provided an indicative list of aggravating and mitigating factors, designedly kept non-exhaustive, so as "[n]ot to fetter judicial discretion".[19]

The aggravating factors included:

a) [I]f the murder has been committed after previous planning and involves extreme brutality; or

b) If the murder involves exceptional depravity; or

c) If the murder is of a member of any of the armed forces of the Union or of a member of any police force or of any public servant and was committed –

   i. while such member or public servant was on duty; or

   ii. in consequence of anything done or attempted to be done by such member or public servant in the lawful discharge of his duty as such member or public servant, as the case may be, or had ceased to be such member or public servant; or

d) If the murder is of a person who had acted in the lawful discharge of his duty under Section 43 of the Code of Criminal Procedure, 1973, or who had rendered assistance to a magistrate or a police officer demanding his aid or requiring his assistance under Section 37 and Section 129 of the said Code.[20]

Likewise, the Court also laid down an illustrative, non-exhaustive

[18] *Bachan Singh* (n 1) [177].
[19] ibid [203].
[20] ibid [202].

list of the following mitigating factors:

a) [T]hat the offence was committed under the influence of extreme mental or emotional disturbance.

b) The age of the accused. If the accused is young or old, he shall not be sentenced to death.

c) The probability that the accused would not commit criminal acts of violence as would constitute a continuing threat to society.

d) The probability that the accused can be reformed or rehabilitated. The State shall by evidence provide that the accused does not satisfy conditions (c) and (d) above.

e) That in the facts and circumstances of the case the accused believed that he was morally justified in committing the offence.

f) That the accused acted under duress or domination of another person.

g) That the condition of the accused showed that he was mentally defective and that the said defect impaired his capacity to appreciate the criminality of his conduct.[21]

It is imperative to note here the Court's crucial direction to sentencing judges to ensure that mitigating factors "[r]eceive a liberal and expansive construction".[22] Notably, no such prescription was made for aggravating factors.[23] Finally, in the event of aggravating factors outweighing mitigating factors, despite the latter being constructed 'liberally and expansively', the Court provided for a final leg of enquiry. This was the determination into whether all the sentencing options, alternative to death, stood 'unquestionably foreclosed'. If the answer to this final enquiry were to be in the affirmative, then the case was to fall under the category of "rarest of rare", meriting the award of the death sentence.[24]

It would thus seem that the sentencing framework evolved in *Bachan Singh* seemingly sets an extremely high threshold for applying the death penalty. Not

[21] ibid [206].
[22] ibid [209].
[23] Project 39A (n 3) 15.
[24] *Bachan Singh* (n 1) [209].

only are aggravating factors to be weighed against mitigating ones, but the latter are to be constructed liberally and expansively – no similar requirement has been laid down for aggravating factors.[25] Furthermore, if despite such exercise the aggravating factors outweigh mitigating factors, the Court must satisfy itself, in the final leg of the enquiry, that all sentencing options other than death, stand 'unquestionably foreclosed'.

Furthermore, subsequent decisions of the Supreme Court have strengthened the *Bachan Singh* framework by expanding the choices of alternative sentencing options available to a court, that are required to be 'unquestionably foreclosed' under the final enquiry. This has further raised the threshold of choosing a death sentence in a given case.[26] These decisions come in the wake of the real problem that under the statutory scheme of the 1973 Code, life prisoners become eligible for remission after the completion of a fourteen-year term,[27] with the practical consequence that often a sentence of life is reduced only to a sentence of fourteen years.[28] Therefore, while a case may be grave enough to merit imprisonment beyond fourteen years, it may yet not be grave enough to merit the death sentence, leaving a court with real-time sentencing dilemmas in such situations.

The Supreme Court resolved this dilemma in *Sriharan*,[29] wherein by a three to two majority, it affirmed its previous decision in *Swamy Shraddananda*[30] and clarified that where death sentence is one of the statutorily prescribed punishments, it is open to the High Courts and the Supreme Court to consider the following two alternative sentences:

1) a life sentence specified to extend till the rest of the entire natural life of the convict, without any possibility of remission; and

2) If a life sentence without remission appears to be disproportionately high, and yet a life sentence reduced to fourteen years after remission appears to be grossly inadequate, as a "middle ground"31 the Court can specify a fixed term imprisonment exceeding fourteen years but falling short of life (say for twenty, thirty, forty years, and so on), and

[25] Project 39A (n 3) 15.

[26] Centre on the Death Penalty, 'Litigating Death Penalty Cases: A Consultation' <https://criminallawstudiesnluj.files.wordpress.com/2020/08/4b439-issuesinlitigatingdeathpenaltycases.pdf> accessed 10 February 2021.

[27] CrPC 1973, s 433A.

[28] *Swamy Shraddananda v State of Karnataka* (2008) 13 SCC 767 [89].

[29] *Union of India v Sriharan* (2016) 7 SCC 1.

[30] *Swamy Shraddananda* (n 28).

[31] Project 39A (n 3) 20.

place it beyond the statutory remission.[32]

Accordingly, as alternatives to the death sentence, the High Courts or the Supreme Court may, depending upon the circumstances of the case, choose to instead impose either a life sentence beyond the scope of remissions; or a fixed-term sentence of choice between fourteen years and life, again placed beyond the scope of remissions. These two sentences further expand the scope of the 'alternative options' that are required to be 'unquestionably foreclosed' under *Bachan Singh*'s final leg of enquiry, before a sentencing court could proceed to choose the death sentence.[33]

It may also be stressed here that even where life or fixed-term imprisonment is under consideration, it may further be of two descriptions – simple, rigorous, or a combination of both. This provides additional flexibility in sentencing.[34] This has the effect of further increasing the already high threshold of the *Bachan Singh* sentencing framework by expanding the scope of the alternative sentences that must be 'unquestionably foreclosed' before the death sentence can be pronounced.

The *Bachan Singh* sentencing framework may, therefore, be encapsulated in the following flowchart:

IV. The Deviations From The Bachan Singh Sentencing Framework

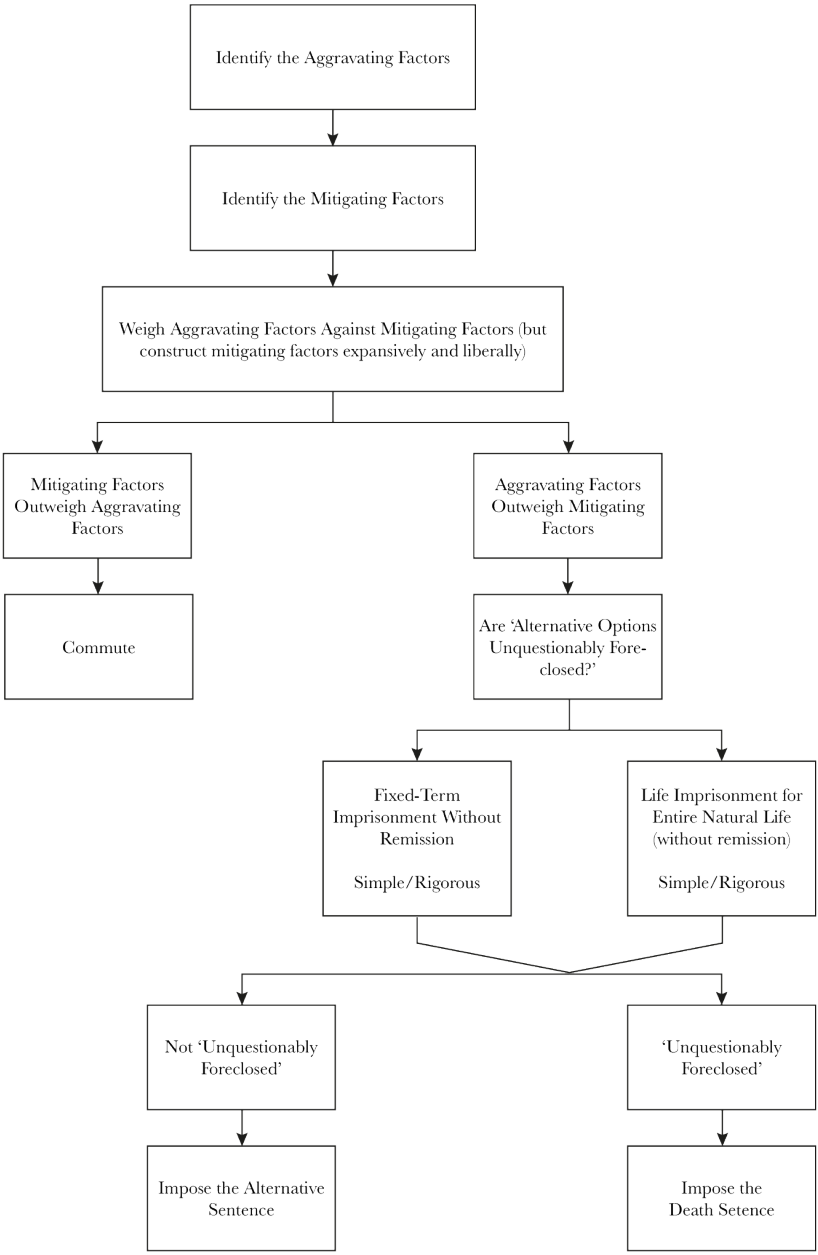
This part of the article aims to evaluate the actual application of the *Bachan Singh* sentencing framework, and analyse how such application measures against the framework's high threshold for choosing the death sentence. The specific deviations and gaps between the theory and the practice of the sentencing framework are discussed, and this analysis is supported by four primary sources of empirical evidence. The first is sentencing data from the trial courts of three major Indian states which administer the highest number of death sentences (a total of two-hundred and fifteen trial court decisions where death sentence was awarded), published in the year 2020.[35] The outcomes of this study record several deviations

[32] *Sriharan* (n 29) [106], [114]; Project 39A (n 3) 20. The constitutional powers of pardon by Governors and the President would still apply.

[33] Project 39A (n 3) 20.

[34] Indian Penal Code, 1860, section 53.

[35] Project 39A (n 3) 10.


Identify the Aggravating Factors
Identify the Mitigating Factors
Weigh Aggravating Factors Against Mitigating Factors (but construct mitigating factors expansively and liberally)
Mitigating Factors Outweigh Aggravating Factors
Aggravating Factors Outweigh Mitigating Factors
Commute
Are 'Alternative Options Unquestionably Fore-closed?'
Fixed-Term Imprisonment Without Remission
Simple/Rigorous
Life Imprisonment for Entire Natural Life (without remission)
Simple/Rigorous
Not 'Unquestionably Foreclosed'
'Unquestionably Foreclosed'
Impose the Alternative Sentence
Impose the Death Setence

from the *Bachan Singh* framework, which I report and analyse in this part of the article, and build upon to formulate a best-case scenario in Section V of this article.

The second source is the opinion study of sixty former justices of the Indian Supreme Court on their experiences in applying the death penalty.[36] Published in 2018, this is a particularly instructive source given that of the sixty judges interviewed, forty-seven had adjudicated sixty-three death penalty cases and confirmed a total of ninety-two death sentences.[37] The third and fourth sources, respectively, include the Death Penalty India Report (published in 2016);[38] and the Law Commission of India Report No. 262 (published in 2015)[39] which recommended abolition of the death penalty except for terrorism-related offences. These two form the final section of this part of the article, and reveal the consequences of the specific deviations from the *Bachan Singh* framework. These consequences take the shape of rampant inconsistencies in applying the death penalty, with a disproportionate impact upon the vulnerable sections of the society.

These four sources collectively reveal the following instances of deviations from the sentencing framework evolved in *Bachan Singh*.

A. Introduction of 'collective conscience' and crime-categories

Over the past several years, a new judicially-evolved test of the crime 'shocking the collective conscience of the society', has emerged as a significant basis for choosing the death sentence. This part explores how the new test of 'collective conscience', despite falling squarely outside the *Bachan Singh* framework, has found acceptance by, both, the Supreme Court and trial courts. Before proceeding into how this has been done, it is worthwhile to note here that not only did the *Bachan Singh* sentencing framework never provide for the criteria of collective conscience, it had, in fact, explicitly guarded against it by forewarning that "[j]udges should not take upon themselves the responsibility of becoming oracles or spokesmen of public opinion".[40]

However, this clear prohibition against courts interpreting community standards while applying the death sentence, seems to have been disregarded,

[36] National Law University, Delhi, 'Matters of Judgment: A Judges' Opinion Study on the Death Penalty and the Criminal Justice System' *Issuu* <https://issuu.com/p39a/docs/combined231117> accessed 9 February 2021.

[37] ibid 14.

[38] Centre on the Death Penalty, 'Death Penalty India Report, Volume I' (National Law University, Delhi Press) <https://static1.squarespace.com/static/5a843a9a9f07f5ccd61685f3/t/5b68a29a6d-2a73cbec1ec89f/1533584200675/Vol.I_Death+Penalty+Report.pdf> accessed 9 February 2021.

[39] Law Commission of India (n 3).

[40] *Bachan Singh* (n 1) [126], [176].

when, just three years after the decision in *Bachan Singh*, a smaller Supreme Court bench in *Machhi Singh*[41] introduced the test of 'collective conscience of the society being shocked' as the basis of imposing the death penalty. This new element of 'shocking the collective conscience' was divided into five sub-categories[42] by the Court, indicating when it comes into play. These are: (a) manner of commission of the crime (which includes brutality); (b) motive of commission of murder; (c) anti-social or socially abhorrent nature of the crime; (d) magnitude of the crime; and (e) personality of victim of the murder.

A society's 'collective conscience is shocked', therefore, depending upon the role played by the above sub-categories. *Machhi Singh* then provided an illustrative list under each of the above categories to determine when these sub-categories can be said to 'shock the collective conscience'. A closer inspection, however, reveals that this illustrative list is, in turn, a compilation of several pre-determined categories of crime.[43] For example, the second sub-category of '*motive*' includes murders by hired assassins for the sake of rewards; murders designed to inherit property, *etc.* Similarly, the third sub-category of '*anti-social crimes*' includes murder of members of Scheduled Castes or minority communities; cases of dowry-deaths, *etc.* The fourth sub-category of '*magnitude*' includes murders of multiple members of a family or a large number of persons from one caste community. The final sub-category of '*victim's personality*' includes murders of innocent children; of those rendered helpless by age, *etc.*[44] Accordingly, the resultant effect of the 'collective conscience' framework is existence of pre-set crime-categories which invite the court to impose the death sentence. For instance, if X is a crime-category that has been identified as a shock to the collective conscience of the society, then an accused A convicted of X would deserve the death sentence.

This manner of sentencing falls foul of the comprehensive sentencing framework laid down in *Bachan Singh*, which required a careful weighing of aggravating and mitigating factors unique to the crime and the criminal. It replaces such individualised sentencing with a straight-jacket factual enquiry into the existence of pre-identified crime-categories. For instance, in Madhya Pradesh[45], trial courts were found to have applied the *Machhi Singh* crime-categories as a matter of "literal adherence"[46], i.e. "[s]imilarity was drawn between the case before the court and *Machhi Singh*'s five categories. Using the circumstances of the

[41] *Machhi Singh v State of Punjab* [1983] SCC 470.
[42] ibid [33]–[37].
[43] Project 39A (n 3) 16; *Machhi Singh* (n 41) [33]-[37].
[44] *Machhi Singh* (n 41) [33]-[37].
[45] Project 39A (n 3) 41.
[46] ibid.

crime, compliance with either all or one of the categories was shown to impose the death sentence".[47]

Even generally, *Machhi Singh*'s formulation of 'collective conscience' has had a lasting legacy on how the death sentence is applied. 'Collective conscience' was found to have been invoked by the trial courts in 72% of the cases in Delhi, 43% of the cases in Madhya Pradesh, and 51% of the cases in Maharashtra.[48] Moreover, in the cases in which 'collective conscience' was invoked, it seemed to have played a determinative role in deciding the sentence.[49]

In an even bigger deviation from the *Bachan Singh* sentencing framework, the existence of 'collective conscience being shocked' has often been the cause of complete non-consideration of mitigating factors. For instance, in all the one-hundred and twelve cases in which collective conscience was a factor influencing death sentencing, no mitigating factors were considered at all in sixty-three of those cases (i.e., in roughly 56% of the cases).[50] This is a disturbing deviation from the *Bachan Singh* sentencing framework, which not only mandates weighing of both aggravating and mitigating factors, but also requires mitigating factors to be constructed expansively and liberally.

Although a perversion of the *Bachan Singh* framework, factors like 'collective conscience' and crime-categories laid down by *Machhi Singh*, have been found to hold considerable sway over judicial attitudes towards the death penalty. For instance, the opinion study of sixty former justices of the Indian Supreme Court on their experiences in applying the death sentence[51] starkly reveals judicial acceptance of the crime-categories approach, although it falls outside the sentencing framework in *Bachan Singh*. No less than thirteen former judges of the Indian Supreme Court who decided eighty death penalty appeals between them and confirmed forty-one death sentences, endorsed the 'crime-categories' approach, leaving hardly any room for mitigating circumstances.[52] Thus, for a significant number of the judges interviewed, the *Bachan Singh* sentencing framework was often understood as reducible to crime-categories or description of offences alone.[53]

Further, eleven former judges of the Court in their responses to the opinion study, considered 'collective conscience' to be a crucial aggravating factor, revealing the entrenchment of the 'collective conscience' framework in judicial

47 ibid.
48 ibid 33, 75.
49 ibid 32.
50 ibid 33, 75.
51 National Law University, Delhi (n 36).
52 ibid 59.
53 ibid 17.

attitudes towards death sentencing. These eleven judges confirmed forty-one death sentences by invoking 'collective conscience'.[54]

Therefore, *Machhi Singh*'s legacy of 'collective conscience' clearly continues to influence death sentencing in India. This repeated use of 'collective conscience', however, should not obscure this criteria's explicit exclusion, and lack of fit, in respect of the *Bachan Singh* sentencing framework which requires judges to consider only the circumstances unique to the crime and the criminal.[55] The gradual incorporation of 'collective conscience' into the Indian death penalty jurisprudence is, therefore, a significant deviation from the *Bachan Singh* sentencing framework and the principle of individualised sentencing which lies at its heart.

B. The dominance of 'brutality'

Another legacy of *Machhi Singh* is the central place given to the aggravating factor of brutality of a crime in death sentencing by trial courts. *Machhi Singh* listed brutality of the crime as the first sub-category indicative of when a society's collective conscience is shocked, and this factor has often received a central place of analysis in death sentencing, often trumping other considerations.

This part of the article highlights how brutal manner of commission of the crime has emerged as one of the main aggravating factors on the basis of which mitigating factors have been regularly dismissed – both by the Supreme Court and the trial courts. This again contravenes the *Bachan Singh* sentencing framework which did not prescribe any inherent hierarchy between aggravating and mitigating factors, except for the latter which were to be expansively and liberally constructed. It is worth remembering that the involvement of extreme brutality was just one of the many aggravating factors listed out by the Court in *Bachan Singh*. No particular aggravating factor was accorded a trumping value over a mitigating factor, and yet, that is precisely how the aggravating factor of brutality seems to have been consistently applied.

First, the influence of brutality as a key aggravating factor, on judicial attitudes within the Supreme Court, is revealed by the opinion study of the former Supreme Court judges:

> "[B]rutality of the crime emerged as a dominant theme in discussions on aggravation. For 21 judges, the nature of the crime, or the manner of its commission, were not just aggravating factors, but bordered on being determinative of the question whether the accused deserved to be sentenced to death. Additionally, for an

[54] National Law University, Delhi (n 36) 67.
[55] Project 39A (n 3) 75.

> almost equal number, the brutality of the crime weighed very heavily in the balancing between aggravating and mitigating factors. One judge who decided nearly 130 murder cases as an appellate court judge said, 'The heinous nature of the crime certainly colours our judgement".[56]

Not only does this reveal a judicial psyche that views brutality as a key basis behind choosing the death sentence, but it also reveals an attitude of suspicion towards mitigating factors in general. For instance, six of these former Supreme Court judges were reported to doubt the very concept of mitigation in cases involving extreme brutality, and for such cases, felt that the concept of mitigation itself was "irrelevant".[57] These judges often came to view mitigating circumstances as an "excuse" for the crime,[58] forgetting that "[m]itigation is not meant to have any effect on the guilt, but is instead meant to act as a tool to individualise punishment".[59]

Second, not only has this been a view influencing judicial attitudes of India's top court, but it has also 'trickled down' to sentencing courts applying the death sentence in the first instance. The data from trial courts suggests that brutality was the "[m]ost commonly argued and considered aggravating factor", having been raised in one-hundred and thirteen of the two-hundred and fifteen cases in which death sentence was awarded in the states of Delhi, Madhya Pradesh, and Maharashtra between 2000-2015 (i.e., in roughly 53% of the cases).[60] The data from the states of Maharashtra and Madhya Pradesh, in particular, suggests that the basis of awarding the death penalty has often been a trumping enquiry into the brutality of the crime, on which basis alone courts have been categorising a case as "rarest of rare".[61]

This trend of dismissing mitigating factors simply in face of one aggravating factor of brutal manner of commission of the crime, squarely falls outside the *Bachan Singh* sentencing framework. It incorrectly attaches a trumping value to one aggravating factor over all other mitigating factors, and also fails to construct each mitigating factor 'expansively and liberally', as explicitly directed by *Bachan Singh*. These decisions do not even venture into the final leg of enquiry into the alternative options being "unquestionable foreclosed". It appears that the entirety

[56] National Law University, Delhi (n 36) 67.
[57] ibid 71.
[58] ibid.
[59] ibid.
[60] Project 39A (n 3) 31.
[61] ibid 27.

of the *Bachan Singh* sentencing framework has been reduced to a single-point enquiry into the brutal manner of commission of the crime.

It must be clarified here that the argument is not that brutality cannot be a reasonable indicator in death sentencing or that it is inherently problematic, but rather that it should not be the sole indicator, that leads to a complete non-consideration of mitigation and a breakdown of the principle of individualised sentencing. The discussion above, however, reveals that sentencing judges deploy brutality to dismiss all mitigating factors, *en masse*, in a large number of cases as their default approach. This trumping value attached to brutality turns death sentencing into a single-point enquiry into the brutal manner of commission of the crime. Therefore, not only does it fail the test of the law as laid down by *Bachan Singh*, but also compromises the principle of individualised sentencing by failing to accord due weightage to mitigating factors.

Machhi Singh's sway over the death penalty jurisprudence is particularly puzzling as it seems to have substantially altered the *Bachan Singh* framework despite being a bench of a smaller strength of three judges, and being legally bound to operate within the confines of *Bachan Singh* – a Constitution Bench of five judges. *Machhi Singh*, in fact, seems to have acted in contravention of the well-settled principle that a decision delivered by a bench of a larger strength is binding on any subsequent bench of a lesser or co-equal strength.[62]

C. Inadequate consideration of mitigating factors

Due consideration of mitigating factors is integral to the sentencing framework evolved in *Bachan Singh*. Hence, it is the mitigating factors, and not the aggravating ones, that need to be constructed 'expansively and liberally'. Evidence from how trial courts impose the death sentence, however, presents a completely different picture.

The trial courts in India's three states with the highest statistics of imposing the death sentence, show a poor record of appreciating mitigating factors during sentencing hearings, relying mainly upon aggravating factors to impose the death sentence. The state that fares the worst in this regard is Madhya Pradesh where no mitigating factors were considered at all in fifty-one of the eighty-two judgements imposing the death penalty (i.e., in about a staggering 62% of cases) between 2000-2015. This is followed by Maharashtra, where forty-one out of the ninety cases showed that there was no appreciation of mitigating factors (i.e., in about 46% of the cases). Finally in Delhi, eighteen out of the forty-three cases failed to

[62] *Central Board of Dawoodi Bohra v State Of Maharashtra & Anr* (2005) 2 SCC 673; *Official Liquidator v Dayananad & Ors* (2008) 10 SCC 1; *Subhash Chandra & Anr v Delhi Subordinate Services Selection Board* (2009) 15 SCC 458.

consider any mitigating factors (i.e., about 42% of the cases).[63] This signals a near breakdown of the *Bachan Singh* sentencing framework.

What may, however, be said about the cases in which mitigating factors were actually considered? The trial courts' study does not necessarily paint an encouraging picture for such cases, either. The study finds that often where mitigating factors were pleaded to seek leniency in sentencing, this exercise was carried by defence counsels in an extremely perfunctory manner by a mere listing of the facts pertaining to the accused, without meaningfully contextualising these facts in the accused's psychological or social circumstances, or detailing how those mitigating factors actually shaped the personality of the accused or impacted their life-choices, to deserve a more lenient sentence.[64] This lack of contextualisation and mere listing of mitigating factors made it rather easy for trial courts to not duly consider them.[65] This practice was observed in forty-nine out of the ninety cases in Maharashtra (i.e., in about 54% of the cases); in twenty-five out of the forty-three cases in Delhi (i.e., in about 58% of the cases); and in thirty-one out of the eighty-two cases in Madhya Pradesh (i.e., in about 38% of the cases).[66]

For instance, in the cases where existence of dependents was argued as a mitigating factor, no arguments were advanced on either the nature of the relationship of such dependents to the accused or how they were dependent upon the accused for their survival.[67] This inadequate manner of presentation of mitigating factors severely prejudices the accused in the sentencing hearings. The perfunctory nature of such arguments does not "[c]arry weight as they do not enable the trial court to assess how the accused is placed within their family and how a death sentence would prejudice them through collateral or consequential damage"[68], and therefore, the particular mitigating factors fail to be accorded the weight they deserve.

The situation is even worse for cases involving multiple accused persons standing trial together in the same proceedings, as evidence suggests that individualized mitigating factors unique to each accused are hardly brought forth in these cases. Again, in the three states of Delhi, Madhya Pradesh, and Maharashtra, of the fifty-two cases in which multiple accused persons stood trial in the same proceedings, individual mitigating circumstances pertaining to each accused were argued in only nine cases, i.e., in only about 17% of the cases.[69] Further, where mitigating circumstances were presented for multiple accused persons, they were

63 Project 39A (n 3) 27, 70.
64 ibid 27.
65 ibid 26.
66 ibid 70.
67 ibid 26.
68 ibid.
69 ibid.

again done in a perfunctory manner without a meaningful contextualisation, just as in cases involving single accused persons.[70]

In figures that reveal the completely broken nature of mitigation hearings in Indian trial courts, where individualised mitigating factors were indeed considered for multiple accused persons standing trial together, this was done only in about 6% cases in Madhya Pradesh, 17% cases in Maharashtra, and 33% cases in Delhi.[71]

These figures conclusively prove that death sentencing in trial courts does not nearly accord the same value to mitigating circumstances that was intended by the framework evolved in *Bachan Singh* – which seems to be followed more in breach than in observance. However, these figures beg the question of why are mitigating factors not adequately presented before the courts in the first place? The answer seems to lie in the rampant practice of same-day sentencing. Section 235(2) of the 1973 Code bifurcates a criminal trial into two stages of conviction and sentencing, with separate oral arguments to be advanced for each stage.[72] The object of having an independent stage of sentencing hearing is to "[e]nable the court to have information relevant to arriving at a decision on the choice of the appropriate sentence"[73], as affirmed by the Supreme Court on multiple occasions, as well.[74] This information includes important mitigating factors that comprehensively bring into consideration the circumstances about characteristics and background of the offender.[75] The primary objective of independent sentence hearings is, therefore, to facilitate a comprehensive presentation of mitigating circumstances.

The quality of mitigation arguments, however, would be drastically reduced in cases of 'same-day sentencing' or passing of sentences on the same day when the conviction proceedings are concluded, as this leaves inadequate time for arguing on mitigating factors, let alone finding enough data to properly contextualise them in the personal circumstances of the accused. It has been found that in no less than 44% of the cases in Delhi, Madhya Pradesh, and Maharashtra, sentencing hearings took place on the same day as the pronouncement of guilt.[76] The practice of same-day sentencing was particularly rampant in Madhya Pradesh, where it was observed in 76.9% of the cases; and in Maharashtra, sentencing on the same

[70] ibid.

[71] ibid 75.

[72] CrPC 1973, s 235 (2).

[73] S Muralidhar, 'Hang Them Now, Hang Them Not: India's Travails With The Death Penalty' (1998) 40 Journal of the Indian Law Institute 143, 155.

[74] *Santa Singh vs State Of Punjab* (1976) 4 SCC 190; *Muniappan vs State Of Tamil Nadu on 18 March, 1981* (1981) 3 SCC 11; *Allauddin Mian & Ors v State Of Bihar* (1989) 3 SCC 5; *Malkiat Singh And Ors v State Of Punjab* (1991) 3 SCC 11.

[75] Muralidhar (n 73) 155.

[76] Project 39A (n 3) 26, 67.

day or with just a twenty-four-hour gap was observed in 57% of the cases.[77] In fact, the data from the trial courts establishes a direct correlation between same-day sentencing and non-consideration of mitigating circumstances, given that

> "[n]o mitigating circumstances were considered in 41 out of the 60 same-day sentencing cases in Madhya Pradesh and in 16 of the 31 same-day sentencing cases in Maharashtra. The two same-day sentencing cases in Delhi did not consider any mitigating circumstances either".[78]

The principle that becomes a casualty in same-day sentencings is that adequate time is vital to present the mitigating circumstances of an accused in a comprehensive manner. These concerns extend beyond the cases of same-day sentencing and into those where there is meaningfully very little gap between hearings on conviction and sentencing. Therefore, it has been found that even in cases where the gap between conviction and sentencing hearing was more than a day's, such time duration

> "[w]as not sufficient for an in-depth mitigation exercise. The median of duration between conviction and sentencing hearing across the three states was one day. It was 0, 2 and 7 days respectively for the state of Madhya Pradesh, Maharashtra, and Delhi".[79]

Given the central importance that the *Bachan Singh* framework attaches to mitigating factors, the practices of same-day sentencing, and insufficient time for arguments on mitigation have been criticised as the "most antithetical" to the *Bachan Singh* framework, for they prevent the mitigating factors relevant to the convict from being adequately presented before the sentencing court.[80] This has the effect of leaving out a key component of the *Bachan Singh* framework from even being considered during sentencing.

D. Summary dismissal of mitigating factors

As problematic as the inadequate consideration of mitigating circumstances is, there has been discerned an even more worrying trend of summary and whimsical dismissal of mitigating factors – an approach surprisingly legitimated by the Supreme Court itself. For instance, in the case of *Sevaka Perumal*, the Supreme

77 ibid 67.
78 ibid 69.
79 ibid.
80 ibid.

Court dismissed the mitigating factors of young age, and the convict being the 'sole bread winner' in the family, on the basis that such "[c]ompassionate grounds would always be present in most cases and are nor relevant for interference".[81] Similarly, in *Shimbhu*, the Court held that the punishment should always be proportionate to the gravity of the offence, and factors like religion, race, caste, socio-economic status of the accused "[c]annot be construed as special factors for reducing the sentence".[82] In an inexplicable development, the Court in *Krishnappa* made a sweeping declaration that "[s]ocio-economic status, religion, race, caste or creed of the accused or the victim are irrelevant considerations in sentencing policy".[83] This blanket dismissal of mitigating factors is a complete reversal of the *Bachan Singh* framework which not only required due weightage to be given to mitigating factors but also required them to be construed expansively and liberally.

Further, such summary dismissal of mitigating factors by the Supreme Court set a dangerous trend which came to be perpetuated by trial courts in the name of applying precedents.[84] The approach seems to have been: if there is a precedent in which an aggravating factor A has been held to outweigh a mitigating factor M, then it gets to be relied upon by future cases to hold that A outweighs M in the latter cases too. This is done regardless of the factual complexities unique to both the cases. For instance, the abovementioned decisions of the Supreme Court in *Krishnappa*[85] and *Sevaka Perumal*[86] were frequently cited by trial courts to dismiss mitigating factors considered as irrelevant in these two decisions: *Krishnappa* was cited in five cases by the trial courts of Madhya Pradesh to dismiss the socio-economic conditions of the accused as a mitigating factor;[87] and likewise, *Sevaka Perumal* was relied upon by several trial courts of Maharashtra to dismiss young age as a mitigating factor.[88]

This manner of reliance upon precedents has usually been outcome-based in as much as it replicates the treatment of mitigating factors, without critically appreciating the facts and circumstances unique to each individual case. It goes against "[t]he grain of individualised sentencing […] given the importance of factual specificity when it comes to both the accused and the crime".[89] Moreover, it allows sentencing judges to abdicate their duty of individually weighing the

81 *Sevaka Perumal v State of Tamil Nadu* (1991) 3 SCC 471 [12].
82 *Shimbhu & Anr v State of Haryana* (2014) 13 SCC 318 [19].
83 *The State of Karnataka v Krishnappa* (2000) 4 SCC 75 [18].
84 Project 39A (n 3) 39-40, 41-43.
85 *The State of Karnataka v Krishnappa* (n 83).
86 *Sevaka Perumal v State of Tamil Nadu* (n 81).
87 Project 39A (n 3) 28.
88 ibid.
89 ibid.

aggravating and mitigating factors unique to a case, by substituting a facts-specific determination with an outcome-based reliance on precedents.

This is obviously against the prescription of the sentencing framework in *Bachan Singh* which embodies the principle of individualised sentencing through the careful weighing of aggravating and mitigating factors specific to a case, and at the cost of repetition, with the latter even being construed liberally and expansively.

E. Life imprisonment not 'unquestionably foreclosed'

It may be recalled here that the final leg of enquiry under the *Bachan Singh* sentencing framework is to satisfy that any alternative sentences in the circumstances stand 'unquestionably foreclosed'. It has also been seen how the scope of the phrase 'alternative options' (and consequently the sentencing choices available) was expanded by the Supreme Court in *Sriharan*, by holding that High Courts and the Supreme Court may impose a special category of sentence of either life imprisonment without remission, or imprisonment exceeding fourteen years but short of life, without remission.[90] While this should have had the effect of a detailed consideration by sentencing courts on the applicability of such alternative sentences, data from the trial courts reveal that such alternative options were seldom considered.[91]

It has been found that of the forty-three capital cases decided by trial courts in Delhi between 2000-2015, life imprisonment as an alternative option was 'considered' (and not necessarily chosen) in just eight cases (i.e., in only about 19% of the cases). This was also found to be the case in twenty-two out of the eighty-two cases in Madhya Pradesh (i.e., in about 27% of the cases), and in twenty-seven out of the ninety cases in Maharashtra (i.e., in 30% of the cases).[92]

What is more worrying is that in all or 100% of the cases in which life imprisonment was considered or discussed, it was dismissed on the basis of brutality of the crime alone.[93] This seems to be consistent with the trend of attaching a trumping value to the single aggravating factor of brutality, as seen previously where brutality alone was deemed sufficient by sentencing courts to dismiss a host of mitigating factors. What the above data, however, reveals is that

[90] *Sriharan* (n 29) [176]-[178].
[91] Project 39A (n 3) 35.
[92] ibid.
[93] ibid.

brutality is also used to dismiss the alternative options that *Bachan Singh* required to be 'unquestionably foreclosed' in its final enquiry.

F. Non-consideration of the probability of reformation

The *Bachan Singh* sentencing framework categorically lists the "[p]robability that the accused can be reformed and rehabilitated" as a mitigating factor to be considered during death sentencing.[94] The framework's dictum on construing mitigating factors 'expansively and liberally' applies to the probability of reformation, as well. The framework is also clear on the matter of burden of proof to establish the probability of reformation, and requires the state to prove by adducing evidence that this criterion is not met for the accused.[95]

The data from trial courts paints a rather bleak picture of the application of the above standard, revealing that the mitigating factor of probability of reformation is "[h]ardly ever considered".[96] In a finding consistent with the previous trend of attaching a trumping value to brutality, the study finds that even where the probability of reformation is considered, "[i]t is incorrectly tied to the brutality of the crime".[97]

In each and every one (i.e., in 100%) of the cases forming part of the study of trial courts in Delhi, Madhya Pradesh, and Maharashtra, the prosecution failed to discharge its burden of adducing evidence to prove that the accused is beyond the probability of reformation. Further, the trial courts too failed to take any initiative to consider the probability of reformation as a crucial mitigating factor. The figures are staggeringly high – probability of reformation was not considered in about 83% of the cases in Madhya Pradesh, 77% of the cases in Delhi, and 58% of the cases in Maharashtra.[98]

Unsurprisingly, in the cases in which probability of reformation was considered, it was mostly dismissed on the basis of the crime having been committed in too brutal a manner – another instance where brutality seemed to have gained a trumping value in death sentencing.[99]

This trend, however, does not seem to be restricted to trial courts alone, and is rather reflective of judicial psyche towards death sentencing as large. This is reflected from the opinion study of the former Supreme Court judges wherein

"10 former judges including two former Chief Justices of India, were of the view that the probability of reformation was to be deduced from the brutality,

94 *Bachan Singh* (n 1) [206(4)].
95 ibid.
96 Project 39A (n 3) 72.
97 ibid.
98 ibid 73.
99 ibid.

and heinousness of the crime […] 14 other former judges believed either that reformation generally had no role to play in penological theory, or it had no application to death penalty cases. One judge who decided nine death penalty cases in six years at the Supreme Court dismissed the entire notion, calling it 'astrology'. Another judge who presided over 13 death penalty cases in five years at the Supreme Court, did not see the point of reformation in serious crimes […]".[100]

This approach of dismissing the mitigating factor of probability of reformation (when it gets to be considered) on the basis of the aggravating factor of brutality alone, is unfortunately in line with the consistent trend of brutality enjoying a trumping value, as also discussed in the previous parts of this article. Nevertheless, that is not the mandate of the *Bachan Singh* sentencing framework, which requires the determination of the probability of reformation to be "[i]ndependent of the circumstances of the crime",[101] and done on the basis of personal circumstances unique to the accused. To that extent, the framework seems to have hardly been complied with.

A possible way forward for a sentencing reform, where due consideration is accorded to the mitigating factor of probability of reformation, and in fact to a mitigation investigation generally, is suggested by the Delhi High Court's decision in *Bharat Singh*.[102] The High Court's approach in *Bharat Singh* stands as a rare exception to the kind of cases discussed thus far in their treatment of mitigating factors, including probability of reformation. The court's approach may be briefly discussed here. *Bharat Singh* was a case concerning the rape and murder of a three-year-old girl where the trial court awarded the sentences of rigorous imprisonment for life with a fine of Rs. fifty-thousand (and in default of payment of fine, to undergo further imprisonment for one year) for the offence of rape of a minor; and sentenced the convict to death for the offence of murder, along with a payment of another fine of Rs. fifty-thousand.

Given that as per Section 366 of the 1973 Code, a death sentence passed by the trial court has to be confirmed by the High Court, the matter of *Bharat Singh*'s sentence came up for confirmation before the Delhi High Court. The court went on to appoint a Probation Officer to examine whether the accused is likely to commit such crimes in the future, and also enquire into the probability of whether the accused can be reformed and rehabilitated – these were the two specific mandates framed by the Court for the Probation Officer.[103] Moreover, the Court provided

[100] National Law University, Delhi (n 36) 74.

[101] Project 39A (n 3) 73.

[102] *State v Bharat Singh, Death Sentence Ref No 1 of 2013 And Criminal Appeal No 509 of 2013* (High Court of Delhi), Orders dated 17 April 2014 and 31 October 2014.

[103] ibid, Order dated 17 April 2014 [68]; *Anil @ Anthony Arikswamy Joseph v State of Maharashtra* (2014) 2 SCALE 554; *Birju v State Of Madhya Pradesh* (2014) 2 SCALE 293.

specific guidance to the Probation Officer for preparing and presenting a detailed report on the above two aspects, *viz*., the 'Social Investigation Report'.[104] It directed the Probation Officer to obtain a report from the jail administration about the accused's conduct in jail; seek inputs on the behavioural traits of the accused from his family and local residents in his native village; and consult two professionals with at least ten years' experience in clinical psychology, and sociology.[105]

The Court also directed the Probation Officer to take note of the United Nations Office on Drugs and Crimes' handbook on 'Prevention of Recidivism and Social Reintegration of Offenders' brought out in December 2012; and the HM Prison Service's document entitled 'Offender's Assessment and Sentence Management, 2005', published by the UK Government.[106] Additionally, the Court passed certain procedural directions to maintain the report's confidentiality, subject to the right of the defense counsel to seek instructions upon the report before making submissions.[107]

The Social Investigation Report as submitted by the Probation Officer – which relied upon interviews with the convict himself; the local residents in the convict's native town; fellow prison inmates; jail authorities; reports from the local police station; opinion of a seven-member medical board of clinical psychologists and psychiatric social workers – revealed multi-faceted aspects about the accused's personality and personal socio-economic circumstances, apart from medical evidence relating to "[m]aladaptive personality traits or disorder especially anti-social personality". The Court relied upon the Report to note the "[d]efinitive unanimous conclusion that there is nothing to suggest that the Appellant cannot be reformed and reintegrated, and put on to the reformative process through social correctional measures".[108] Significantly, even the jail authorities had submitted a positive feedback as regards "[t]he convict's conduct in jail and his preparedness to render services to his old and ailing inmates".[109] Accordingly, the Court declined to confirm the death sentence, instead sentencing the convict to life imprisonment for the offence of murder.

The manner in which the Delhi High Court ensured that a multifaceted mitigation investigation is carried out to determine the probability of reformation is a possible solution to the existing practice of the summary dismissal of this

[104] ibid, Order dated 17 April 2014.
[105] ibid [69].
[106] ibid [70].
[107] ibid [71].
[108] ibid, Order dated 31 October 2014 [8].
[109] ibid.

crucial mitigating factor. This would shape some of the discussion in the next part of the article on the 'best-case scenario'.

G. Outcomes of the deviations from the *Bachan Singh* framework

The end-result of the preceding deviations is a distortion of the *Bachan Singh* framework beyond recognition, that has removed any principled basis for death sentencing and made it an extremely subjective and judge-centric exercise. This has led to death sentencing being beset by extremely varying and inconsistent applications. The Law Commission of India has noted that in a large number of cases where the Supreme Court imposes the death sentence, there is often no unanimity among the judges themselves either on the guilt of the accused, or on whether the *Bachan Singh* framework applies, or both.[110] Therefore, it does not come across as overstating the case when the trial courts' study, on the basis of the data discussed in this section, concludes that there has been a "[c]omplete breakdown of the sentencing framework developed in *Bachan Singh*",[111] and that "[t]he lack of compliance with the requirements laid down in *Bachan Singh* shows that death penalty sentencing is no longer carried out on any principled basis".[112]

More astonishingly, *Bachan Singh*'s application has varied even in factually similar cases when heard by different judges. This problem is best illustrated by Dr. S. Muralidhar's analysis of the Supreme Court's decision in *Harbans*.[113] He highlights how the appeals of three convicts arising from the same crime were treated differently as each appeal got to be heard by three different benches each of the Supreme Court. One accused had his appeal dismissed and was subsequently executed; the second accused's death sentence was commuted; while both the appeal and the mercy petition of the third accused were dismissed, but ultimately his sentence was commuted by the Court.[114] He also analyses many instances of disparate outcomes on similar facts,[115] and powerfully remarks that "[t]he gnawing uneasiness that the same case if heard by a different set of judges may have resulted in a different punishment will always rankle in the minds of those unsuccessful death row convicts facing the noose".[116]

This inconsistency has, in particular, put the socially and economically disadvantaged sections of the Indian society at a greater risk, who cannot afford

[110] Law Commission of India (n 3) paras 5.4.19-5.4.20.
[111] Project 39A (n 3) 74.
[112] ibid.
[113] *Harbans Singh v State of Uttar Pradesh & Others* (1982) 2 SCC 101.
[114] Muralidhar (n 73) 150.
[115] ibid 150-154; Project 39A (n 3) 44.
[116] Muralidhar (n 73) 154.

quality legal representation. The Death Penalty India Report (published in 2016),[117] notes that of the approximately three hundred and sixty-seven prisoners in India who were sentenced to death between July 2013 and January 2015, a staggering 76% belonged to a population identifying with both economically backward classes and religious minorities.[118]

Similarly, the Law Commission of India has noted that in the application of the death penalty, assumptions relating to caste and class have often been made, rendering the death sentence to operate in the larger context of persistent social prejudice.[119] This view was also echoed in Bhagwati J's dissent in *Bachan Singh* when he held that death sentence has a class bias, which led him to declare it as unconstitutional for its disproportionate impact upon the socially and economically disadvantaged sections of the society.[120]

Another major outcome of the deviations is that, with the watering down of the *Bachan Singh* threshold, death sentences have come to be awarded at a disproportionately high rate by trial courts. This becomes manifestly evident when we note that an overwhelming majority of cases in which trial courts award the death sentence, end up in either acquittals or commutations at the appellate level. For instance, between 2000-2015 where trial courts had imposed the death sentence, in 62.8% of such cases the appellate courts commuted the sentence; in roughly a third of such cases, acquittals were granted at the appellate stage; and it was only in 4.3% of such cases that the death sentence was actually upheld at the appellate stage.[121] This problem of awarding of death sentences at a disproportionately high rate by trial courts is manifest evidence of the breakdown of the *Bachan Singh* framework that sought to limit the use of the death penalty to only extreme cases.

Therefore, the deviations from the *Bachan Singh* framework have clearly led to the death penalty being administered in an extremely inconsistent manner, often resulting in the socially and economically disadvantaged sections being at its receiving end. This speaks of a crisis in the administration of the sentencing framework evolved in *Bachan Singh*.

V. The Best-Case Scenario

The discussion on the death penalty, thus far, has focused on the gaps and deviations in its application, when measured against the sentencing framework laid down in *Bachan Singh*. This has been brought forth by multiple sources of empirical evidence relied upon in Section IV. Therefore, in this Section, I venture

[117] Centre on the Death Penalty (n 38).
[118] ibid.
[119] Law Commission of India (n 3), para 5.3.9.
[120] Bachan Singh Dissenting Opinion (n 3) [81] (Bhagwati J).
[121] Law Commission of India (n 3), para 5.2.70.

to hypothesize possible responses to these deviations which may be forthcoming from a retentionist perspective. Given the credible nature of empirical evidence discussed in the preceding part, I envisage a retentionist who is at the very least open to acknowledge, when presented with such evidence, that the application of the death sentence has significantly varied from what was intended in *Bachan Singh*. This acceptive attitude towards the inconsistencies in the actual practice of death sentencing, defines who I call for the purposes of the present discussion, a 'rational retentionist'. Therefore, while a rational retentionist concedes the inconsistencies in applying the existing framework, they might very well argue that such inconsistencies, at best, only reflect a wrongful application of a framework otherwise meant to work well, and this doesn't necessarily translate into an argument for abolition, *per se*. If anything, such inconsistencies and deviations call for further strengthening of the sentencing framework by removing the gaps in its application – for how can one ask for abolishing a system that has not worked at its full potential yet? It is an argument worth engaging with.

Accordingly, an engagement with the rational retentionist would presuppose the existence of a system best suited to implement the sentencing framework to its fullest potential; in other words, a best-case scenario that is capable of applying the death penalty consistently and in line with *Bachan Singh*, while preventing the deviations previously observed. It is only after the failure of such an undertaking, that a value judgement on abolition could enter the analysis, i.e., an abolitionist shall need to demonstrate that the death penalty does not work even in the best possible scenario imagined to apply it. This part of the article, therefore, endeavours to make a genuine attempt at devising a 'best-case scenario' of administering the death sentence in as consistent a manner as possible, under a system that is best suited to remove and avoid the recurrence of the deviations discussed in the preceding parts.

The shape and form of the best-case scenario, in turn, would be a response to the probable cause behind the deviations from the sentencing framework. It is my submission that such cause lies in lack of codification of the sentencing framework in a single place of reference. The origins of the framework, of course, do not lie in any legislation, but in a five-judge bench judgement of the Indian Supreme Court – an unlikely place for sentencing guidelines. Therefore, even though *Bachan Singh* laid down the sentencing framework for administering the death sentence, each step of the framework and its position upon the finer aspects of sentencing can only be culled-out after a careful reading of the two-hundred and eleven paragraphs of the majority opinion. Moreover, the framework is only complete if the *Bachan Singh* judgement is read in conjunction with other judgements such as *Swamy Shraddananda* and *Sriharan* which added strength to it

by expanding the scope of the alternative sentencing options required by *Bachan Singh* to be 'unquestionably foreclosed'. This leaves us with a scenario where there is not one single point of reference where the exact steps to be followed as part of the death sentencing framework are codified in a logical arrangement. This is compounded by the problem of contrary subsequent Supreme Court decisions which have inaccurately applied the *Bachan Singh* framework, such as in *Machhi Singh*. The net result of this judicial confusion is lack of clear guidelines and a very wide room of flexibility available to sentencing courts in interpreting and defining for themselves, what the *Bachan Singh* framework requires of them.

Accordingly, the first step in constructing our best-case scenario is to address this lack of clarity and provide a single-point of reference for what the sentencing framework requires, in clear terms. For this endeavour, the sentencing framework must first be moved out of the multitudes of judicial opinions, and into one single codified law – the usual source for sentencing guidelines. This might be achieved through the codification of the *Bachan Singh* framework through its statutory incorporation in the 1973 Code itself. For this purpose, we may envisage a hypothetical amendment to the 1973 Code where after Section 354(3), a new Section 354(*3A*) is added.

Now, this new sub-section (*3A*) added to Section 354 could be dedicated solely to incorporating the *Bachan Singh* sentencing framework, explicitly listing out the steps of the framework as illustrated in the flowchart on page 9, Section III, along with the illustrative list of aggravating and mitigating factors provided in *Bachan Singh* and reproduced on pages 5-6. Section III. This would ensure that all the steps of the *Bachan Singh* sentencing framework (including the additions made later by the Court in *Sriharan* and *Swamy Shraddananda*) are to be found in a logical arrangement in one single place, on the back of statutory authority, so as to not leave room for any confusion for sentencing courts about what the framework requires of them, and minimize unguided discretion.

For the best-case scenario to work, however, it would need to go beyond a mere statutory incorporation of what the Court held in *Bachan Singh*, rather, it must also incorporate safeguards necessary to prevent the specific deviations discussed in Section IV. These may be best be achieved through the common practice, under the Indian statutory scheme, of appending 'explanations' to legislative provisions (or in our hypothetical, the proposed new Section 354(*3A*)). These 'explanations' are incorporated as a part of the legislative provision itself and thus are very much a part of the written text of the statute. They are frequently relied upon by courts as a key internal aid for statutory interpretation. The Supreme Court has held that 'explanations' appended to statutory provisions serve as a tool to explain the meaning and effect of the main legislative provision by clarifying

doubts and removing confusions.[122] It is settled law that 'explanations' are a common way of explaining legislative intent; removing obscurity or vagueness in the main provision; and providing additional support to the dominant object of the legislative provision, to make it meaningful and purposeful.[123] Therefore, they are a useful guide for courts while applying a statutory provision or in interpreting it; and can similarly provide the necessary clarity in applying the statutorily incorporated *Bachan Singh* framework, by preventing the specific deviations previously discussed.

Accordingly, explanations can be attached to the hypothetical Section 354(*3A*) that clarifies the effect of the *Bachan Singh* framework and incorporates a statutory prohibition for each of the deviations – such as the use of collective conscience, public opinion, and crime categories; the dominance of brutality; summary dismissal and inadequate consideration of mitigating factors; inadequate time between conviction and sentencing hearings; *etc.* Each such explanation would declare how the sentencing framework is not meant to work (in addition to the main sub-section (*3A*) itself that will declare how it is meant to work), thereby addressing the deviations that the four decades of death sentencing in India, since *Bachan Singh*, have revealed to exist. This is meant to be our best-case scenario for the reason that it is a single statutorily-codified point of reference that provides clear directions to sentencing courts, as opposed to the need of referring to multiple, detailed, and often contrary judicial opinions, to 'interpret' what the sentencing framework actually requires.

Further, as far as the probability of reformation is concerned, it has been discussed in Section IV.F that the Delhi High Court's approach in *Bharat Singh* can act as a model in the best-case scenario, capable of guiding a comprehensive mitigation investigation. Therefore, the court's directions to the Probation Officer in that case may very well also be attached to Section 354(*3A*) as a proviso.

Keeping the above principles in mind, our new Section 354(*3A*) of the 1973 Code may read as follows (for the sake of continuity, Section 354(3) is also reproduced below):

> (3) When the conviction is for an offence punishable with death or, in the alternative, with imprisonment for life or imprisonment for a term of years, the judgment shall state the reasons for the sentence awarded, and, in the case of sentence of death, the special reasons for such sentence.
>
> *(3A)* For the purposes of sub-section (3), where a court chooses

[122] *Dattatraya Govind Mahajan & Ors v State of Maharashtra & Anr* (1977) 2 SCC 548.
[123] *S Sundaram Pillai v V.R. Pattabiraman* (1985) 1 SCC 591.

to impose the death sentence for special reasons, it shall:

1) identify the aggravating factors, including but not limited to:

   (a) whether the murder has been committed after previous planning and involves extreme brutality;

   (b) whether the murder involves exceptional depravity;

   (c) whether the murder is of a member of any of the armed forces of the Union or of a member of any police force or of any public servant and was committed –

      (i) while such member or public servant was on duty; or

      (ii) in consequence of anything done or attempted to be done by such member or public servant in the lawful discharge of his duty as such member or public servant, as the case may be, or had ceased to be such member or public servant; and

   (d) whether the murder is of a person who had acted in the lawful discharge of his duty under Section 43 of the Code of Criminal Procedure, 1973, or who had rendered assistance to a magistrate or a police officer demanding his aid or requiring his assistance under Section 37 and Section 129 of the said Code

2) identify the mitigating factors including but not limited to:

   (a) the offence was committed under the influence of extreme mental or emotional disturbance;

   (b) the age of the accused. If the accused is young or old, he shall not be sentenced to death;

   (c) the probability that the accused would not commit criminal acts of violence as would constitute a continuing threat to society;

(d) the probability that the accused can be reformed or rehabilitated;

The State shall by evidence provide that the accused does not satisfy conditions (c) and (d) above.

(e) that in the facts and circumstances of the case the accused believed that he was morally justified in committing the offence;

(f) that the accused acted under duress or domination of another person; and

(g) that the condition of the accused showed that he was mentally defective and that the said defect impaired his capacity to appreciate the criminality of his conduct.

3) weigh the aggravating and mitigating factors against each other, but construct only the mitigating factors expansively and liberally.

4) commute the death sentence if the mitigating factors so constructed outweigh the aggravating factors.

5) if the aggravating factors outweigh the mitigating factors, satisfy that all the alternative sentencing options are unquestionably foreclosed. These alternative sentences include but are not limited to

5.1) life sentence for the entire natural life of the convict, without statutory remission, notwithstanding anything contained in Section 433A; and

5.2) imprisonment exceeding fourteen years but short of life, without statutory remission, notwithstanding anything contained in Section 433A.

6) choose the death sentence only if the alternative options

are unquestionably foreclosed.

Provided that an enquiry into the probability of reformation shall not be deemed to have been carried unless the sentencing court has appointed a Probation Officer to submit a Social Investigation Report, and

a) the Probation Officer has enquired from the jail administration and sought a report about the conduct of the accused while in jail. The jail authorities will extend their full cooperation to the Probation Officer in this regard;

b) the Probation Officer has met the family of the accused and the local people including from the place where the accused hails. The Probation Officer will seek their inputs on the behavioural traits of the accused with particular reference to the probability of the accused committing acts of violence as constituting a continuing threat to society; and the probability that the accused can be reformed and rehabilitated.

c) the Probation Officer has consulted and sought specific inputs from two professionals with not less than ten years' experience from the fields of Clinical Psychology, and Sociology.

d) the Probation Officer has taken note of the United Nations Office on Drugs and Crimes' handbook on 'Prevention of Recidivism and Social Reintegration of Offenders' (December 2012); and the HM Prison Service's 'Offender's Assessment and Sentence Management, 2005'.

e) the State, through the Secretary, Home Department, has made appropriate arrangements and reimbursed the expenses incurred for the Probation Officer to discharge their functions.

f) the report of the Probation Officer has been submitted within a period of ten weeks to the court in a sealed cover. As soon as the sealed cover is received, it will be opened by the court and four copies made thereof, two for the court which will be kept along with the original in the cover and resealed and two given to each of the learned counsel for the parties, both of whom shall maintain confidentiality of the said document. Nothing here shall prejudice the right of the learned counsel for the accused to seek his instructions on the report before making submissions on the

next date of hearing.

Explanation 1: For the removal of doubts, it is hereby declared that 'collective conscience' or public opinion in any form is not a relevant aggravating factor in death sentencing.

Explanation 2: For the removal of doubts, it is hereby declared that standardized and pre-determined categories of the crime are not relevant to the sentencing process.

Explanation 3: For the purposes of this sub-section, brutality or manner of commission of the crime shall not be deemed to always take precedence over mitigating factors; and shall always be weighed in the individual circumstances of a case.

Explanation 4: For the purposes of this sub-section, brutality or the manner of the commission of the crime shall be irrelevant to the enquiry into the probability of reformation.

Explanation 5: For the purposes of this sub-section, no sentence hearing shall take place unless a time of at least thirty days has lapsed since the pronouncement of the verdict on guilt.

Explanation 6: For the purposes of this sub-section, the court shall assign detailed reasons in writing to justify the apportionment of weight to each aggravating and mitigating factor; and a mere listing or mentioning of such factors shall not suffice.

Explanation 7: For the removal of doubts, it is hereby clarified that a sentencing court shall weigh aggravating and mitigating factors in the unique circumstances of each case, and not replicate the treatment of such factors simply on the basis of precedents.

VI. Inevitable Inconsistency

We have thus far discussed the deviations from the *Bachan Singh* framework and made an attempt at a best-case scenario that most closely approximates to applying the death penalty in as consistent a manner as possible, and as envisaged by *Bachan Singh*. This now equips us to engage with a rational retentionist on whether the death penalty can be consistently administered under the best possible

framework envisaged to apply it, so that the deviations from the sentencing framework are avoided.

This may be the right stage to clarify that inconsistency is not *per se* an undesirable value, and, indeed, may be welcomed in a discretion-based criminal justice system that aims to individualise justice. That each case is be treated on the merits of its own circumstances – involving a consideration of aggravating and mitigating factors unique to an accused; a personalized Social Investigation Report etc. – necessarily implies inconsistent or dissimilar outcomes. This is, in fact, reflective of the conundrum of retaining discretion in the administration of criminal law – on one hand, discretion is necessary to individualise justice and produce desirable inconsistencies, and on the other hand, it may lead to an element of unguided discretion producing unprincipled inconsistencies. The discretion conundrum is spelled out by Padfield and Gelsthorpe who have observed that

> "[t]he exercise of discretion can work in a *negative* way where it leads to unwarranted disparity and discrimination. At the same time, however, it might be suggested that discretion can be linked to justice. Decision-makers may discriminate in a *positive* way too. Thus there is an inherent tension".[124]

Therefore, while discretion has the potential to produce desirable inconsistencies, the aim of this article is to investigate the extent to which the best-case scenario leaves room for undesirable inconsistencies which would not be acceptable in any rule-based system. For instance, those inconsistencies which may make death sentencing dependent upon criteria which are irrelevant or extraneous to the law – such as personal predilections, predispositions, value-preferences or other subjective factors influencing a sentencing judge. It is this kind of inconsistency that would be considered unacceptable to the rule of law. They were found to be the reason influencing the kinds of deviations from the *Bachan Singh* framework which we explored in Section IV, and accordingly, the best-case scenario is meant to prevent this level of unacceptable inconsistencies.

The resultant argument, then, is that in the event the best-case scenario is unable to prevent such inconsistencies, we are only left to conclude the practical impossibility of applying the death sentence in a consistent manner. This is what is meant by the framework of the 'inevitable inconsistency' of the death penalty, which this article explores. The framework posits that there is something in the very nature of death sentencing that makes its application inconsistent at a large scale – which even the best-case scenario cannot salvage. This section, thus, is an

[124] Loraine Gelsthorpe and Nicola Padfield, *Exercising Discretion: Decision-Making in the Criminal Justice System and Beyond* (Taylor and Francis Group 2012) 5."

inquiry into the failures of even the best-case scenario, to the extent that it allows large-scale inconsistencies, leaving the gaps in theory and practice unamendable.

The logical import of the inevitable inconsistency framework should be a shift away from the conversation on the ways and means of infusing consistency into death sentencing, and towards arguing for its abolition on the ground that (an unacceptable level of) inconsistency involved in its application is inevitable – i.e., incapable of being salvaged through legislative or judicial safeguards internal to the criminal justice system. Therefore, the 'inevitable inconsistency' framework posits that unacceptable inconsistencies are an inevitable outcome of the retention of the death penalty in law, and on such basis alone, the death penalty is of suspect constitutionality; and its desirability should be re-evaluated.

I aim to establish the inevitable inconsistency framework in this part of the article by highlighting that the best-case scenario fails to prevent such inconsistencies in two major ways. These pertain to the exercise of weighing aggravating and mitigating circumstances – which lies at the heart of the best-case scenario; and the final leg enquiry into alternative options being unquestionably foreclosed.

A. The weighing of aggravating and mitigating factors

The first major source of inconsistency lies at the very heart of the sentencing framework, and remains seemingly unaddressed by the best-case scenario. This relates to the exercise of weighing aggravating and mitigating factors – requiring sentencing judges to apportion weight to each factor and, on a balance, make a judgement call about which set of factors 'outweigh' the other. There is, however, no objective criteria on the basis of which weights are to be apportioned to each factor in a particular case.

There is no particularly consistent or objective basis, for example, as to why the aggravating factor of brutality should outweigh the mitigating factor of young age in a particular case, unless a sentencing judge's value-preference dictates their belief that young age does not justify the brutal nature of the crime. These value judgements are bound to differ with each sentencing judge, and weights would be ultimately apportioned to each factor depending upon each judge's value-preferences, predispositions, or personal attitudes towards particular aggravating and mitigating factors. This makes the weighing exercise extremely subjective and judge-centric.

Moreover, the best-case scenario is incapable of guiding such choices without risking standardization of the sentencing process, through pre-determined weights attached to aggravating or mitigating factors. This would be antithetical to individualised sentencing which lies at the scenario's very heart. In fact, the clear prescription to the contrary is that aggravating and mitigating factors must be

weighed in the individual circumstances of each case, and pre-determined crime-categories are not relevant. This is indeed the mandate of Explanations 2[125] and 7[126] of the best-case scenario, respectively, and also what the Court in *Bachan Singh* explicitly held.[127]

The one guidance that both *Bachan Singh*, and the best-case scenario offer, is to construct mitigating factors 'expansively and liberally', but one can hardly term this as a real guidance for the weighing exercise – the requirement being sufficiently vague and nebulous. It, at best, signals to a sentencing judge that mitigating factors are more important than aggravating ones, but it does not necessarily offer any particular degree of consistency to how weights should be apportioned to each factor in individual cases.

While any sentencing framework can provide an illustrative list identifying aggravating and mitigating factors, it is nearly inconceivable for any such framework to anticipate the weight that would be apportioned to such factors in the unique facts and circumstances of every future case. Therefore, by leaving the weighing exercise to individual predispositions and sensibilities of each sentencing judge, the best-case scenario fails to prevent inconsistencies in death sentencing, rendering the death penalty to remain arbitrary in a crucial way.

B. The determination of 'unquestionably foreclosed'

The final leg of the enquiry under *Bachan Singh*, and the best-case scenario is to determine whether all the alternative sentencing options are 'unquestionably foreclosed'. An answer in the affirmative enables a sentencing court to choose the death sentence. This enquiry includes the expansive scope of the 'alternative options' available after the decisions in *Sriharan*[128] and *Swamy Shraddananda*[129], as discussed previously.

The expression 'unquestionably foreclosed' is identified here as the second major source of inconsistency, as it fails to provide guidance on when an alternative option can be considered to have been foreclosed, let alone 'unquestionably' so. These are not objective choices to make, as appeals to an alternative option being 'unquestionably foreclosed', are also appeals to individual value-judgements on whether the alternative sentences adequately respond to the perceived gravity of the crime. Inconsistencies are bound to arise when each sentencing judge is left to

[125] The bar on relying upon pre-set crime categories including on the basis of aggravating factors alone.

[126] The bar on replicating the treatment of aggravating and mitigating factors on the basis of precedents.

[127] *Bachan Singh* (n 1) [201].

[128] *Sriharan* (n 29).

[129] *Swamy Shraddananda* (n 28).

exercise personal value-judgements on whether the ends of justice in a case are served by choosing the alternative option instead of the death penalty.

Further, there is a similar lack of guidance on which alternative option is to be chosen. After the decisions in *Sriharan*[130] and *Swamy Shraddananda*,[131] two additional options of a life sentence without statutory remission, and a sentence exceeding fourteen years but falling short of life again without statutory remission, are now available to be considered in alternative to the death sentence. However, there is little to clearly define the boundaries of these two alternative sentences, and when one becomes better suited to the other. These determinations, again, narrow down on the view a sentencing judge takes on the gravity of the crime and which of the alternative options more closely respond to the same – an exercise as subjective as the determination of 'unquestionably foreclosed'.

Accordingly, as even the best-case scenario leaves room for inconsistencies at a magnitude that wouldn't be acceptable in a rule-based system, the death penalty proves itself to be inevitably arbitrary, making its consistent application a practical impossibility.

It may be also clarified here that while the charge of inevitable inconsistency may be laid against any penalty administered by a criminal justice system that relies upon judicial discretion to individualise sentencing, the presence of such inevitable inconsistency in death sentencing must be judged to a higher threshold of constitutional scrutiny. This is for the simple reason that the stakes involved in death sentencing are incomparably high with life itself facing wrongful extinguishment, which is impossibly to restore. This was acutely illustrated when in three of its decisions[132], the Supreme Court acknowledged its six previous judgements as *per incuriam*. This meant that thirteen convicts were wrongly put on the death-row; but unfortunately the decision for two of them came too late as they had already been executed by such time.[133]

VII. Concluding the Search for Consistency

Our search for consistency in death sentencing has revealed the failures of even the best-case scenario envisaged to apply the death penalty in as consistent a manner as possible. The said scenario fails because the weighing exercise and the enquiry into alternate options being 'unquestionably foreclosed', are inherently

[130] *Sriharan* (n 29).

[131] *Swamy Shraddananda* (n 28).

[132] *Santosh Kumar Satishbhushan Bariyar and Ors. v State of Maharashtra*, (2009) 6 SCC 498; *Dilip Premnarayan Tiwari & Anr v State of Maharashtra* (2010) 1 SCC 775; *Rajesh Kumar v State Through Govt of NCT of Delhi* (2011) 13 SCC 706.

[133] V. Venkatesan, 'A Case against the Death Penalty' (*Frontline*) <https://frontline.thehindu.com/cover-story/article30167180.ece> accessed 10 February 2021.

subjective and judge-centric undertakings. Moreover, this raises an unresolvable dilemma – attempts to infuse consistency into these two undertakings compromise the value of individualised sentencing and risk standardization. This unresolvable dilemma makes the best-case scenario doomed for failure, by leaving room for widespread inconsistencies at odds with the rule of law. Hence, it would appear that the inevitable inconsistency cannot be resolved, and the best-case scenario cannot be salvaged, even in other *sui generis* ways.

Now, the failure of the best-case scenario leaves us with only two options left to achieve consistency – either mandatory death sentences which are not subject to choice and discretion of sentencing judges; or no death sentence at all. Any middle-ground, even when perfected to the best-case scenario, shall produce inevitable inconsistencies not acceptable to the rule of law.

The choice between these options, however, is easily resolved in the Indian context, where the Supreme Court has already declared that mandatory death sentences for an offence are unconstitutional.[134] In *Mithu*, the Court struck down Section 303 of the Indian Penal Code, 1860, which provided the death sentence for a convict who commits murder while serving a sentence of life imprisonment. The Court reasoned:

"[T]he legislature cannot make relevant circumstances irrelevant, deprive the courts of their legitimate jurisdiction to exercise their discretion not to impose the death sentence in appropriate cases, compel them to shut their eyes to mitigating circumstances and inflict upon them the dubious and unconscionable duty of imposing a pre-ordained sentence of death".[135]

The Court further noted that a mandatory death sentence also deprives a convict of the benefit of Section 235(2) of the 1973 Code (noted in Section IV.C) to argue on sentence and show why they should not be sentenced to death.[136] Even aside from settled law, choosing a mandatory death sentence would run counter to the principle of individualised sentencing that the *Bachan Singh* sentencing framework was keen to preserve, and arguably, put us in a situation more disadvantaged than one that the said framework unwittingly led to.

Accordingly, while the search for consistency and the failure of the best-case scenario led us to a cross-road between a mandatory death sentence and no death sentence at all, at least in the Indian context, this choice already stands resolved by virtue of the Supreme Court's decision in *Mithu*. This ends our search for consistency with abolition as the only remaining option.

The 'inevitable inconsistency' framework, therefore, provides a valuable tool to re-evaluate the desirability of the death penalty by exposing the realities of

[134] *Mithu v State of Punjab* (1983) 2 SCC 277.
[135] ibid [12].
[136] ibid.

its administration and the inescapable inconsistency in its application, that won't find place in any rule-based system. While it might be present for other penalties, given the irrevocable nature of the death sentence, 'inevitable inconsistency' is proposed to be considered as a *sui generis* ground for constitutional scrutiny of the death penalty.

Cambridge Law Review (2021) Vol VI, Issue ii, 62–85

Clash of Dilemmas: How Should UK Copyright Law Approach the Advent of Autonomous AI Creations?

Oways A Kinsara *

Abstract

The advancement of artificial intelligence (AI) and deep learning (DL) has given rise to pressing dilemmas and long-standing debates, primarily in the US, on how AI-generated works ought to be conceived in copyright law. Meanwhile, there is a widely held presumption among scholars that the UK, particularly through section 9(3) of its Copyright, Designs and Patents Act 1988, is well-prepared to accommodate such works. Although today's AI differs from yesterday's generative computers, there is limited research doubting the presumption or suggesting alternative UK frameworks to address AI-generated works. Aiming to fill the gap, this article revisits the different manners in which today's AI creations encounter copyright law and explains why the current UK approach fails to address the issue. In doing so, it analyses, inter alia, the provision's legal fiction, legislature intention and judicial interpretation. It then turns to the path forward by carefully inspecting various proposed approaches to the question of copyright ownership for AI-generated works in view of the UK regime. Upon examining different models, the article highlights numerous dilemmas in each and thus argues in favour of

* Saudi Arabian Legal Researcher. LLB (King Abdulaziz University), LLM Candidate (London School of Economics). Alongside my own panel of trusted advisers, I thank Professor Andrew Murray FRSA, Dr. Martin Husovec and Dr. Luke McDonagh from the LSE Law Department whose teaching and discussions were instrumental in my understanding of the different intersections between copyright and information technology laws. All errors are my own. E-mail address: O.kinsara@gmail.com

entrance into the public domain as the least dilemmatic and most appropriate solution for AI-generated works, with promising economic and social benefits.

Keywords: AI-generated works, deep learning, copyright law, UK, public domain

I. Introduction

Technological advances, such as in artificial intelligence (AI), may be regarded as legally disruptive in many ways. They are not, however, disruptive in the sense that they are unregulable,[1] nor are they "an outside force acting upon the law".[2] Indeed, to the extent that AI challenges the law, "it does so because of how it encounters existing features of the law, both doctrinal and theoretical".[3]

To this end, the UK government requested views on the new implications of AI developments on copyright protection, with the key issue of "how we should treat works created […] by AI systems and whether the current approach is right".[4] Likewise, AI has given rise to frustration for both the US government[5] and the WIPO.[6]

While AI-generated works have engendered long-lasting debates primarily in the US, there is surprisingly extensive research arguing for the preparedness of the UK, specifically of its 1988 Copyright, Designs and Patents Act (CDPA) to accommodate such works through section 9(3) on computer-generated works (CGWs). Consequently, with today's AI diverging from yesterday's generative computers, published research suggesting alternative UK frameworks to address AI-generated works is lacking.

This article aims to fill the gap by examining the ways in which today's generative AI and its ever-increasing advances of deep learning (DL) encounter copyright law, focusing primarily on the UK regime. First, it briefly overviews the evolution of AI and clarifies what is meant by "today's AI". Second, it discusses the extent to which, if any, AI disrupts the traditional copyright framework. It then turns, in Part IV, to the UK's CGW approach, assessing its viability against new

1 Margot Kaminski, 'Authorship, Disrupted: AI Authors in Copyright and First Amendment Law' (2017) 51 UC Davis L Rev 589, 615.

2 ibid 591.

3 ibid 590.

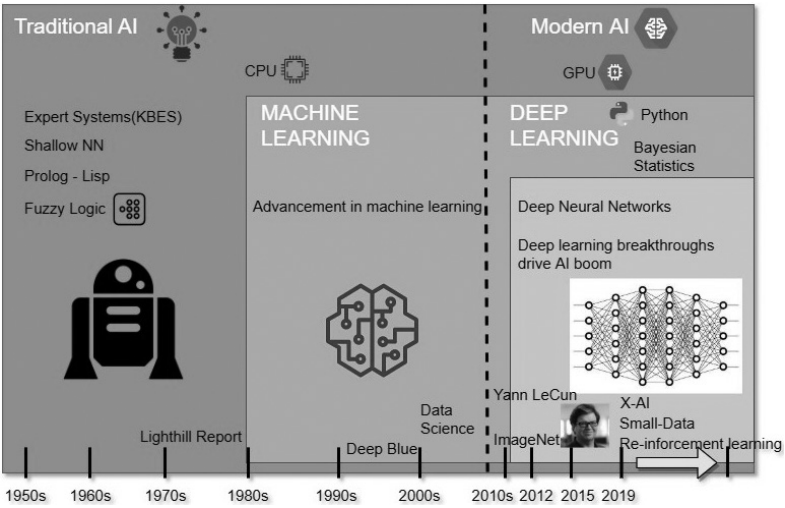
4 UK Intellectual Property Office (IPO), 'Artificial Intelligence Call for Views: Copyright and Related Rights' (*GOV.UK*, 7 September 2020) <www.gov.uk/government/consultations/artificial-intelligence-and-intellectual-property-call-for-views/artificial-intelligence-call-for-views-copyright-and-related-rights> accessed 1 March 2021.

5 US Copyright Office, 'Copyright in the Age of Artificial Intelligence' (*Copyright.gov*, 5 February 2020) >www.copyright.gov/events/artificial-intelligence< accessed 14 April 2021.

6 World Intellectual Property Organization, 'WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI)' WIPO/IP/AI/GE/19 (27 September 2019)

AI creations by analysing, inter alia, its legal fiction, legislature intention and judicial interpretation Eventually, Part V critically analyses various alternative propositions on how the law should treat AI-generated works. Upon inspecting different models, traditional and recently proposed, the article finds numerous dilemmas in each and thus argues in favour of entrance into the public domain as the least dilemmatic and most appropriate solution for AI-generated works, with promising economic and social benefits.

II. AI Evolution in Brief



70-year evolution of AI by Awais Bajwa[7]

The umbrella term AI "simply means making computers act intelligently".[8] It encompasses numerous technologies ranging from expert systems and machine learning (ML), which largely depend on a rule-based (if x then y) approach, as the primary focus from the 1950s onwards, to the advent of DL in the 2000s.[9] DL is an advancement of ML, which utilises brain-like neural networks and has the ability to learn by itself, making it more autonomous, whereas "non-deep ML is dependent on

[7] Awais Bajwa, 'Traditional AI vs. Modern AI' (*Towards Data Science*, 5 December 2019) <http://towardsdatascience.com/traditional-ai-vs-modern-ai-5117b469a0c9> accessed 8 March 2021.

[8] Michael Schmidt, 'Clarifying the Uses of Artificial Intelligence in the Enterprise' (*Tech Crunch*, 12 May 2016) <https://techcrunch.com/2016/05/12/clarifying-the-uses-of-artificial-intelligence-in-the-enterprise/?guccounter=1> accessed 9 March 2021.

[9] Bajwa (n 7).

human intervention to learn".[10] The latest trend in DL techniques is generative adversarial network (GANs), which were introduced in 2014 and are sometimes broadly referred to as generative AI. GANs are often described by their ability to "use existing content […] to create new plausible content".[11] In 2018, the MIT Technology Review classified GANs as one of the most promising AI advances in the past decade.[12] GAN architecture enables applications to be highly autonomous in generating strikingly unique outputs, which are indistinguishable from human intellectual creations, in the realm of copyright subject matter. These include *Portrait of Edmond de Belamy*, which sold for $432,500 in New York,[13] AWS DeepComposer (music generation)[14] and InferKit (text generation including articles, novels and poetry).[15] These examples are indicative of a substantial shift in how computers generate creative works, which has raised important questions on how such works encounter copyright law.

III. Creativity, Authorship and AI: Legal Disruption or Technological Misunderstanding?

This part discusses how AI-generated works encounter traditional copyright frameworks of creativity, authorship and originality and the extent to which and how they may be disrupted.

A. What is creativity?

A preliminary point is to ascertain whether works generated by today's AI are creative. While many appear so at first sight, the question under debate is whether computational creativity is different from human creativity owing to how it is generated. It is then a matter of how creativity is defined. This philosophical question is beyond the scope of this article; nevertheless, a conceptual definition is needed to continue exploring the authorship issue of AI-generated works. To

[10] Eda Kavlakoglu, 'AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?' (*IBM*, 27 May 2020) <www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks> accessed 7 March 2021.

[11] Huzaifah Saleem, 'What Is Generative AI and How Much Power Does It Have' (*IBM*, 20 August 2020) <https://developer.ibm.com/technologies/artificial-intelligence/blogs/what-is-generative-ai-and-how-much-power-does-it-have/> accessed 7 March 2021.

[12] MIT Technology Review, '10 Breakthrough Technologies' (*Technology Review*, 21 February 2018) <www.technologyreview.com/10-breakthrough-technologies/2018/> accessed 7 March 2021.

[13] Gabe Cohn, 'AI Art at Christie's Sells for $432,500' (*New York Times*, 25 October 2018) <www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.amp.html> accessed 7 March 2021.

[14] Amazon Web Services Inc, 'AWS DeepComposer Concepts and Terminology' (*AWS Amazon*, 2021) <http://docs.aws.amazon.com/deepcomposer/latest/devguide/deepcomposer-basic-concepts.html> accessed 9 March 2021.

[15] InferKit Website, 'FAQ' <https://inferkit.com/docs/generation> accessed 9 March 2021.

this end, AI researchers and computer scientists often approach the question of creativity either (a) in terms of absolute novelty[16] or (b) in operational terms which hold that AI can never be creative because it merely follows algorithmic orders.[17] In essence, their logic may be right but nonetheless irrelevant to creativity in copyright. Arguably, human creativity depends likewise on algorithmic, or algorithm-like, methods. In the words of some non-AI-sceptic researchers, "all human thought is completely algorithmic, that is, it can be broken down into a series of mathematical operations".[18] Even as regards the most intuitive romanticised literary works, Calvino characterises his creative process as:

"[A] constant series of attempts to make one-word stay put after another by following certain definite rules; or, more often, rules that were neither definite nor definable, but that might be extracted from a series of examples, or rules made up for the occasion—that is to say, derived from the rules followed by other writers".[19]

Accordingly, he sees writers as "writing machines" fed by the appropriate logic.[20] Alternatively, in considering novelty as a proxy of creativity, one must consider the latter's relative nature on a contextual basis. Bridy presented Boyden's conceptualisation of creativity as relevant to intellectual property (IP) discourse, as it differentiates between physiological (P) and historical (H) creativity.[21] The former is found in a work "that's new to the person who generated it",[22] whereas "historical novelty [...] is one that is P-creative and has never occurred in history before".[23] As Bridy found, Boden's H-creativity reflects the originality threshold in patents (absolute novelty), while P-creativity – focused on the work's novelty per se relative to its originator – "aligns with the originality standard in copyright law... [where all] work can still be considered original [...] even if another person has already created it, as long as the second work is not copied from the first".[24] Bridy here is particularly referring to the US originality doctrine as established through the *Fiest* case[25]. This standard of P-creativity likewise speaks to the traditional

[16] Roger Schank and Christopher Owens, 'The Mechanics of Creativity' in Raymond Kurzweil (ed), *The Age of Intelligent Machines* (MIT Press 1992) 395.

[17] Selmer Bringsjord, 'Chess is Too Easy' (1998) 101 MIT Tech Rev 23.

[18] ibid.

[19] Italo Calvino, *The Uses of Literature: Essays* (Harcourt Brace Jovanovich 1986) 15.

[20] ibid.

[21] Annemarie Bridy, 'Coding Creativity: Copyright and the Artificially Intelligent Author' (2012) 5 Stan Tech L Rev 1, 12–13.

[22] Margaret Boden, 'Computer Models of Creativity' (2009) 30 AI Magazine 23, 24.

[23] ibid.

[24] Bridy, 'Coding Creativity' (n 21).

[25] *Feist Publications Inc v Rural Telephone Service Co* Inc. 499 US 340 (1991).

UK originality requirement of skill, labour and/or judgment;[26] in the *University of London Press* case, Peterson J clarified, "The Act does not require that the expression must be in an original or novel form, but that the work must not be copied from another work".[27] This is often seen as a lower standard compared to *Feist*.[28] The traditional criteria applicable in the UK have nevertheless changed after the EU 2009 *Infopaq* case[29] in which the "author's own intellectual creation" standard became harmonised throughout Member States; the court explained: "It is only through the choice, sequence and combination of [the] words that the author may express his creativity in an original manner and achieve a result which is an intellectual creation".[30]

In principle, putting aside the authorship conundrum for a moment, the creativity standard "has been [only] lifted slightly by the Court of Justice European Union (CJEU) in a US *Feist*-like manner".[31] Therefore, it can be argued that AI-generated works are intellectual-like creations as they show a nexus with human intellectuality.

A dilemma, nevertheless, appears in subsequent cases, such as *Painer*[32] and *Football Dataco,*[33] where the court emphasised that originality is satisfied only if the author is employing her creative ability or creative choices that "stamp his 'personal touch'".[34] Therefore, the EU case law interlinked the intellectual creativity threshold with the reflection of one's personality, which indicates a strong association between creativity and authorship. The question of authorship will be examined in the following section. It is worth noting for now that it is yet unclear

[26] Courts have used the phrase separately and sometimes cumulatively; see Lionel Bently and others, *Intellectual Property Law* (OUP 2018) 97giving a broad context for exploring the key principles of the subject. In this fifth edition, the introduction has been updated to take account of Brexit. Important developments covered include the introduction of a doctrine of equivalents into UK patent law, the reforms of EU trade mark law (particularly with respect to 'representation' of marks, and the 'functionality exclusions'.

[27] *University of London Press v University Tutorial Press* [1916] 2 Chapter 601, 608–609.

[28] Andreas Rahmatian, 'Originality in UK Copyright Law: The Old "Skill and Labour" Doctrine Under Pressure' (2013) 44 IIC 4, 14.

[29] *Case C–5/08 Infopaq International A/S v Danske Dagblades Forening*, ECLI: EU:C:2009:465.

[30] ibid [45].

[31] Rahmatian (n 28) 15.

[32] C-145/10, *Painer v Standard Verlags GmbH*, ECLI: EU:C:2011:798, [92].

[33] C-604/10, *Football Dataco Ltd v Yahoo! UK Ltd*, ECLI:EU:C:2012:115.

[34] ibid [38].

whether the UK courts will revert to the traditional originality standard (i.e. pre-Infopaq) after Brexit.

B. Authorship dilemma

Although international treaties have been silent on defining authorship, there is a strong presumption that authors are humans as a default rule. For example, the Explanatory Memorandum to the EU Software Directive Proposal states, "In common with all literary works, the question of authorship of the program is to be resolved in favour of the natural person or group of persons who have created the work".[35]

The derogation from this entrenched rule is seen in the granting of authorship to legal persons. In the UK, section 9(1) of the CDPA 1988 states that the author "in relation to a work, means the person who creates it". The use of "person" clearly affirms the understanding that an author can either be a human being or exceptionally, such as in works by employees, a non-human *legal* person; in either case, personhood is a prerequisite for authorship.

In EU case law, authorship, like creativity, seems to be attached with the originality requirement. The interlinkage with "personal touch" infers a human origin requirement, which is resolutely affirmed by the Advocate General in *Painer*: "Only human creations are […] protected".[36] Therefore, if the work in question is acknowledged as having been autonomously originated by AI, even if deemed creative, it collides with the authorship aspect of originality in the EU and, therefore, would not be copyrighted. AI-generated works of today, therefore, are not congruent with the authorship framework in both EU and UK copyright regimes.

(i) Tool vs Autonomous Creator

Copyright works generated by the "aid" of yesterday's computers do not raise any problems concerning authorship, as such computers can easily be regarded as tools insofar as human beings are involved in the creative process of the resulting work.

A prominent example is the 1985 English *Express Newspapers* case[37], which concerned a lottery competition that was run through sequences of letters incorporated into a grid of five columns and rows. Liverpool Daily contended that copyright did not subsist in the work as it was produced by a computer, that

[35] Commission, 'Proposal for a Council Directive on the Legal Protection of Computer Programs' COM (1989) 816, [1989] OJ C91/9, 20.

[36] *Painer* (n 32) Opinion of AG Trstenjak, [121].

[37] *Express Newspapers Plc v Liverpool Daily Post & Echo Plc and Others* [1985] 1 WLR 1089.

is, without a human author. Mr Justice Whitford, however, emphasised that the involvement of a human author in addition to his intellectual skill and labour were necessary to arrive at such sequences as precise as necessary for the lottery competition to not be "hopelessly uneconomic".

Therefore, in his rejection, Mr Justice Whitford clarified that the computer is merely an aid:

> "The computer was no more than the tool by which the varying grids of five-letter sequences were produced to the instructions [...] it is as unrealistic as it would be to suggest that, if you write your work with a pen, it is the pen which is the author of the work rather than the person who drives the pen".[38]

Furthermore, according to the Advocate General's opinion in *Painer*, protected human creations may "include those for which the person employs a technical *aid*, such as a camera"[39] insofar as

> "the photographer still enjoy[s] sufficient formative freedom [...] [such as determining] the angle, the position and the facial expression of the person portrayed, the background, the sharpness, and the light/lighting. To put it vividly, the crucial factor is that a photographer 'leaves his mark' on a photo".[40]

As such, yesterday's AIs are viewed as tools since a human author's input is present in the resulting work itself, in which her stamp, skill, labour or creative choices are reflected. Conversely, today's AI cannot sit comfortably with this feature of copyright law. It is autonomous in its creations; it learns from its own experience and performs creative choices. To borrow a precise description from a computational creativity authority:

> "[The] performance [...] [is] a stand-alone matter, wherein the computer generates the result all by itself. Having written the program, the human artist then stands back, hands off, to let it run […] [thus] where G-art [(generative art)] is involved, it's especially likely that the AI system itself [...] will be credited with

[38] ibid 1093.
[39] *Painer* (n 32) [121].
[40] ibid [124].

creativity".[41]

To be further assured, let us briefly take the GAN type of DL algorithms as a case study alongside the exercise of judgment as an indication of autonomy. As conceptualised by a recent study,[42] GANs, which mimic the human mind, come from two main networks: generative and discriminator – both working against the other. The former creates new data simulating the artistic pattern of the training data, while the latter ensures that the process does not produce fake art and also improves its performance through feedback, which "maps the human process of trial and error".[43] This means that the AI judges its output throughout the creation process independently, "[u]nlike other algorithms which make use of human judges".[44]

Notwithstanding this study, one may take the ability to break free from rules and historical patterns as the required autonomy level under the EU originality-authorship standard and thus argue for the impossibility of AI systems, howsoever sophisticated, to cross that "free choices" threshold. This argument, however, does not take into account a corresponding AI case study, particularly called the creative adversarial networks (CANs), which increase the novelty of DL outputs and move away from the patterns of training datasets, thereby creating new artistic styles, where

> "the generator does not only need to fool the discriminator to think that the image that it produces is 'art', it needs also to confuse the discriminator about the style of the generated work [...] [which means that] CANs 'create' instead of merely 'emulate' [artistic creativity]".[45]

If, additionally, a work's predictability shall be the sought linkage to human authorship, today's AI has been proven to eliminate this possibility in many instances. One example is DeepDream, a generative art DL by Google, for which the programmers and training specialists could not predict the works in any great precision:

> "The results surprised Google's team. It turns out that much in the same way a child can look at a fluffy cloud and see a duck carrying

41 Boden (n 22) 31.
42 Caterina Moruzzi, 'Measuring Creativity: An Account of Natural and Artificial Creativity' (2020) 11 Eur J Philos Sci 1, 14.
43 ibid 15.
44 ibid.
45 ibid 14.

> a top hat, so can a computer. The resulting psychedelic images are like little drawings from the AI's imagination".[46]

Therefore, "[b]y any rational measure AI systems can be said to produce works creatively [...] in an accurate, logical and independent way".[47]

C. Is there anything new under the sun?

In light of the foregoing, are today's AI creations disruptive to a sufficient extent to compel reconsideration of our copyright frameworks? Grimmelmann argues that "old-fashioned pen-and-paper works raise all of the same issues; there is nothing new under the sun".[48] However, his view was informed by intrinsic human involvement in the creative process,[49] whereas works generated by today's

[46] Jack Clark, 'Trippy AI Art Jumps from Internet to TV Screens, Music Videos' (*Bloomberg*, 21 October 2015) >www.bloomberg.com/news/articles/2015-10-21/google-deepdream-s-ai-art-will-star-in-music-videos< accessed 10 April 2021.

[47] Enrico Bonadio and Luke McDonagh, 'Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity' (2020) 2 IPQ 112, 114.music and literature. There is no doubt that, as has often happened in the past during previous waves of technological advances, AI platforms-and especially, machine learning-have brought with them new opportunities as well as challenges. Machine learning is an AI application enabling programs to learn and progress automatically from experience. Its main feature is accessing data and often using it for the purpose of creating outputs, including music, literature, movies and art. Amounts of data are observed and analysed by the machine, which enables the latter to learn and then make creative decisions leading to final outputs that, as precise works of art, are often not foreseeable by the people who developed and started the initial program. Such a process is characterised by the absence of substantial human intervention or assistance after the program is operated, and by the use of algorithms-namely a sequence of instructions aimed at solving a problem or performing a computation. 1 It can be deemed \"algorithmic creativity\", or the way by which AI/machines create new works.","author":[{"dropping-particle":"","family":"Bonadio","given":"Enrico","non-dropping-particle":"","parse-names":false,"suffix":""},{"dropping-particle":"","family":"McDonagh","given":"Luke","non-dropping-particle":"","parse-names":false,"suffix":""}],"container-title":"Intellectual Property Quarterly","id":"ITEM-1","issued":{"date-parts":[["2020"]]},"page":"112-137","title":"Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity","type":"article-journal","volume":"2"},"uris":["http://www.mendeley.com/documents/?uuid=815f1eb3-0dd9-460b-b686-dbe3fc91a57b"]}],"mendeley":{"formattedCitation":"Enrico Bonadio and Luke McDonagh, 'Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity' (2020

[48] James Grimmelmann, 'There's No Such Thing as a Computer-Authored Work - And It's a Good Thing, Too' (2016) 39 Colum JL & Arts 403, 404.

[49] Especially as Grimmelmann could only imagine five differences of AI creations, which informed his analysis: "(1) […] embedded in digital copies. (2) People create them using computers rather than by hand. (3) Programs can generate them algorithmically. (4) Programmers, as well as users, contribute to them. (5) Programs can generate them non-deterministically". See Grimmelmann (n 48).

AI are different in that they "destabilize copyright law's approach to authorship by obscuring the connection between the creative process and the work".[50]

Boyden sees the emergence of algorithmic authorship in general as a "novel problem" not only in that it "fails to fall well in existing doctrinal categories",[51] such as the tool designation, but also disrupts the rationale behind the originality doctrine. Therefore, in Boyden's words, CGWs disrupt the authorship framework because "it [is] no longer […] possible to simply assume that all minimally creative elements stemmed from the mind of one or more human authors".[52] The US Copyright Office confirms Boyden's view by stating that "[copyright] law […] excludes photographs and artwork created by animals or by machines without human intervention".[53]

In between Grimmelmann's denial and Boyden's assertion of disruption, Bridy argues that US copyright law shifted long ago from the romantic notion of authorship that centres on humanness, as evidenced by the work-made-for-hire (WMFH) doctrine, which grants copyright to corporations.[54] Thus, in Bridy's eyes, AI creations do not pose a novel disruption but rather one that can easily be accommodated through extending the doctrine's legal fiction.

Regardless of whether the disruption is deeper in kind or a minor one merely requiring doctrinal tweaks, as Kaminski respectively classifies Boyden's and Bridy's stances,[55] how today's AI encounters the law should impel us to revisit its doctrines and rationales – and, more importantly, our understanding of them – for the sake of legal certainty. Therefore, the following Parts revisit various untraditional features of the law to assess whether they may effectively accommodate AI-generated works. These includes existent features, such as section 9(3) of the UK CDPA 1988, and newly invented propositions by legal scholars.

IV. Does the UK Provision on CGWs Accommodate the Advent of the Autonomous Creator?

The current UK trend took off through the 1988 CDPA, wherein section 9(3) states that, where a literary, dramatic, musical or artistic (LDMA) work is computer-generated, "the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken". The CDPA

50 Bruce E Boyden, 'Emergent Works' (2016) 39 Colum JL & Arts 377, 380.

51 ibid 379.

52 ibid.

53 US Copyright Office, 'Compendium of US Copyright Office Practices' (3rd edn., US Copyright Office 2019) Section 313.2.

54 Annemarie Bridy, 'The Evolution of Authorship: Work Made by Code' (2016) 39 Colum JL & Arts 395, 400.

55 Kaminski (n 1) 603.

defines a CGW as a work "generated by computer in circumstances such that there is no human author".[56] Lord Yang, then Secretary of State for Trade and Industry, described the Act as "the first copyright legislation anywhere in the world which attempts to deal specifically with the advent of artificial intelligence".[57] Nevertheless, today's AI may challenge the legal certainty of determining who, if anyone, made the necessary arrangements to create a particular work. This Part critically analyses the UK model and assesses its viability.

A. Categorisation paradox and legislative intention

The first dilemma lies in the wording of the designated category for CGWs, which has misled recent scholarship to suppose that CGWs are different from computer-aided works. For the latter, analogies with a camera, pen and paper or a keyboard are usually used to illustrate the idea of AI being a mere tool versus CGWs, which, by contrast, are considered to be "autonomously created by AI".[58] It is indeed tempting to draw a binary distinction, as such understanding ensures the UK approach does away, as some scholars believe, "with most potential debates about the creative works produced by artificially intelligent agents"[59] and "in many scenarios where original works are produced by computers or robots with no or little human input".[60] While this solution may sound rational at the outset, it is problematic when examined through the copyright lens of the 33-year-old Act for several reasons.

First, it contradicts the legislature's stipulation that copyright is granted to "the person by whom arrangements are undertaken",[61] which presupposes a causal link between human intervention and creative output. Thus, the CGW

[56] Copyright, Designs and Patents Act 1988 (CDPA 1998), section 178.

[57] HL Deb 12 November 1987, vol 489, col WA1477.

[58] Enrico Bonadio, Luke McDonagh and Christopher Arvidsson, 'Intellectual Property Aspects of Robotics' (2018) 9 EJRR 655.with many social settings now entailing and increasingly requiring the use of robots to support a variety of human activities. Unsurprisingly, robots' form and shape, their level of intelligence and intended purpose can vary significantly depending on the relevant industry. 1 Domestic robots are already a reality in a growing number of family homes. They include both humanoid robots which support those in need (such as the elderly, people with disabilities or children

[59] Andres Guadamuz, 'Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works' (2017) 2 IPQ 169, 175.

[60] Bonadio, McDonagh and Arvidsson (n 58) 670.with many social settings now entailing and increasingly requiring the use of robots to support a variety of human activities. Unsurprisingly, robots' form and shape, their level of intelligence and intended purpose can vary significantly depending on the relevant industry. 1 Domestic robots are already a reality in a growing number of family homes. They include both humanoid robots which support those in need (such as the elderly, people with disabilities or children

[61] CDPA, section 9(3).

regime "still trace[s] back authorship to human intervention [...] so, to a certain extent, computers are still tools in this construction".[62] Second, the expansive interpretation resulting from this categorisation error contradicts the fact that the technology coinciding with the 1988 enactment was no more than a rule-based AI and, therefore, subject to significant human intervention. It would be unreasonable to argue that the enactment intended to deal with advanced non-existent technologies, since such an argument unjustifiably credits the legislature with regulating what was, at the time, mere science fiction. Third, as analysed later, there is no judicial approach to support this understanding. Fourth, it collides with the legislature's intention as manifested in the parliamentary discussion that preceded the enactment, which asserted:

"[T]he correct approach is to look on the computer as a mere tool in much the same way as a slide rule or even, in a simple sense, a paintbrush. A very sophisticated tool it may be, with considerable powers to extend man's capabilities to create new works, but a tool nevertheless".[63]

One might wonder why, if the provision is thus interpreted, it was then legislated in the first place. Is it not enough to apply the courts' traditional pen-and-paper analogy? Such an argument is often supplemented with an additional compelling question: How do we then read section 9(3) in conjunction with section 178(b), which defines CGWs as works generated "in circumstances where there is no human author"? The response can be found in the Whitford command paper:

"On that basis it is clear that the author of the output can be none other than the person, or persons, who devised the instructions and originated the data used to control and condition the computer to produce the particular result. In many cases it will be a matter of joint authorship. We realise this in itself can cause problems, but no more than in some other fields, and we are not convinced there is a need for special treatment".[64]

It follows that the provision was indeed considered unnecessary, and if one wishes to read it in conjunction with the definition, it can only make sense to interpret it as referring to circumstances where it is too unfathomable to identify a person's contribution as distinct from that of the computer, in which case the protection would otherwise be (in a conventional sense) given jointly.

Accordingly, the CGW provision is often seen as legal fiction in that it "derogates from the general rule that defines the author as the one who creates the

[62] Ana Ramalho, 'Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems' (2017) 21 J Internet L 12, 17.

[63] Mr Justice Whitford Committee, *Report of the Committee to Consider the Law on Copyright and Designs* (Cmnd 6732, 1977) [514].

[64] ibid [515].

work".[65] The legal fiction can be further understood as a way to avoid the factual determination of ownership, where otherwise there would be joint authorship between the person and the machine. In other words, the legal fiction is an extension of the pen-and-paper tool analogy – not an expansion to the WMFH doctrine in the sense that it "vest[s] ownership of such a work in the person for whom it was prepared",[66] as some scholars wrongly believe.

B. Necessary arrangements and their responsible person

Interpreting the vague term "arrangements" is another dilemmatic task. Ramalho considered the term to be equivalent to "preparing or organising something so that the work may be created".[67] Bently and others offer greater clarity by suggesting that, in "an appropriate case", the arrangements would include the operation of the computer, the input feed or the programming.[68]

In seeking further explanation from the only available case law, the High Court had to determine whether copyrights had been infringed in graphical frames generated and displayed when playing a video game.[69] The computer software, on the one hand, plays its role by overlaying a pre-created set of bitmap images together in different orientations with a relevant cue and storing them in the memory so that it can later select and display a certain image matching the cue when prompted. The videogame developer has devised "the appearance of the various elements of the game",[70] namely by creating bitmap files, "the rules and logic by which each frame is generated and [...] the relevant computer program".[71] The player's essential arrangement is to press a specific button at a certain moment, which determines, to some extent, the frames displayed. The player, as per Mr Justice Kitchin, "is not, however, an author of any of the artistic works created in the successive frame images. His input is not artistic in nature and he has contributed no skill or labour of an artistic kind [...]. All he has done is to

65 Ramalho (n 62).

66 Bridy, 'Coding Creativity' (n 21) 27.

67 Ramalho (n 62).

68 Bently and others (n 26) 128.giving a broad context for exploring the key principles of the subject. In this fifth edition, the introduction has been updated to take account of Brexit. Important developments covered include the introduction of a doctrine of equivalents into UK patent law, the reforms of EU trade mark law (particularly with respect to 'representation' of marks, and the 'functionality exclusions'

69 *Nova Productions Ltd v Mazooma Games Ltd* [2007] EWCA Civ 219; [2007] EMLR 14, 427.

70 ibid [105].

71 ibid.

play the game".[72] Thus, it was held that it is the programmer by whom the necessary arrangements are undertaken and, therefore, who is entitled to authorship.

Three distinct, though related, observations can be made based on this orphan case, which seem to have not been highlighted in legal scholarships. First, the technology underlying the graphical output is not sophisticated. It does not involve ML or DL, thereby calling into question whether it can be called AI. Indeed, the court's judgment begins with exploring the technicalities in detail; Mr Justice Kitchin states, "Mr Jones did his programming work in Visual Basic. In writing the game he switched between writing aspects of the program and producing the graphics in the form of bitmap files. In order to create the bitmaps he used software called Adobe Photoshop".[73] Suffice it to say, *Visual Basic* was declared legacy (i.e. outdated) soon after the case. Second, the court applied a high threshold by requiring a direct causal link between the intellectual creation of the human input and that of the output, notably, by highlighting the player's input as not being artistic. Third, there are only two persons at stake in the case: a developer and an end-user.

Such a two-fold determination is too onerous to be implemented on generative DL, where multiple persons are involved, and the technical intermediaries blur the lines between the input of the human and that of the software. If the previous UK case had involved state-of-the-art DL, there would likely have been various persons at stake: an investor who funds the development of the software, a developer, a person who inserts the big data, a person who trains the system and an end-user who prompts the content.

C. The originality of CGWs

Reconciling the CGW provision with section 1(1), which requires all LDMA works to be original, begs another dilemma: Can CGWs be original? The CDPA does not establish any special test of originality for CGWs, nor does it explicitly exempt them from the originality requirement, as explained above. Meanwhile, some scholars only acknowledge the undefined challenge,[74] whereas others

[72] ibid.

[73] ibid [16].

[74] Bently and others (n 26) 117.giving a broad context for exploring the key principles of the subject. In this fifth edition, the introduction has been updated to take account of Brexit. Important developments covered include the introduction of a doctrine of equivalents into UK patent law, the reforms of EU trade mark law (particularly with respect to 'representation' of marks, and the 'functionality exclusions'

argue that the section 9(3) provision "evidently constitute[s] an exception"[75] and "divorce"[76] from originality. Others are more creative in their conclusion that originality is "self-standing and independent of authorship".[77] In McCutcheon's words, by a hypothetical analogy, "if the work had been authored by a human, or if that human could be identified, would it be original?"[78] Lacking consideration by the English courts, these understandings unjustifiably eschew the explicit language of the CDPA, which does not imply any discrimination in CGWs' favour.

As demonstrated, it is neither viable to rely on the 33-year-old provisions to solve a complexity involving increasingly autonomous AI, nor did the legislature aim to regulate such advancement at the time. Indeed, "even where they are a solution to less autonomous AIs, it is unclear who the person responsible for the arrangements is".[79] In sum, while the UK CDPA 1988 may accommodate AI-generated works in circumstances such that the AI, if a human, would be at most a joint author, it does not address circumstances where the AI, if a human, would be

[75] Bonadio and McDonagh (n 47) 120.music and literature. There is no doubt that, as has often happened in the past during previous waves of technological advances, AI platforms-and especially, machine learning-have brought with them new opportunities as well as challenges. Machine learning is an AI application enabling programs to learn and progress automatically from experience. Its main feature is accessing data and often using it for the purpose of creating outputs, including music, literature, movies and art. Amounts of data are observed and analysed by the machine, which enables the latter to learn and then make creative decisions leading to final outputs that, as precise works of art, are often not foreseeable by the people who developed and started the initial program. Such a process is characterised by the absence of substantial human intervention or assistance after the program is operated, and by the use of algorithms-namely a sequence of instructions aimed at solving a problem or performing a computation. 1 It can be deemed \"algorithmic creativity\", or the way by which AI/machines create new works.","author":[{"dropping-particle":"","family":"Bonadio","given":"Enrico","non-dropping-particle":"","parse-names":-false,"suffix":""},{"dropping-particle":"","family":"McDonagh","given":"Luke","non-dropping-particle":"","parse-names":false,"suffix":""}],"container-title":"Intellectual Property Quarterly","id":"ITEM-1","issued":{"date-parts":[["2020"]]},"page":"112-137","title":"Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity","type":"article-journal","volume":"2"},"uris":["http://www.mendeley.com/documents/?uuid=815f1eb3-0dd9-460b-b686-dbe3fc91a57b"]}],"mendeley":{"formattedCitation":"Bonadio and McDonagh (n 46

[76] Guadamuz (n 59) 176.

[77] Ramalho (n 62).

[78] Jani McCutcheon, 'Curing the Authorless Void: Protecting Computer-Generated Works Following IceTV and Phone Directories' (2013) 37 MULR 46, 51.

[79] Ramalho (n 62) 18.

eligible for copyright protection. Therefore, a new UK approach to AI-generated works is vitally needed.

V. The Path Forward: Inspecting Alternative Proposals

Since the CGW model is not useful, legal certainty may require "doctrinal tweaks" or the introduction of further legal fictions. This Part inspects three propositions on how the law should treat AI creations: (a) AI as the author of its own generated works; (b) employer (either programmer, end-user or company) authorship; and (c) entrance into the public domain.

A. AI as the author

Today's AI can autonomously produce works that would be copyrightable had they been created by a human; should authorship then be vested in the AI system itself? This solution, though radical, may arguably solve all the copyright allocation dilemmas in AI-generated works and "incentivize the creation of new and valuable creative output".[80] It would also align with the fundamental authorship principle of copyright law, which provides that the author is the person who creates the work.[81] Authorship, however, cannot be given to AI systems, at least at present, for two reasons: (a) lack of legal personhood; and (b) inability to respond to incentives.

The first reason belongs to a different territory of law, but its discussions, interestingly, are not so far from the presently discussed authorship dilemma. For scholars such as Solum, legal personhood must not be granted to AI essentially upon two grounds: (a) AI is simply not human and should be treated as property; and (b) AI lacks crucial aspects of personhood (such as consciousness).[82] Thus, Solum's reasons echo the concept of the romantic author and its associated arguments against protection for AI-generated works. His justification regarding the critical elements lacking in AI also aligns with the romantic belief that "belonging to the category of 'author' requires participation in the social, relational and dialogic practice of authorship".[83] On the other hand, Teubner finds no reason why AI

[80] Ryan Abbott, 'I Think, Therefore I Invent: Creative Computers and the Future of Patent Law' (2016) 57 BCL Rev 1079, 1121. While the article is mainly about patent law, the author did designate a section to 'Lessons for Copyright Law'.

[81] E.g., CDPA1998, section 9(1).

[82] Lawrence Solum, 'Legal Personhood for Artificial Intelligences' (1991) 70 NCL Rev 1231.

[83] Carys Craig and Lan Kerr, 'The Death of the Al Author' (2019) Osgood Legal Studies Research Papers 1, 7 <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3374951> accessed 27 March 2021.

cannot be a legal actor, drawing analogies with corporate entities.[84] This aligns with those who argue for protection on the basis that the law has already shifted from the humanness of authorship by vesting copyright in nonhuman entities.

Had the AI been considered an author of its generated works, the copyright purpose of incentives could easily be served by having the AI enter a contract of service with other natural or legal persons that respond to monetary and exclusive right incentives. Or, perhaps, by designing it to imitate a response to such incentives. Nonetheless, for now, AIs cannot (and should not) be authors until the law decides to personify them. Otherwise, the copyright law would initiate uncertainties and create unsettled questions.[85] Therefore, until such a time, it may be more reasonable to consider granting copyright to whom is already a person.

B. Employer authorship

There are plenty of arguments supporting the WMFH model by expanding and reinterpreting or, as Kaminski perceived, "tweaking"[86] it to cover AI-generated works. Under the WMFH doctrine, an employer is considered the author for copyright purposes even when an employee is the author-in-fact.[87] In other words, the approach deems autonomous AI outputs as employee works under section 11(2) of the CDPA 1988 – the English equivalent of the US WMFH. The section explains that, if a "work is made by an employee in the course of his employment, his employer is the first owner of any copyright".

Notwithstanding this, employment in this sense cannot exist unless it is between two humans or between a legal person (i.e. a corporation) and a human being.[88] Proponents of the WMFH approach, however, want the relationship between the generative AI and, say, the programmer to be deemed employment for copyright purposes by amending the doctrine definition in the sense that it "would vest ownership of such a work in the person for whom it was prepared".[89] The equivalent solution for the UK copyright law would be to broaden its definition of employment to include circumstances where a person employs an AI system to generate works with no human involvement in the creative process. Hence,

84 Gunther Teubner, 'Rights of Non- Humans? Electronic Agents and Animals as New Actors in Politics and Law' (2006) 33 J Law Soc 497.

85 E.g. Victor Palace, 'What If Artificial Intelligence Wrote This? Artificial Intelligence and Copyright Law' (2019) 71 Fla L Rev 217, 234 ("Who enforces the right? What remedies should artificial intelligence be granted? What other rights should artificial intelligence receive?"). See also Kalin Hristov, 'Artificial Intelligence and the Copyright Dilemma' (2017) 57 IDEA 431, 441.

86 Kaminski (n 1).

87 Bridy, 'Coding Creativity' (n 21) 26.

88 CDPA 1988, section 178 defines "employment" as to "refer to employment under a contract of service or of apprenticeship".

89 Bridy, 'Coding Creativity' (n 21) 27.

by assigning copyright to natural or legal persons while disregarding traditional originality analysis, the WMFH solution "avoids the predicament of vesting rights in a machine",[90] which lacks personhood and does not respond to copyright incentives, as explained above. The solution also avoids entrance into the public domain of such works. The ultimate advantage of the WMFH approach is "[p]roviding financial incentives in order to encourage the growth and development of the AI industry".[91]

The approach, however, is not free of dilemmas. From a conceptual level, such a change is not a mere "tweaking" that the confines of the WMFH doctrine can easily realise; rather, it goes beyond the doctrine to touch upon the rationale for which it has emerged. To illustrate this point, the purpose of the doctrine is to transfer the copyright of works between two parties with personhood: the real author is a human being, and her employer, who would be the author-in-law, may be a human or a corporate entity, but a person nevertheless. The proposed approach is at the end of a different spectrum: it transfers the copyright to a natural or legal person from a machine that has no personality nor any right whatsoever. Thus, the proposition "flips the purpose of the doctrine on its head".[92] Additionally, such a proposition may conflict with an essential concept associated with copyright, that is, "with rights comes responsibilities".[93] If the AI *employer*, whoever this may be, is deemed the author-in-law for an AI-generated work, she must bear responsibility when that work involves infringement, defamation or any harm. Will AI stakeholders accept such a burden for a creative process they cannot predict or control? This article joins Gervais in suggesting that "it is safer to answer in the negative".[94]

From a practical perspective, the proposition leaves uncertain who is to be deemed the AI employer. In Bridy's model, "that person would generally be the programmer in the first instance, although one could imagine situations in which it could be either the user of the program or the programmer's employer";[95] suggesting that the determination could be solved by the courts. Nevertheless, such a proposition would likely bring us back to square one, similar to the CGW model, as previously analysed, which is riddled with uncertainties. In the CGW model, we at least had human involvement in creativity as a determinate standard for

[90] ibid 25.
[91] Hristov (n 85) 444.
[92] Daniel J Gervais, 'The Machine as Author' (2020) 104 Iowa L Rev 2053, 2094.
[93] ibid 2085.
[94] ibid 2087.
[95] Bridy, 'Coding Creativity' (n 21) 27.

authorship allocation, whereas, in the proposed WMFH expansion, no certain standards seem to be agreed upon for determining the supposed AI employer.

It has been argued, instead, that copyright should be vested in a certain party, but such arguments for either the programmer, the end-user or the company as the employer (and, therefore, the author) vary. The most notable of these lend support for the AI developer, based on the assumption that AI proliferation depends on "the investment of time and skills by AI programmers and the financial backing of the companies for which they work".[96] They also contend that end-users contribute the least, if anything, to the construction and dissemination of AI systems, and thus their copyright demands should be deemed the weakest.[97] Therefore, Hristov warns that, "[b]y losing copyright claims to end users, owners and programmers may restrict the use of AI by third parties",[98] which would hinder the very objective that the copyright rationale of incentives seeks to achieve. He also suggests that there would be no conflict of authorship allocation between the AI developers and companies under the WMFH model, since independent developers would obtain copyright, while those whose services are employed would concede the copyright to their employer, who is a company or natural person.[99]

Nonetheless, arguments for end-users seem equally compelling. End-users undertake a considerable risk in purchasing the AI system hoping to end up with valuable work for commercial purposes or otherwise.[100] Furthermore, it has been argued that end-users "are in the best position to take the initial steps that will bring a work into the marketplace"; after all, it will be only at their insistence that the sought creations are brought to life.[101] Therefore, awarding them authorship of AI-generated works achieves the copyright rationale in that it will "most efficiently promote the proliferation of the devices and the works they produce".[102]

What is to be deduced here is that, even if the proposed WMFH tweak is rekindled to award authorship to a *particular* employer for the sake of certainty, there is no certain way to find a player who is more deserving than the other candidates of this incentive. In each case, it would be over-rewarding. Developers or their companies already enjoy copyright in the literary work embedded in their program and the payments consequent to their sales. They would also enjoy noneconomic

96 Hristov (n 85) 444.
97 ibid.
98 ibid.
99 ibid 445.
100 Shlomit Yanisky-Ravid, 'Generating Rembrandt: Artificial Intelligence, Copyright, and Accountability in the 3A Era – The Human-like Authors Are Already Here – A New Model' (2017) 2017 Michigan State L Rev 659, 712.
101 Pamela Samuelson, 'Allocating Ownership Rights in Computer-Generated Works' (1986) 47 U Pitt L Rev 1185, 1227.
102 Yanisky-Ravid (n 100).

fruits, such as an enhanced reputation within their industry and the wider society. End-users, on the other hand, can freely make use of such works in many ways: they may enjoy the generated poetry, art, music or drama for their pleasure, build upon it as a source of inspiration[103] and share it to raise publicity.

There is no plausible candidate who, if attained copyright, would facilitate the ultimate public policy objective of AI proliferation. On the contrary, the approach may provoke access limitations, inequality and promote "a grab-all environment".[104] To this effect, consider the case in which the programmer or her company has been deemed to be the AI employer and, therefore, the copyright owner of its generated works; thusly "enticed with the highly lucrative opportunity of obtaining copyrights at an unprecedented rate",[105] what would their response to such an incentive likely be? It is highly likely that they would obscure access to autonomous generative AIs so that they could maintain the advantage of their "AI employer" status and, in so doing, become the authors of countless commercially valuable works.[106] Otherwise, AI programmers and companies might risk losing many copyright entitlements in the hands of end-users, some of whom would withhold the work or lie about its creation source. On the other hand, if the approach decides to award end-users the copyright, it is again fairly rational to presume that AI developers and companies would hoard accessibility so as to remain users and authors in perpetuity. They might as well increase the prices of such generative DL systems at a tremendous rate instead of squandering their potential copyright advantages.

In sum, the approach introduces potential uncertainties for judges and distorts the purpose to uphold the copyright inventive whoever is the incentivised party. Furthermore, it would inevitably undermine AI growth and access equality. In all circumstances, no stakeholder would be happy under the new sun. Therefore, without any author worthy of being granted copyright, it is necessary to consider entrance into the public domain.

C. Public domain

The dilemmas of the above-inspected models lead us to favour the entrance of AI-generated works into the public domain. This approach is arguably the least dilemmatic and most appropriate solution to maintain the objective of

[103] Palace (n 85) 237.
[104] ibid 238.
[105] ibid 237.
[106] ibid.

creativity proliferation while, at the same time, eliminating uncertainties regarding authorship, rights allocation and public benefit.

Opposition to the public domain solution primarily takes the form of concern that it would inevitably diminish AI innovation and discourage creativity within the sector. Nevertheless, is it actually the case that AI pioneers and developers view a copyright monopoly as the principal means to reap the fruits of their efforts? In addition to the rewards already enjoyed by various stakeholders, as previously discussed, the incentives for companies and programmers extend far beyond copyright, which, except for the literary elements of the written software itself, is not a determining factor to AI development in this century. Indeed, AI advancement is set to move forward, with or without copyright, as it has become a race between countries, whose outcome will affect both national pride and policy;[107] the issue was put succinctly by Russian President Putin: "Who becomes the leader in this sphere will be the ruler of the world".[108]

On the other hand, the public domain solution for AI-generated works is an underappreciated mechanism. A recent empirical study found multivalued benefits for British companies using public domain works.[109] The study also illustrated that such works can attract a higher rate of funding and foster greater creativity.[110] The study, therefore, strongly opposed the notion that "overgrazing will diminish the value of public domain works".[111] Another study on the social value of the public domain suggested that, notwithstanding the protection rationale, "promoting and expanding the public domain in several key areas would yield large benefits for society in the form of increased access, greater development of complementary goods and services, and the ability to decentralise and widen the innovation process".[112] This article argues that the sphere of AI creations is ideal to unlock these advantages.

A final and important point is that the public domain solution for AI-generated works neither negates the work's creator, nor does it mean that there is nothing new under the sun; moreover, it does not contradict our previous analysis that AI systems are autonomously capable of creativity. It rather acknowledges

[107] ibid 239.

[108] CNBC, 'Putin: Leader in Artificial Intelligence Will Rule World' (*CNBC*, 4 September 2017) <www.cnbc.com/2017/09/04/putin-leader-in-artificial-intelligence-will-rule-world.html> accessed 8 April 2021.

[109] Kris Erickson and others, 'Copyright and the Value of the Public Domain' (UK Intellectual Property Office 2015) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/561543/Copyright-and-the-public-domain.pdf> accessed 5 April 2021.

[110] ibid 67.

[111] ibid.

[112] Rufus Pollock, 'The Value of the Public Domain' (Institute for Public Policy Research 2006) 15 <www.ippr.org/files/images/media/files/publication/2011/05/value_of_public_domain_1526.pdf> accessed 8 April 2021.

that an AI can rightly be an author-in-fact but cannot be the author-in-law insofar as the law does not assign it any legal right or personhood, let alone the incentive-related dilemma until the personification day comes. The solution moreover does not prevent incentivising users to publish the works generated by their AI as is currently done under the EU regime for unpublished public domain works.[113] In sum, the public domain solution serves us here as a technique to minimise uncertainties while maintaining copyright concepts, theories and rationales against potential undermining effects.

VI. Conclusion

When facing creative works, lawyers and judges promptly search for someone to award copyright to; the feasibility of vesting this right, however, is rarely questioned. This entrenched idea of protection necessity can be understood in circumstances where all forms of creative works are presumed to be human-created. Nevertheless, today's circumstances, under which computational creativity races humankind, compel us to understand how today's AI encounters copyright law and to revisit the law's ultimate objectives, not to be misguided by outdated presumptions.

By examining the operation of creativity in today's AI through numerous interdisciplinary lenses, this article emphasised that AI is drifting further away from the law's old understanding, in whose framework AI could be viewed as a mere tool, judicially analogous to a pen, involving the intervention of a human being throughout the creative process. Generative AI in relation to humans indeed reflects the *Star Wars* quote from Darth Vader: "When I left you, I was but the learner, now I am the master". Following this deduction, the article discussed how the autonomous capability of creative AI encounters traditional doctrines of copyright and concluded that, in some aspects, incompatibility exists. While AI-generated works easily meet the creativity aspect of originality, they fall short in terms of authorship. This is not merely because of the latter's romance but, more so, the personhood requirement. Likewise, the article challenged the widely held assumption that the UK CDPA 1988 provision on CGWs accommodates today's AI creations. In doing so, it analysed the provision's legal fiction, legislature intention and judicial interpretation.

Various newly invented frameworks were subsequently discussed to assess whether any could accommodate today's AI-generated works. The article rejected the AI-as-the-author proposal insofar as the law does not currently personify such systems; therefore, recognising their copyright would inevitably raise serious uncertainties undermining the legal system. While protection proponents are

[113] See Ramalho (n 62) 22.

mostly supportive of the employer-authorship solution, the article demonstrated its poor viability. Putting aside the conceptual flaws in expanding the employment doctrine, the entrenched disagreements over determining the AI *employer* reflect that copyright would be, in all circumstances, a wasted incentive to already-rewarded players, while potentially leading to access inequality and a grab-all environment and also risking public policy objectives alongside copyright rationales.

Therefore, it was found that entrance into the public domain is the least dilemmatic and most promising solution for AI-generated works. This solution would prevent over-rewarding, align with public policy and ensure fair access to an unprecedented volume of creative works. As discussed, technology innovation will inevitably continue to advance under all circumstances.

While undeniably presenting dilemmas, AI-generated works also offer an opportunity to rethink copyright doctrines, theories and rationales and our understanding of them. The article, having undertaken such rethinking, urges judiciary, legislature and legal scholars to consider a public domain model for AI creations, thereby unlocking its significant potential value, socially and economically, both to the public as well as to various stakeholders.

Cambridge Law Review (2021) Vol VI, Issue ii, 86-118

Individuals Under Observation: The Law Responds to (Live) Facial Recognition Technology

Ana Rosenthal*

Abstract

'Facial recognition' is an artificial intelligence tool that has the potential to identify individuals in real-time. The technology forms part of the growing system of mass surveillance, itself a multibillion-dollar industry which promises to bolster public safety and security. Police forces in England and Wales began testing the technology in 2017. In the absence of a statutory framework, an individual named Edward Bridges challenged the legality of its use and issued judicial review proceedings against the South Wales Police (SWP) in 2019. The legal challenge was lost by Bridges in the first instance, but the Court of Appeal overturned the decision in August 2020, finding that the use of facial recognition by the SWP had been unlawful. Most importantly, the Appellate Court held that, as a novel technology, the lack of a clear legal framework infringed the right to privacy under Article 8 of the European Convention of Human Rights (ECtHR). The Court also noted that the SWP had failed to safeguard the rights and freedoms afforded to individuals by the Data Protection Act 2018, and found that in each deployment, the police had overlooked whether the technology had a gendered or racial bias. The implications of the judgment are significant, not least because it maps out how Parliament may want to legislate on such technologies in the future. This article explores the theoretical and legal implications behind facial recognition,

* LLM, Birkbeck, University of London. Thank you to Dr Bernard Keenan, my research supervisor, as well as the anonymous reviewers from the Cambridge Law Review. I am also grateful to Eva, Norman, Manuela, and Andrej for listening to my thoughts. Any errors are my own. anarosenthalm@gmail.com.

particularly at a time when individual and fundamental rights have been brought into even sharper focus as a result of the global pandemic.

Keywords: facial recognition technology, surveillance, individuals, data collection, privacy

I. Introduction

Edward Bridges lives in Cardiff, South Wales. He is a white British man, and a father of two. He works in an office job at Cardiff University.[1] Ed is the Claimant in *Bridges v South Wales Police*,[2] the world's first legal challenge in the courts concerning the use of Automatic Facial Recognition (AFR) technology. Ed claimed his face was caught by an AFR camera on two deployments by the South Wales Police (SWP); the first time, he had been shopping in Cardiff, and on the second, he was peacefully protesting against the Cardiff Arms Fair. Today, Ed, together with the human rights group Liberty, has successful campaigned and challenged the SWP in the Court of Appeal for the way AFR was being used.[3]

Robert Williams lives in Detroit, in the US state of Michigan. He is a black African American, and also a father of two.[4] He is an officer worker at an automotive supply company. Robert was the first known individual in the US to be wrongfully arrested because of an incorrect alert on a facial recognition system. He was accused of shoplifting in January 2020. He spent "30 hours in custody and was released on a $1,000 personal bond".[5] Today, Robert, alongside the American Civil Liberties Union and the University of Michigan Law School's Civil Rights

1 Steven Morris, 'Office worker launches UK's first police facial recognition legal action' (*The Guardian*, 21 May 2019) <https://www.theguardian.com/technology/2019/may/21/office-worker-launches-uks-first-police-facial-recognition-legal-action> accessed 28 July 2021. See also, Ed Bridges, 'End lawless and dangerous police use of facial recognition technology' (*Crowdjustice* blog, 11 June 2020) <https://www.crowdjustice.com/case/facial-recognition/> accessed 28 July 2021.

2 *R (Bridges) v Chief Constable of the South Wales Police* [2019] EWHC 2341 (Admin).

3 *The Queen (on the application of Edward Bridges) v The Chief Constable of South Wales Police & others* [2020] EWCA Civ 1058.

4 Kashmir Hill, 'Wrongfully Accused by an Algorithm' (*The New York Times*, 24 June 2020) <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html> accessed 28 July 2021.

5 ibid.

Litigation Initiative, has filed a lawsuit against the Detroit Police Department to obtain compensation and ban the technology for the way it operates in the US.[6]

It is useful to consider how facial recognition has impacted both Ed and Robert because their cases indicate the complexities affecting an individual, including issues of racial bias.[7] This article, however, concentrates on the topical case of Ed Bridges because its aim is to survey the attitude of the UK and its legal system. The legal issues surrounding facial recognition in the UK relate to fundamental rights and data protection laws, particularly in the absence of statutory regulation on the matter. This article considers what it means to be an "individual" as facial recognition systems begin to alter the parameters of mass surveillance in the public sphere, even though the topic has immense legal implications for all countries, those largely democratic and those dictatorial.

This article is primarily library-research based, and relies on internet research and news articles because of the limitations posed by the Covid-19 restrictions in 2020. It involves a doctrinal analysis of recent cases that reveal the justifications for the deployment of AFR. Doctrinal research contrasts primary sources, i.e. case law, with secondary academic papers to help frame how the law has begun to develop. This article pays particular attention to how the *Bridges* case evolved in the courts, using the Divisional Court's judgment to contrast what was later decided by the Court of Appeal – a judgment the SWP chose not to challenge any further.

The article is structured as follows: Section II aims to define AFR. The section explains what the courts have said about the deployment of AFR, and reflects on why this particular technology has become an issue of both public and legal importance. The idea is to start looking more deeply into the socio-legal context behind its use to be able to situate the role of the individual within the debate. Section III outlines useful frameworks that help analyse some of the effects of AFR. It focuses on ideas put forward by theorists such as Michel Foucault or Gilles Deleuze, and situates them amongst more recent ideas developed by writers such as Jackie Wang and Evgeny Morozov. It introduces the work of Claudio Celis Bueno, a researcher who also adopts Deleuze's work to explain why facial recognition is tied to an inherent contradiction, i.e., "the weakening the processes of individualisation on the one hand and the growing centrality of the face on the

6 Tate Ryan-Mosley, 'The new lawsuit that shows facial recognition is officially a civil rights issue', (*MIT Technology Review,* 14 April 2021). <https://www.technologyreview.com/2021/04/14/1022676/robert-williams-facial-recognition-lawsuit-aclu-detroit-police/> accessed 28 July 2021.

7 For further research on race, as well as gender and class issues see: Joy Buolamwini, MIT Media Lab. See also, Steve Lohr, 'Facial Recognition is Accurate, if You're a White Guy' (*The New York Times*, 9 Feb 2018) <https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html> accessed 28 July 2021.

other".[8] Section IV looks more closely at why facial recognition might begin to "weaken" the idea of the individual. We explain this in the context of the growth of mass surveillance. In particular, this section delves deeper into the *Bridges* case and evaluates how human rights law and data protection regulation create highly procedural norms on new surveillance technologies. Section V reflects on why the "face" remains pertinent to the debates on facial recognition. The tool accentuates structural issues in society as it can categorise individuals into different profiles. In particular, we focus on the 'watchlist' – an apparatus that has given facial recognition its unique characteristics and power to individuate. Finally, Section V concludes.

II. Facial Recognition Technology: What Is It and What Have the Courts Said About It?

Automated (or live) facial recognition (hereinafter referred to as 'AFR')[9] can be described as an artificial intelligence tool which measures facial features to develop a unique facial code for an individual. The algorithm then uses those measures and matches them to other facial images that will be stored on a database or 'watchlist'. The result generated by the algorithm is based on a percentage of matching features i.e. a threshold of similarity, rather than a straightforward 'yes' or 'no' test.[10]

AFR is a tool that varies in use: it can open your phone, tag photos on social media or if in operation, may detect that you are at either an airport or busy location. For example, between May 2016 and March 2018, AFR was used by private developers in quasi-public spaces, namely Kings Cross, London.[11] In the future, experts indicate that AFR will have the potential to start analysing our emotions.[12] However, what is certain today is that the tool has become increasingly

[8] Claudio Celis Bueno, 'The Face Revisited: Using Deleuze and Guattari to Explore the Politics of Algorithmic Face Recognition' (2020) 31 Theory, Culture & Society 73-91.

[9] AFR is the technical name used by the SWP in the *Bridges* case.

[10] Silkie Carlo et al., 'Faceoff: The Lawless Growth of Facial Recognition in UK policing' (*Big Brother Watch*, May 2018) <https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf> accessed 28 July 2021.

[11] Dan Sabbagh, 'Facial Recognition Technology Scrapped at King's Cross Site' (*The Guardian*, 2 September 2019) <https://www.theguardian.com/technology/2019/sep/02/facial-recognition-technology-scrapped-at-kings-cross-development> accessed 28 July 2021. See also, Zoe Kleinman, 'King's Cross developer defends use of facial recognition' (*BBC*, 12 August 2019) <https://www.bbc.co.uk/news/technology-49320520> accessed 28 July 2021.

[12] Hannah Devlin, 'AI systems claiming to 'read' emotions pose discriminatory risks' (*The Guardian*, 16 February 2020) <https://www.theguardian.com/technology/2020/feb/16/ai-systems-claiming-to-read-emotions-pose-discrimination-risks> accessed 28 July 2021; Madhumita Murgia, 'Emotion recognition: Can AI detect human feelings from a face?' (*FT*, 12 May 2021) <https://www.ft.com/content/c0b03d1d-f72f-48a8-b342-b4a926109452> accessed 28 July 2021.

pervasive in modern society, largely as public bodies have tested AFR for the purposes of security control and crime detection.

Since 2017, various police departments across the UK have had the capacity to deploy automatic facial recognition, and it can be used in two ways. First, a tool referred to as "AFR Identify" allows law enforcement agencies to upload pictures of an unidentified suspect or by using a "probe image" that may relate to a previous crime incident. This is then matched against images held by a police custody database system containing approximately 500,000 pictures.[13]

Second, the technology has capabilities to be used in real-time using a method known as "AFR Locate".[14] Cameras are mounted onto police vehicles to survey an area, and as people pass within the designated catchment area, every face is filtered through the AFR system. This system includes a 'watchlist': a list of wanted people that has been compiled by the operator in advance of its deployment. If a match alert is triggered, then the system operator, for example, a police officer, would need to decide in real time how to respond based upon whether they consider the alert to be a true positive or not. If no match is made, which will occur in the vast majority of cases, then the AFR Locate system deletes the facial biometrics or images that have been filtered through the live system.[15]

The capabilities of the first model – AFR Identify – incited controversy in early 2020 when *The New York Times* published its investigation into a relatively unknown start-up called Clearview AI, a company that had been "mining" public images on Facebook, YouTube and other well-known portals.[16] It was revealed that Clearview AI's facial recognition tool was holding over three billion images in its database, and the tool was being sold to law enforcement agencies and various private businesses for security purposes.[17] The Clearview-AI model, with its immense database collection and retention system, is not in operation in the UK.[18]

[13] *Bridges* (n 2) 27.

[14] ibid 28.

[15] ibid 37.

[16] Kashmir Hill, 'The Secretive Company That Might End Privacy as We Know It' (*The New York Times*, 18 January 2020) <https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html> accessed 28 July 2021.

[17] ibid. See also, Kashmir Hill, 'Facial Recognition State-Up Mounts a First Amendment Defence' (*The New York Times*, 11 August 2020) <https://www.nytimes.com/2020/08/11/technology/clearview-floyd-abrams.html> accessed 28 July 2021; CCN Business, 'Clearview AI's Founder Hoan Ton-That speaks out [Extended interview], (*YouTube*, 6 March 2020) <https://www.youtube.com/watch?v=q-1bR3P9RAw> accessed 28 July 2021.

[18] 'Clearview AI' (*Clearview AI*, 12 May 2021) <https://clearview.ai/law-enforcement> accessed 28 July 2021.

In fact, the legal issues in *Bridges* focus specifically on the second 'live' model: AFR Locate.[19]

To test AFR, the Home Office first awarded a contract to the South Wales Police (SWP) – which then deployed it at the Champions League final, which took place in Cardiff in June 2017. Thereafter, additional trials were set up by other police departments, including by London's Metropolitan Police in early 2020.[20] As stated above, in 2019, Ed Bridges – represented by the civil liberties group Liberty – brought the first legal challenge concerning the use of AFR against the SWP.[21]

At first instance, the Divisional Court held that whilst AFR interferes with privacy rights, the current legal regime does provide adequate safeguards that guarantee the "appropriate and nonarbitrary use" of the technology, adhering both to the Human Rights Act 1998 and the data protection legislation.[22] In rejecting the Claimant's arguments for Judicial Review, the Court of Appeal decided to review the legal matter – with the three day trial taking place over Skype in June 2020 because of the pandemic.

For a case adjudicating on new technologies, the online trial was replete with technical difficulties – from muffles to echoing sounds, to sudden interruptions and some delay.[23] In fact, as a participant observer, I recall how the trial was initially scheduled to be broadcasted on YouTube, but the sounds coming from the three judges in the courtroom meant it was impossible to hear the Appellant's QC, who also appeared from his echoey office. Shortly after starting, Sir Terence Etherton, the Master of the Rolls, had to temporarily suspended the sitting as he ordered the search for technical aides. The rest of the Court awaited instruction, at which point Dame Victoria Sharp, the President of the Queen's Bench Division, and Lord Justice Singh, both sitting as justices, could be heard in light conversation. I noticed Singh cautioned silence, noting "we might potentially be live on YouTube", whilst gently nodded towards the cameras.[24] It provoked a sense of irony that the judge

[19] *Bridges* (n 2) 27-28; *Bridges* (n 3) 11.

[20] Damien Gayle, 'Met Police Deploy Live Facial Recognition Technology' (*The Guardian*, 11 February 2020) <https://www.theguardian.com/uk-news/2020/feb/11/met-police-deploy-live-facial-recognition-technology> accessed 28 July 2021.

[21] *Bridges* (n 2).

[22] ibid 1, 159.

[23] Owen Bowcott, 'UK's facial recognition "breaches privacy rights"' (*The Guardian*, 3 June 2020) <https://www.theguardian.com/technology/2020/jun/23/uks-facial-recognition-technology-breaches-privacy-rights> accessed 28 July 2021.

[24] Court of Appeal Trial, R (*Bridges) v CC South Wales & ors* (C1/2019/2679) N.p. 2020. Web. 23-25 June 2020.

himself felt very aware that he was being watched – a slight indication as to what would follow.

On 11 August, 2020, the CA released its judgment. The CA held that the SWP's use of AFR was not "in accordance with the law" – specifically, as the legal framework did not support what it called 'the "who" and the "where" question', i.e. in what circumstances can AFR be used.[25] In other words, the technology had the capability to violate our Article 8 rights to privacy because there were insufficient safeguards used to determine who might be on the 'watchlist'.[26] In addition, the SWP had failed to take reasonable steps to make enquires about whether AFR has racial or gender bias.[27] Liberty took to twitter, posting: "it's almost lunchtime, the sun is shining, and discriminatory police #FacialRecognition tech is unlawful".[28]

III. Situating Facial Recongition: A Conceptual Framework

On 24 January, 2020, some weeks prior to the first Covid-19 lockdown, the Metropolitan Police announced that it would start using live facial recognition operationally for the first time on the streets of London. Twitter reacted. Silkie Carlo, Director of the civil liberties group, Big Brother Watch, reshared the Metropolitan's post with the caption: "see you in court".[29]

Perhaps what was most surprising, certainly from a legal perspective, was that the Metropolitan Police had overlooked the fact that the first facial recognition case, *Bridges*,[30] was already making its way through the courts. The Court of Appeal had only recently granted its permission for appeal on the basis that *Bridges* has a "real prospect of success".[31] It seemed as though the Metropolitan Police was using the Divisional Court's judgment as authority for a full-scale operational deployment that moved beyond the trial phase.

A. Theories to surveil

To analyse the effects of the rapid deployment of AFR by the state, a useful starting point is the conceptual triad developed by Gilles Deleuze in his

[25] *Bridges* (n 3) 91.

[26] ibid.

[27] Ibid 182-199.

[28] Liberty (*Twitter*, 11 August 2020). <https://twitter.com/libertyhq/status/1293145042253742081> 11 August 2020.

[29] Silkie Carlo, (*Twitter*, 24 January 2020). <https://twitter.com/BigBrotherWatch/status/1220686136806445057> accessed 28 July 2021.

[30] *Bridges* (n 3)

[31] Monidipa Fouzder, 'Court of Appeal to hear facial recognition technology challenge' (*The Law Society Gazette*, 20 November 2019) <https://www.lawgazette.co.uk/news/court-of-appeal-to-hear-facial-recognition-technology-challenge/5102241.article> accessed 28 July 2021.

short essay 'Postscript on the Societies of Control'.[32] Deleuze recounts the way in which new societies of 'control' have evolved from what Michel Foucault had previously identified as 'discipline' societies. Foucault, as Deleuze contends, had recognised that spaces of enclosure, for example, the factory, the prison, the hospital, had disciplined societies away from what Foucault termed societies of 'sovereignty', which functioned to "tax rather than organise production, to rule on death rather than to administer life". Accordingly, discipline power produces a form of subservience, intended to form an interiorised type of behaviour. But now, Deleuze argues, it is the *disciplinary* systems themselves that have fallen into crisis – to be replaced by societies of control. Writing in 1990, Deleuze did not live to witness the exponential growth of CCTV, let alone the current trialling of facial recognition by the state, but he did envisage a shift taking place within the network of surveillance.

According to Deleuze, societies of control move away from the structure of enclosure and allow for an illusion of freedom. The apparatus of power can exert control over us precisely by provoking a sense of variation and continuous change. The contemporary writer Jackie Wang, who has examined the racial, economic and legal influence of the US carceral state, writes that: "it is possible that as technologies of control are perfected, carcerality will bleed into society".[33] She implies that the physical structures of prisons might be superseded by new surveillance methods because those methods can be portrayed as both economically viable as well as a benefit to individual freedom. Wang argues that this development may bring about the "birth of a more all-encompassing police state".[34] To her mind, spaces that are policed develop into carceral spaces.

In *Discipline and Punish*, Foucault famously developed a theory of modern urban *panopticism* to describe a process of observation that is utilised to guarantee order. Foucault details how: "the panoptic mechanism arranges spatial unities that make it possible to see constantly and recognise immediately".[35] He notes that the key effect of the Panopticon is: "to induce […] a state of conscious and permanent visibility that assures the automatic functioning of power".[36] However, the subtlety to Foucault's argument reveals that the Panopticon is essentially a tool of power

[32] Gilles Deleuze, 'Postscript on the Societies of Control' (1992) 59 MIT Press 3-7. <https://cidadeinseguranca.files.wordpress.com/2012/02/deleuze_control.pdf> accessed 28 July 2021.

[33] Jackie Wang, *Carceral Capitalism* (Semiotext(e) Intervention Series 2018) 39.

[34] ibid 40.

[35] Michel Foucault, *Discipline and Punish: The Birth of the Prison* (Penguin 1991) 200.

[36] ibid 201.

used to supervise individual conduct to increase profitability and the productivity of an activity.

The notion of productivity, therefore, is useful for thinking about the way norms on new surveillance have been framed. In fact, an example of this appeared during the *Bridges* trial at the Court of Appeal in an argument made by the SWP – the defendants. In their submissions, the SWP openly declared that AFR should be understood as a tool for policing "lower-level suspected persons".[37] This point was made on the basis that current forms of policing remain too resource intensive and require a large workforce to trace those who fail to engage with the criminal process, for example, people on warrants. To that end, Jason Beer QC, appearing for the SWP, submitted that AFR serves as "a tool that narrows the pool of individuals".[38]

In his lectures on *The Birth of Biopolitics*, Foucault traces the origins of productivity in what he denotes to be the liberal art of government, which originated during the eighteenth century. Foucault argues that liberalism is not aimed at securing freedom, but it is instead closely related to the freedoms provided for by the market. He writes: "the formula of liberalism is not: 'be free'. Liberalism simply formulates the following: I am going to produce what you need to be free".[39] In other words, it emphasises a technique of governance that is less focused on the imperatives of individual freedoms and more on "the management […] of the conditions in which one can be free".[40] Foucault continues: "freedom is something which is constantly produced", but notes that the cost to manufacturing this freedom is "security".[41]

Liberalism, therefore, develops as a mechanism that continually needs to: "arbitrate between the freedom and security of individuals by reference to [a] notion of danger".[42] Indeed, the "problem of security" requires finding the balance between the collective and individual interest. This framework certainly applies to the way in which the courts produce and apply legal norms to surveillance technologies. This is because the courts will limit themselves on interfering on issues of national security. Thus, the law will engage itself in a process that is less about "ruling" and more about finding a balance between liberty and the protection of rights on the one hand, and the power of state to surveil on the other.

Interestingly, the philosopher Frédéric Gros developed a theory on security in his book *The Security Principle*. Gros formulates four definitions for security, but

[37] Court of Appeal Trial, *R (Bridges) v CC South Wales & ors* (C1/2019/2679). N.p. 2020. (23-25 June 2020).

[38] ibid.

[39] Michel Foucault, *The Birth of Biopolitics: Lectures at the Collège de France 1978-79* (Palgrave Macmillan 2008) 63.

[40] ibid 63-64.

[41] ibid.

[42] ibid.

relevant here is the notion of 'biosecurity' – which is defined as a measure that is necessary to "protect, control and regulate the individual".[43] Importantly, he suggests 'control' is one of the practices of biosecurity as it seeks to trace human beings through unique biological traits.[44] In other words, Gros suggests that "control" is not about observing nor correcting individual behaviour through a centralised gaze or institution, but rather a means to track movements and actions through the numerous traces that individuals leave behind.

For Gros, facial recognition would be another identifier that "helps things run more fluidly".[45] Gros indicates how tracking devices now allow security forces to identify and locate individuals, serving as "irrefutable evidence".[46] Gros adds that the other major function of the new techniques of control is the "compiling of digital data".[47] In his words: "our acts today have acquired an almost indestructible memory".[48] In this context, Gros comes close to describing what the watchlist might be: "the traces can be brought up again for anyone at any moment – and the justice system can itself lay hold of them".[49]

In a manner akin to Hannah Arendt's concept of "objective enemies" that occur under totalitarian regimes, Gros says that the countless files that exist mean we are all now "objective subjects". However, Gros dismisses the idea that new technologies of control are lifting the "spectre" of totalitarianism.[50] This is because modern forms of control fall short of spying on people to verify their ideological conformity. In fact, technology lacks an ideological undertone in the sense that informed consumers seem willing to share their biometric identity so that they can access a "mode of sociability".[51]

Importantly, as Gros suggests, totalitarianism would follow a principle of nonreciprocity between those that controlled the surveillance apparatus versus those that were *subjected* to it.[52] In fact, totalitarian forms of surveillance were rooted in highly centralised and hierarchical forms of power. Today, the difference is that

[43] Frédéric Gros, *The Security Principle* (Verso 2019) chapter 4.
[44] ibid 152.
[45] ibid 157.
[46] ibid 153.
[47] ibid 154.
[48] ibid.
[49] ibid.
[50] ibid. 159.
[51] ibid.
[52] ibid.

modern technologies of control also exist in more democratic and participatory systems, though this is perhaps an increasingly "privatised" hybrid model.

B. Eternal technological amelioration?

Evgeny Morozov has contributed to the argument by outlining how the ideological imperative of new technologies are in fact rooted in their quest to problem solve. Morozov states that the fall back to the promise of "eternal amelioration" brought about by technology inevitably carries less scrutiny in regard to its ethics.[53] Morozov does not oppose technology itself, but he seeks to alert the reader to the idea that not every problem will demand a technological fix. In fact, any belief in technological perfection that fails to address the intricacies of the human existence, he argues, is likely to create new problems.[54]

Race, for example, has become a central problem in the debate surrounding the rapid deployment of AFR by the state. Evidence gathered in an independent report for the Metropolitan Police Trial of AFR, by Professor Peter Fussey and Dr. Daragh Murray, suggests that indirect discrimination appears in two ways. First, the technology may react differently depending on an individual's sex, race, or colour, which is an issue linked to technical performance as the technology carries an inbuilt bias. The second, and perhaps more long-standing issue, relates to the way in which the technology is used because it often correlates with potentially discriminatory policing processes.[55] The legal scholar Andrew Guthrie Ferguson offers a detailed account as to why machines get racial or gender bias wrong.[56] Specifically, algorithms replicate biases as machine-learning tools learn by correlating past data sets. Ferguson explains:

> "[…] even if race were completely stripped out of the model, the correlation with communities of colour might still remain because of the location. A proxy for racial bias can be baked into the

[53] Evgeny Morozov, *To Save Everything Click Here: Technology, solutionism, and the urge to fix problems that don't exist* (Penguin, 2014) xiii.

[54] ibid 1-16.

[55] Murray P. Fussey and Daragh Murray, *Independent Report on the London Metropolitan Police Service's Trial of Live Facial Recognition Technology* (The Human Rights, Big Data and Technology Project 2019) 40. <https://48ba3m4eh2bf2sksp43rq8kk-wpengine.netdna-ssl.com/wp-content/uploads/2019/07/London-Met-Police-Trial-of-Facial-Recognition-Tech-Report.pdf> accessed 28 July 2021.

[56] Andrew Guthrie Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement* (New York University Press 2020).

system, even without any formal focus on race as a variable".[57]

Therefore, a key limitation that is likely to occur is the potential for the algorithm to misidentify an individual as a result of racial or gender biases being built into the model. Jackie Wang elegantly summarises the issue: "A wrong 'You may also like […]' product recommendation on Amazon is one thing, but a wrong prediction in the arenas of punishment, policing, and finance is quite another".[58]

The rapid development of data science is creating methodologies for policing that focus on anticipating and predicting crime.[59] Anticipatory methods that aim to prevent crime have always been used, but new data tools may increasingly allow police officers to make decisions based on an algorithm. Interestingly, Evgeny Morozov has labelled predictive policing as the "epitome of solutionism" as it appears to be an easy and logical step, but it does not adequately consider the complex "environmental vulnerabilities" that encourage crime to begin with.[60] In addition, any underreporting in crime can also lead to huge discrepancies that limit the capabilities of algorithmic policing.[61] The fear is that labelling subjects as potential risks could actually end up producing individuals as such.[62] Ferguson reminds us that data is not blind. In fact, he says: "data is us, just reduced to binary code".[63]

Here, we ought to recall Deleuze's argument on how new modes of power brought about by digital technologies no longer sustains the individual as the product – which arose more clearly under Foucault's disciplinary mode of power. To Deleuze, in the society of control, individuals have become "*dividuals*".[64] The implication is that the individual can be fragmented into endless data points or codes. Deleuze notes that: "The numerical language of control is made of codes that mark access to information, or reject it".[65] In short, it is the way that data can be collected, and then used as a method of control, that delineates us as "dividuals".[66] However, Deleuze's digital theory of control perhaps sits at odds with the pervasive

57 ibid 122.
58 Wang (n 33) 49.
59 ibid 42.
60 Morozov (n 53) 182-185.
61 ibid.
62 Wang (n 33) 43.
63 Ferguson (n 56) 123.
64 Deleuze (n 32) 5.
65 ibid.
66 Celis Bueno (n 8) 79.

Western ideologies of individuality – and this includes the legal logic employed in human rights law, that seeks to protect the individual's rights and freedoms.[67]

Rouvroy and Berns use the term "algorithmic governmentality" to uncover the rationality behind the automated collection and analysis of big data, and its effects on populations.[68] However, central to their argument is that such processes are not focused on individuals, nor subjects, but "on relations".[69] The result is a "reduction of opportunities to challenge forms of 'knowledge' production based on datamining and profiling".[70] Most interestingly, they contend that: "[…] 'power' grasps the subjects of algorithmic governmentality no longer through their physical body, nor though their moral conscience […] but through multiple 'profiles' assigned to them, often automatically, based on digital traces of their existence and their everyday journeys".[71]

It appears that Rouvroy and Berns' analysis anticipates the power of the 'watchlist' in facial recognition technology:

> "The fact of [algorithmic] power having a digital rather than a physical 'grasp' in no way means than individuals are […] reducible to networks of data. Instead, the function of algorithmic power is to draw out: […] 'profiles' (as a potential fraudster, a consumer, a potential terrorist, a student with high potential, etc.)".[72]

AFR does, therefore, seem to create a problematic position for the individual in light of its algorithmic capabilities. In a more recent article, which has also adopted Deleuze's theoretical framework of control, academic Claudio Bueno Celis argues that facial recognition technologies encounter an inherent contradiction: "a weakening of the processes of individualisation on the one hand, and an ever-growing centrality of the face as a mechanism of individualisation on the other".[73]

Importantly, Celis contends that the technology does not simply function as a ubiquitous panopticon, but ought to be understood as an "apparatus of metadata that goes beyond the task of individualisation".[74] And yet, the tool also retains an element that will utilise the individual's "face" as a disciplinary diagram of

67 Paul Bernal, 'Data gathering, surveillance and human rights: Recasting the debate' (2016) 1(2) Journal of Cyber Policy 243-264, 243.

68 Antoinette Rouvroy and Thomas Berns, 'Algorithmic governmentality and the prospects of emancipation' (2013) Réseaux 163-169.

69 ibid v.

70 ibid x.

71 ibid xi-xii.

72 ibid xii.

73 Celis Bueno (n 8) 81.

74 ibid 80.

control.[75] The remainder of this article will explore this contradictory duality through a legal context. In using Celis' argument as a generalised framework, it aims to address what the law is doing for the individual within the context of ever-growing AFR systems.

IV. Mass Surveillance: The 'Weakening of the Individual'

The 'weakening of the processes of individualisation' might be best understood within the context of mass surveillance capabilities. In fact, Celis suggests that AFR can be viewed from the perspective of what he recognises as "machinic enslavement" – the idea that AFR algorithms "do not operate as apparatuses of individualisation but rather as an apparatus of metadata", i.e. "information about information" to create control.[76]

Importantly, the legal and political debate on surveillance intensified in the wake of the Edward Snowden revelations in 2013. Previously, ideas about surveillance levels in the UK were largely a matter for speculation – as intelligence and security agencies can adopt a "neither confirm nor deny" policy.[77] Today, we can differentiate new forms of surveillance from more "traditional" methods like phone-tapping or photography.[78] Present-day techniques capture both "content" data and also what is referred to as "metadata". The former targets communications of known individuals, whilst the latter can involve "bulk" interception of large amounts of data with no specific target in mind.[79] The important point, however, is that "metadata", which is perhaps best described as the "residual-like" part of collected data, is often more revealing as it can be processed and analysed very quickly.

In particular, facial recognition uses machine-learning algorithms to harvest biometric data, which can be collected in bulk from large numbers of people. In essence, these algorithms do not use a pre-given template to match a facial image to a specific person. Instead, they are based on statistical calculations that uncover patterns in the "training datasets".[80] Broadly speaking, the process involves "feeding" millions of images to the algorithm in order to "train" it to identify faces. During the process, however, the face "fragments […] into bits of information that no longer belong to a private individual, but rather constitute a data bank of face templates and training sets".[81] It follows, therefore, that identity

[75] ibid 81.
[76] Celis Bueno (n 8) 84, 80.
[77] Bernal (n 67) 246. Also see, *Bridges* (n 3) 95.
[78] ibid 246.
[79] Bernal (n 67) 248.
[80] Celis Bueno (n 8) 80.
[81] ibid 84.

is reduced to a number of quantitative measures[82] – what Deleuze would define as the "*dividuals*".[83] Accordingly, the next section discusses how the law has responded to the way AFR has increased the 'potential' for mass surveillance by the state, and why it has often developed a problematic position for the individual. We review this in the context *Bridges*[84] – contrasting certain elements from the Divisional Court position with the final judgment by the Court of Appeal.

Importantly, *Bridges* was initiated on the premise that AFR would have profound consequences for privacy (Article 8 of the European Convention on Human Rights – ECHR) and data protection rights.[85] In this section, we will first look at the grounds challenging the breach of Article 8(1) and (2) of the ECHR, which primarily considered if the use of AFR was "in accordance with the law" – per Article 8(2). Second, we briefly review how the courts addressed the data protection claim in *Bridges*, once again showing that judges will engage in a proforma checklist exercise, rather than assessing broader questions in regard to the need to collect large amounts of information belonging to the individual. We argue that the Court of Appeal recognised that AFR is a technology that might "weaken" what it means to be an individual given the potential baleful surveillance capabilities. At the same time, we show how the courts tend to avoid making assessments in regard to the wider "necessity" for bulk surveillance capabilities, favouring instead a procedural approach that relies on formal safeguards, as well as controls set by UK regulatory bodies. In many ways, adopting a procedural approach means the legal discussion has moved beyond a reiteration of Ed Bridges as the individual subject, essentially focusing on the effects of the overall system.

We should note that *Bridges* disputed one final key point in regard to the public sector equality act, though it is not discussed here as the focus of this section is on the laws approach to mass data.

A. Privacy: *Bridges v South Wales Police*

To reiterate, in *Bridges*, the Divisional Court had considered that whilst AFR interferes with the privacy of every individual scanned, i.e. Article 8(1) ECHR, the legal frameworks assumed adequate safeguards for police forces to utilise AFR.[86] The Court of Appeal, however, overturned this decision in part. Most importantly, in a unanimous decision, the judges held that the Divisional Court

[82] ibid 81; see also: Bernal (n 67) 248.

[83] Deleuze (n 32) 5.

[84] *Bridges* (n 2); Bridges (n 3).

[85] Dan Squires QC, Aidan Wills and Megan Goulding 'Skeleton Argument on Behalf of the Claimant for the Substantive hearing on 21-23 May 2019' *R (Bridges v The Chief Constable of South Wales Police*) CO/4085 2.

[86] *Bridges* (n 2) 159.

had erred in finding that South Wales Police's (SWP) interference with Mr Bridge's privacy right was "in accordance with the law" – per Article 8(2). In particular, although the SWP relied on a legal framework comprised of primary legislation, secondary legislation in the form of codes of practice, and local policies, there was no specific guidance as "to 'where' AFR could be used and 'who' could be put on a watchlist".[87] As such, the Court of Appeal recognised the debilitating potential of AFR for an individual, like Mr Bridges, who had never appeared on a watchlist. However, and equally important, the Court of Appeal said that the Divisional Court had been right in its balance of the "proportionality" question. In their view, the benefits of AFR remain important, and the Court of Appeal specified that the impact on Ed Bridges, as an individual, had ultimately been minor in the circumstances.[88] Thus, the general 'use' of AFR could be proportionate under an Article 8(2) assessment if the technology operated under clear safeguards, which alludes to the idea that the individual remains secondary to the Court of Appeal in that the technology has a role in monitoring elements of society.

Importantly, the dispute over whether AFR ever "engaged" Article 8 was settled by the parties before the case got to the Court of Appeal. This is because the collection of biometric data contains personal information of an "intrinsically private" nature.[89] In other words Article 8 matters in circumstances where facial biometrics are retained even for a short time, including the near instantaneous processing of an individual's biometric data where no match occurs. Therefore, privacy rights are merely "triggered by the initial gathering of the information".[90] However, any lawful interference with Article 8 privacy rights will also need to abide by Article 8(2) of the ECHR:

> There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

The legal question that often takes shape in regard to laws that grant powers of "bulk interception" to law enforcement agencies tend not to focus on whether concrete harm has been done to a specific person. Instead, the issue for the courts is whether the law abides by the principles of "legality, legitimacy, [incorporating]

87 *Bridges* (n 3) 91.
88 ibid 131-141.
89 ibid 57.
90 ibid 59.

checks and balances" as a means to safeguard against potential abuses by the state.[91] Broadly speaking, the Court of Appeal's judgment in *Bridges* supports this statement, to the extent that the Divisional Court was wrong to find that SWP's use of AFR was "in accordance with the law" as per 8(2) largely because: "AFR is a novel technology".[92]

Previously, the Divisional Court had found the legal framework to be satisfactory, in large part because the police is "a creature of the common law".[93] The Divisional Court had relied on both *R (Wood)*[94] and *R (Catt)*,[95] to state that the use and retention of photographs by police is justifiable to maintain "public order and identify crime".[96] Moreover, the lower court had suggested that AFR was less intrusive than other technologies, like DNA or fingerprints, which had reinforced the idea that police would not require new and express statutory powers. However, the Court of Appeal overturned this position – on the suggestion that AFR "involves the capturing of the images and processing of digital information of a large number of members of the public […] the vast majority of them will be of no interest whatsoever".[97] In fact, the technology gathers '"sensitive" personal data,' which is then "processed in an automated way".[98] If viewed from Celis' conceptual framework, the Court seems to express concern that AFR might "weaken" the processes of individualisation because the technology operates within a power vacuum, to the extent that "too much discretion is currently left to individual police officers".[99]

Nonetheless, in asking whether a lawful interference with Article 8(1) could be necessary, both courts referred to the four-stage test in *Bank Mellat*,[100] which enlists four broad principles for the objective justification of a limitation on a Convention right, i.e. when might it be necessary to restrict our individual human rights.[101] Before the Court of Appeal in *Bridges*,[102] the Appellants expressed concern over the "fourth" question in *Bank Mellat* within the context of AFR, namely "whether a fair balance has been struck between the rights of the individual and

91 Bart van der Sloot and Eleni Kosta, 'Big Brother Watch and Others v UK: Lessons from the Latest Strasbourg Ruling on Bulk Surveillance' (2019) 5 Eur Data Prot L Rev 252, 256.
92 *Bridges* (n 3) 86.
93 *Bridges* (n 2) 69.
94 *R (Wood) v Commissioner of Police of the Metropolis* [2010] 1 WLR.
95 *R (Catt) v Association of Chief Police Officers* [2015] AC 1065.
96 ibid 7. See also, *Bridges* (n 2) 70-71.
97 *Bridges* (n 3) 87.
98 ibid 88-89.
99 ibid 91.
100 *Bank Mellat v Her Majesty's Treasury* (No 2) [2014] AC 700.
101 ibid; see also, *Bridges* (n 2) 98; *Bridges* (n 3) 132.
102 *Bridges*, CA (n 3) 132.

the interests of the community".[103] On this substantive question, the Divisional Court had previously concluded that the deployment of AFR by SWP "struck a fair balance and was not disproportionate".[104] Various reasons had been given, including that SWP deployed it in an "open and transparent way, with significant public engagement". Furthermore, the Divisional Court had stated that the trials had only ever sought individuals on a "watchlist", and no data had been kept.[105] In fact, the lower court noted that AFR could be used to save resources for the police.[106]

By contrast, the Appellants argued that the Divisional Court had erred in finding that SWP's use of AFR was proportionate according to Article 8(2) because it AFR Locate affects individual rights as well as the rights of every member of the public. Interestingly, the Court of Appeal considered that because the interference with the Appellant's Article 8 rights had never been "in accordance with the law", it should not be expected to adjudicate on this matter, yet the Court felt it was important to briefly address the issue to some degree.[107] The judges suggested that the impact on Mr Bridges – as an individual – had been minor in relation to the potential benefits of this technology, saying that "an impact that has very little weight cannot become weightier simply because other people were also affected. It is not a question of simple multiplication".[108] This assessment by the Court of Appeal implies that AFR might be deemed "proportionate" and "necessary" under Article 8(2) assessment in the future, which could ultimately overshadow many individual concerns.

B. Privacy: procedural v necessity

The European Court of Human Rights (ECtHR) itself has typically avoided questions of 'necessity' in regard to surveillance measures.[109] Instead, the ECtHR favours a more pragmatic/procedural approach and preserves the right for governments to have a "margin of appreciation" in regard to their own security issues.[110] However, the ECtHR has upheld that surveillance is "necessary in a democratic society" if it responds to a "pressing social need", and if the

[103] ibid.
[104] *Bridges* (n 2) 101.
[105] ibid.
[106] ibid 106.
[107] *Bridges* (n 3) 131.
[108] ibid 143.
[109] Maria Helen Murphey, 'A Shift in the Approach of the European Court of Human Rights in Surveillance Cases: a rejuvenation of necessity?' (2014) 5 *European Human Rights Law Review* 507-518. 510. See also, Kirsty Hughes. 'Mass surveillance and the European Court of Human Rights', (2018) E.H.R.L.R. 6, 589-599
[110] Murphey (n 109) 515.

interference is proportionate to the aim pursued and can be readily justified by national authorities.[111]

Accordingly, the ECtHR has also focused on whether the interference with Article 8 rights was "in accordance with the law", which should be noted means an adherence to the "quality" of the law doctrine – i.e. is the law foreseeable, accessible and overall compatible with the rule of law.[112] We know that this was the key dispute in *Bridges*.[113] However, the specificity of this approach shows that the courts will concentrate on evaluating the adequacy of existing procedural safeguards rather than determining with greater detail the 'necessity' of surveillance "in a democratic society".

The question of 'necessity' is important because data gathering – of any kind – continues to have the potential to invade the privacy of a population.[114] In an article on human rights and bulk surveillance, the academics Murray and Fussey propose that a more "nuanced" approach to privacy is required.[115] In particular, the threshold for use of bulk capabilities should only be satisfied by issues that constitute a "active" threat or impairment to the workings of a "democratic society". This means the law needs to change to be able to fully position the difference "between those situations in which bulk powers are useful and those situations in which they are vital".[116] To that end, we submit that "necessity" allows us to situate the individual in the context of mass surveillance more clearly.

Nonetheless, questions of "necessity" have often been superseded by the courts precisely because in a "democratic society" there is expectation that institutions abide by principles of accountability and oversight.[117] Moreover, abiding by a rigorous application of proportionality might interfere with the boundaries marked by the separation of powers as surveillance policy is usually thought to be better suited to executive expertise.[118] In that way, Maria Murphey shows us that the perceived benefit to the procedural approach is that: "it avoids

[111] *Catt v United Kingdom*, ECtHR, (Application nos. 43514/14), 24 Jan 2019, para 109. See also Kyriakos Kotsoglou and Marion Oswald, Marion, 'The Long Arm of the Algorithm? Automated Facial Recognition as evidence and trigger for police intervention' (2020) 2 Forensic Science International: Synergy 86-89, 87.

[112] Murphey (n 109), 508. See also, Keenan (n 109) 4.

[113] *Bridges* (n 3).

[114] Bernal (n 67) 259-260.

[115] Daragh Murray and Pete Fussey, 'Bulk Surveillance in the Digital Age: Rethinking the Human Rights Law Approach to Bulk Monitoring of Communications Data' (2019) 52(1) Israel Law Review 31-60.

[116] ibid 58.

[117] Bernal (n 67) 259-260.

[118] Murphey (n 109) 517.

direct competition between individual rights and the stated public interest in this complex and highly sensitive area" we call surveillance.[119]

This procedural logic in regard to Article 8 appeared most acutely in the case of *Big Brother Watch v UK*,[120] a prolonged legal dispute in response to the Edward Snowden revelations in 2013, now settled by the Grand Chamber of the ECtHR. It is useful to highlight some of its key consequences to outline why favouring a procedural approach on issues of mass surveillance capabilities carries certain limitations. In generalised terms, the procedural method to justice prioritises an oversight system that works to evaluate its own procedures, rather than delineating what forms an acceptable level of surveillance and what is "necessary in a democratic society" to truly protect individuals.[121]

In *Big Brother Watch*, various NGO's raised concerns over the UK's legal regime governing the sharing of foreign intercepted material collected by the US government, as well as the 'bulk' collection of metadata by UK intelligence services GCHQ under the codename TEMPORA.[122] Put very simply, the ECtHR reviewed key questions in regard to the compliance of the UK surveillance framework under the Regulation of Investigatory Powers Act 2000 (RIPA).[123] This act allowed for the interception of internal and external communications, after the Secretary of State had issued the relevant warrants.[124] On this matter – and relevant to our discussion – the Claimants argued that the UK's legal framework on the interception of bulk communications/metadata was not "in accordance with the law" and amounted to an interference with Article 8(2) – the right to privacy.

In its judgment, the Grand Chamber of the ECtHR stated that "Surveillance which is not targeted directly at individuals […] has the capacity to have a very wide reach indeed, both inside and outside the territory of the surveilling State".[125] Interestingly, the Court noted that "the degree of interreference with individuals Article 8 rights will increase as the bulk interception process progresses",[126] even where "the stored material is in coded form, intelligible only with the use of

[119] Murphey (n 109) 511.

[120] *Big Brother Watch and Others v United Kingdom ECtHR* (Application nos. 58170/13, 62322/14 and 24960/15) 25 May 2021.

[121] Keenan (n 109) 8, 17. See also, Murphey (n 109) 511.

[122] *Big Brother Watch* (n 120). See also, Van der Sloot and Kosta (n 91).

[123] RIPA has now repealed by the Investigatory Powers Act 2016 (IPA). See also, Keenan (n 109) 9; Tzanou, Maria, '*Big Brother Watch and others v. the United Kingdom*: A Victory of Human Rights over Modern Digital Surveillance?' (*Verfassungsblog*, 18 September 2018) <https://verfassungsblog.de/big-brother-watch-and-others-v-the-united-kingdom-a-victory-of-human-rights-over-modern-digital-surveillance/> accessed 28 July 2021.

[124] *Big Brother Watch* (n 120).

[125] ibid (n 120) 322.

[126] ibid (n 120) 330.

computer technology".[127] In deciding on the bulk interception of communications, the Grand Chamber said that national authorities withhold a "wide margin of appreciation" to address mounting national security concerns.[128] However, the Court acknowledged that since all interception regimes, whether bulk or targeted, are potentially open to abuse, the discretion "afforded [to Contracting States] must be narrower and a number of safeguards will have to be present".[129] To that end, the Court upheld that surveillance: "must be subject to end-to-end safeguards", i.e. a checklist of requirements for "supervision and independent *ex post facto* review".[130] In short, the Grand Chamber enlisted the *procedure* for how collected data ought to be used, stored, and then destroyed. In a similar manner, the Court of Appeal in *Bridges* considers that at the very least there needs procedural rigour on "who" is placed on a watchlist and "where" AFR can be deployed to be "in accordance with the law".[131]

Importantly, the Claimants in *Big Brother Watch*, had urged the Court to "update" the list of requirements, demanding for both prior "judicial authorisation" to be sought, and for "subsequent notification" be given to a target of surveillance.[132] The lower ECtHR court – the First Chamber – in its 2018 judgement, had rejected this view on the premises that "it would be wrong automatically to assume that bulk interception constitutes a greater intrusion into private life of an individual than targeted interception".[133] The Grand Chamber agreed, noting that "judicial authorisation" might be a useful safeguard against arbitrariness, but 'not a "necessary requirement". The Grand Chamber went further and upheld that "bulk interception should be authorised by an independent body" – separate from the "executive".[134] Such authorising bodies should be "informed" of any bulk interception operations, overriding a system based on "notifying" a target of surveillance.[135] To give just one example, the ECtHR reviewed the supervisory role carried out by the IC Commissioner to examine any complaints concerning unlawful interception.

Notably, the Divisional Court in *Bridges* did the same on the question of "proportionality", for example, saying that the Information Commissioner and

[127] ibid.
[128] ibid 338.
[129] ibid 347.
[130] ibid 350.
[131] *Bridges* (n 3) 91.
[132] *Big Brother Watch and Others v United Kingdom ECtHR* (Application nos. 58170/13, 62322/14 and 24960/15) 18 September 2018. 316. See also, Keenan (n 109) 5-6; Van der Sloot and Kosta (n 91) 257.
[133] *Big Brother Watch* (n 132) 316.
[134] *Big Brother Watch* (n 120) 351.
[135] ibid 357-359.

Surveillance Camera Commissioner both exist to provide oversight.[136] Interestingly, the Court of Appeal concluded that it should not be for the Divisional Court to adjudicate on whether the SWP had an "appropriate policy document" regarding the use of AFR within the meaning of Section 42 of the DPA 2018 for determining the sensitive processing of data as the Information Commissioner exists to issue guidance on that point.[137] However, as mentioned earlier, there are limits to relying on such a highly procedural approach, largely because it reflects a form of surveillance bureaucracy whereby the courts end up limiting their own supervisory role. The overreliance on "oversight" bodies means that verifying the actual "need" for surveillance in general becomes difficult.[138] Moreover, such systems have been described by legal academics as "judicial-executive hybrids", and it is important to note that such systems have failed in their supervisory role in the past: i.e. like the surveillance programmes revealed by Edward Snowden.[139]

In adopting such a procedural approach – the debate on surveillance, or AFR, is then largely shaped through a definition of how "intrusive" a technology is rather than assessing its substantive necessity. In fact, the notion of "intrusiveness" played a key role in the *Bridges* litigation – and the concept shaped much of the Divisional Court's assessment on the "proportionality" of AFR use.[140] Interestingly, the Divisional Court, in accordance with Lord Sumption's judgment in *Catt*, had said that: "intrusive methods" only equate to "entry on private property or acts […] which would constitute an assault".[141] Accordingly, AFR was not like fingerprints or DNA because there exists: "no physical entry, contract or force […] to obtain biometric data".[142] The Court of Appeal agreed, but nonetheless stated that AFR was not "analogous" to photographs or CCTV either,[143] thus pushing the debate to focus on procedural safeguards, i.e. how the collected data is "utilised, stored, and destroyed" rather than addressing the need for collecting mass data and assessing whose biometric data should be stored in the Police National Database.[144]

C. Data protection: *Bridges v South Wales Police*

We briefly turn to how data protection law responds to the growth of data within the context of facial recognition, showing that whilst individual protection

136 *Bridges* (n 3) 108.
137 *Bridges* (n 3) 160-161.
138 Bernal (n 67) 259.
139 Keenan (n 109) 14.
140 Kotsoglou and Oswald (n 111) 87.
141 *Bridges* (n 2) 74; refer also to: *Catt* (n 95) 7.
142 ibid 75.
143 *Bridges* (n 3) 77.
144 Keenan (n 109) 8. See also, Bernard Keenan, 'Automatic Facial Recognition and the Intensification of Police Surveillance' *Modern Law Review* (2021) 1-21, 17.

is considered, the law does not sufficiently control the initial *collection* of that data. Put very simply, regulation protects "personal data" – i.e. the information about an identifiable individual. However, understanding that data "is a commodity challenges the status quo".[145] This is because the gathering of image data "depersonalises the individual data subject; whereby, ultimately the subject becomes a template for future reference or [even] a marketing target".[146] Accordingly, there is a sense that a "weakening of the process of individualisation" might begin to occur, particularly as a highly proceduralised methodology is followed.

In many ways, Edward Snowden has accurately captured how the problematic nature of data protection laws, including GDPR, is simply cited within the name: "the concern is with data protection, not data collection".[147] In this manner, the laws in place also exist to list procedures, rather than addressing wider questions over the necessity of collecting large amounts of data. As the Appellants had contended in their skeleton argument before the court: "The Defendant places considerable reliance on the DPA 2018. […] [T]he data protection principles contain little of specific relevance to determine when and in what circumstances the capturing of facial biometrics through AFR is or is not permissible".[148]

The Divisional Court had explained that the first key issue in regard to the data protection claims, concerned the extent to which AFR involved the processing of personal data according to the Data Protection Act 2018 (DPA). Interestingly, the Divisional Court reasoned that AFR does "individuate" people in the sense that AFR will distinguish individuals from one another: an indication that data protection principles will apply.[149] Though the individual would be taken into consideration once data is processed, the impact of data surveillance is largely "portrayed" to occur when data is examined by humans and not when it is gathered or decoded by an algorithm.[150]

To that end, the Court of Appeal adjudicated on two points: first, whether the SWP had complied with the obligation to undertake an "impact assessment" – in compliance with section 64 of the DPA 2018; and second, given that AFR involves the "sensitive processing" of an individual's biometric data by a public authority, whether the SWP would be expected to comply with three requirements enlisted in section 35(5) per DPA 2018. Specifically, the Court of Appeal was

[145] Ian Berle, 'The Future of Face Recognition Technology and Ethics: Legal Issues' in Berle Ian, *Face Recognition Technology* (Springer 2020) 180.

[146] ibid.

[147] Marius Dumitrescu, 'WebSummit 2019 – Edward Snowden: The problem isn't data protection. The problem is data collection' (*YouTube*, 5 November 2019) <https://www.youtube.com/watch?v=Ezp16KD8dVw> accessed 28 July 2021.

[148] Squires QC (n 85) 68.

[149] *Bridges* (n 2) 122, 124.

[150] Bernal (n 67) 243.

concerned only with the third criteria: if data processing occurs, the "controller", for example, the police, must have an "adequate" policy document in place (section 42 of the DPA 2018) that outlines all procedures relating to the sensitive handling of information, including the retention and deletion of data.

On the first point, and based on their Article 8 justifications, the Court of Appeal held that the "impact assessment" by the SWP had failed to properly assess the "risks to the rights and freedoms of data subjects".[151] There is a sense that the Court of Appeal recognises how AFR can "weaken" the position of the individual given the fact almost anyone could have been placed on the watchlist.[152] However, on the second point, the Court of Appeal dismissed the claim. The Divisional Court had said the SWP had an existing policy document, even though it may have lacked sufficient detail.[153] It was agreed that it was up to the Information Commissioner to give "further guidance" on the level of detail necessary, rather than having the judges intervene. It shows the courts relying on the oversight of regulatory bodies and the existing formal safeguards – a position the Court of Appeal was quick to adopt.[154]

To some extent, we might say that the "political struggles" for data protection (and privacy) are somewhat "losing strategies"[155] as norms for surveillance as both favour a highly procedural method. In this way, the law is not so much "ruling" as it is moderating, regulating, and also responding to the need for balance. The recent Court of Appeal judgment ultimately gives way for the future regulation of AFR, rather than acknowledging wider issues regarding the actual need for AFR use and the impact it may have on individuals.

V. The Face: A Mechanism of Individualisation

This last section remembers the idea of the face as a mechanism of individualisation. To reiterate, Celis indicates that even though the growth of mass surveillance helps us trace a "shift from individual to population" the human face – which "belongs to the human domain of individuality" – is still important to the debate on AFR. He highlights that: "the face is always political", and therefore, it becomes useful to consider the "concrete circumstances which trigger the social production of the face".[156] This section will expand on this analysis through

[151] *Bridges* (n 3) 152-153.
[152] ibid 152.
[153] *Bridges* (n 2) 139-141.
[154] *Bridges* (n 3) 161.
[155] Celis Bueno (n 8) 88.
[156] ibid 77.

its socio-legal perspective, and in particular reflects on the importance of the 'watchlist' given its vital role in selecting faces to help categorise the individual.

In *Discipline and Punish*, Foucault suggests that disciplinary societies "characterize, classify, specialize; they distribute along a scale, around a norm, hierarchize individuals in relation to one another and, if necessary, disqualify and invalidate".[157] Celis, however, draws attention to another important passage in Foucault's book that pinpoints how disciplinary techniques "mark the moment when the reversal of the political axis of individualisation […] takes place".[158] Foucault describes how in feudal regimes, and perhaps certain other societies too, individualisation belonged to the "echelons of power". However, inside the modern, or disciplinary society, as power is "more anonymous and more functional, those on whom [power] is exercised tend to be most strongly individualised". Accordingly, "in a system of discipline, the child is more individualized than the adult, the patient more than the health man, the madman and the delinquent more than the normal and the nondelinquent".[159]

However, as suggested earlier, Foucault, as well as Deleuze, elucidate the ways in which mechanisms of "security" and "control" gradually replaced disciplinary societies, which implies the individual is also replaced by the population as the new mode of power.[160] Nonetheless, as Celis contends, even if the individual might recede in the society of control, the "face" still retains "haunting significance and political consequence".[161] He suggests that the face gathers new meaning if understood from the context of "social subjection".[162] This can be explained as a "diagram of power" that sees humans "subjected" to external objects, such as machines. Celis notes that if facial recognition is viewed as an "apparatus of social subjection [AFR technologies] link the face to a private individual, rendering the face as the sign of a privatized body".[163] This is how subjects become: 'a potential "consumer", "criminal" or "terrorist"'.[164] In other words, social subjection provides

[157] Foucault (n 35) 223.
[158] ibid 192. See also, Celis Bueno (n 8) 79.
[159] Foucault (n 35) 192-193.
[160] Celis Bueno (n 58) 79.
[161] Claudio Celis Bueno, 'The Face Revisited: Using Deleuze and Guattari to Explore the Politics of Algorithmic Face Recognition' (2020) 31 Theory, Culture & Society 73-91, 81.
[162] Celis Bueno (n 8) 81.
[163] ibid 83.
[164] ibid.

roles, it categorises, and it "produces individual subjects with identities – and ID cards".[165]

Therefore, the face itself remains pertinent to the debate on ARF as the technology is still "automating the process of individualisation".[166] In a somewhat more Foucauldian manner, the consequence is that individuals can be easily hierarchized in relation to one another – and selected within the crowd. In regard to AFR, this process occurs most acutely though the formulation of the "watchlist", which simultaneously acts as a legitimating system for the deployment of AFR by the state.

A. The 'watchlist'

Faces within a crowd will be cross-referenced against a prepopulated 'watchlist' of individuals during an AFR operation. The facial geometry of identifiable information is collected from the individual and will then be aggregated against an image on the watchlist made up of specific identifiable faces. This will subsequently generate new data about an individual i.e. whether or not that "face" is of interest to the authorities.[167] If no match occurs, the system will delete the data.

In *Bridges*,[168] the Divisional Court eloquently situated how watchlists are generated from images held on police databases, a practice that is part of "ordinary" policing tactics.[169] In our discussion on mass surveillance, we allude to the notion that compiling watchlists is within the common law powers of the police.[170] For AFR, the SWP would create bespoke watchlist of suspected individuals that may be present at the deployment, anywhere between four-hundred to eight-hundred people.[171] Importantly, the Divisional Court described how the SWP had categorised individuals into separate watchlists, for example, "persons wanted on warrants", "individuals who are unlawfully at large" or "vulnerable persons" – to name a few.[172] Therefore, the watchlist acts as the identifier of faces.

Ed Bridges, the Claimant, was not on the watchlist. In fact, the Divisional Court made the claim that Ed is "not a 'victim' in this regard, and therefore can

[165] McKenzie Wark, 'Maurizio Lazzarato: Machinic enslavement' in Wark McKenzie (ed.), *General Intellects: Twenty-One Thinkers for the Twenty-First Century* (Verso 2017) 77–92, 78-79.
[166] Celis Bueno (n 8) 83.
[167] Joe Purshouse and Liz Campbell, 'Privacy, Crime Control and Police Use of Automated Facial Recognition Technology' (2019) 3 Criminal Law Review 188-204, 208.
[168] *Bridges* (n 2).
[169] ibid 30.
[170] ibid 77. See above (n 81).
[171] ibid 30-31.
[172] ibid 30.

have no personal complaint about the watchlists".[173] The initial decision detailed that there was "no minimum threshold of seriousness for the types of offences" to be included on the watchlist, subject to overarching provisions on proportionality.[174] And yet, there remained a general lack of transparency in terms of which individual faces will be selected for the watchlist, i.e. who are the police targeting.

Today, the Court of Appeal, unlike the Divisional Court, has perhaps given a greater degree of importance to the notion of the individual face: "who" might be categorised on the watchlist. For example, the Court of Appeal recognised that "persons where intelligence is required" was not an objective category as it could cover "anyone who is of interest to the police".[175] In their words, the "fundamental deficiencies" in the legal framework concerned 'the "who" and "where" question'. The judgment suggested that perhaps: "once the 'who' question can be satisfactorily resolved, that will give clear guidance as to the 'where' question".[176] In relation to both questions, the Court of Appeal expressed concern that there is "too broad a discretion vested in the individual police officer to decide who should go on the watchlist".[177]

In a study on the politics of "blacklisting individuals", Margaret Hu suggests that "matches and mismatches" in watchlist systems ultimately creates an inferred guilt because it can "categorise individuals as administratively 'guilty until proven innocent' by virtue of [a] digitally generated suspicion".[178] Importantly, if facial recognition identifies a possible match between a face and the watchlist image, it will be up to the "system operator", for exa the police, to establish whether a match has in fact occurred.[179] The Divisional Court had suggested that "the human eye" acts as a vital safeguard "to ensure that an intervention is justified".[180] However, whether it will ever be a safeguard that is truly objective is somewhat open to debate, particularly when an officer is required to act in real-time to a finding provided by the algorithm.[181] In many ways, it raises the possibility for discrimination, which led the Court of Appeal to agree that the issue "ought to be considered properly […] as human beings can also make mistakes. This is particularly acknowledged in the

[173] ibid 77.
[174] ibid 31.
[175] *Bridges* (n 3) 124.
[176] ibid 96.
[177] *Bridges* (n 3) 124.
[178] Margaret Hu, 'Big Data Blacklisting' (2016) 67 Florida Law Review 1735, 1744.
[179] ibid 33.
[180] ibid.
[181] Kotsoglou and Oswald (n 111) 88.

context of identification".[182] The Court concluded that police forces must take all "reasonable steps" to ensure the software does not have a racial or gender bias.[183]

Interestingly, Jackie Wang writes that police forces in general have recently begun to adopt the logic of "objectivity" and "science" to respond to their critics. She argues this is a clever way to take away agency from individual officers and show the police is being "neutral, unbiased and rational".[184] Objectivity is an essential way to "retain public support […] [and] solve the police's crisis of legitimacy".[185] Nonetheless, the belief that facial recognition algorithms create "objectivity" when they identify persons of interests in real time overshadows the issue that this technology can reproduce, or even intensify, the biases that already exist in modern day policing.[186] As Gutherie Ferguson has said: "the danger, of course, is that human consequences get subsumed in the quest for technological guidance".[187] The visual artist Trevor Paglen adequately captures this feeling when he wrote: "because image operations […] are not (actually) dependent on a human seeing-subject […] they are harder to recognise for what they are: immensely powerful levers of social regulation that serve specific race and class interests while presenting themselves as objective".[188]

B. Beyond the face, the future for the individual

From the above discussion, it is worth asking whether the individual can ever be the 'victim' of facial recognition technologies. I pose this question because it previously seemed highly dependent on whether or not your face was on the watchlist. Today, Ed Bridges has shown that if you are not on the watchlist, there will be safeguards in place to challenge your privacy rights. However, what happens to those individuals whose face is on a watchlist? We know that the Court of Appeal accepts it will not "design" the future policies on it,[189] making it a matter of time before new regulation comes in.

The 'risk' of being on an AFR watchlist is likely to be higher for people with minor or previous convictions. The risk is intensified because of the general legal safeguards that currently oversee the collection and retention of custody images by the police, as this can include pictures of people that may have had contact with

[182] *Bridges* (n 3) 173, 185.
[183] ibid 181, 201.
[184] Wang (n 33) 236-237.
[185] ibid.
[186] Purshouse and Campbell (n 153) 200. See also, Kotsoglou and Oswald (n 11) 88.
[187] Ferguson (n 56) 198.
[188] Paglen, Trevor, 'Invisible Images (Your Pictures Are Looking at You)' (*The New Inquiry*, 9 December 2016) <https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/> accessed 28 July 2021.
[189] *Bridges* (n 3) 94.

the police but were then never convicted. In *Catt*, the Supreme Court held that the retention of personal data by the police for someone with a "clean record" can still be justified for three key reasons – all related to allowing the police to take decisions on what makes individuals a public safety risk.[190] In Lord Sumption's words, it is difficult for the police to determine whether a "piece in the jigsaw is irrelevant". He continues: "the most than can be done is to assess whether the value of the material is proportionate to the gravity of threat to the public".[191] *Catt* cautions that it is not for the courts to interfere with how the police assess risk. In *Bridges*, the Court of Appeal held that though AFR is not analogous to taking photos or CCTV, *Catt* retained importance: "just as the human eye can observe a person in a public place […] the police have the power to take photographs of people".[192] The decision, in fact, implies that the general standard for being placed on a watchlist could remain relatively low.

Interestingly, at the *Bridges* appeal, the Appellant's barrister, Dan Squires QC, said AFR would "radically" alter the way Britain is policed in the future.[193] Earlier, we alluded to the idea that a key fear is the rise of predictive policing and the effects it has on the individual. The extreme is a situation whereby in "marking subjects as potential risks, they are actually produced as such".[194] Civil liberty groups, including Liberty, continue to remind us that it will be marginalised communities that will be the most affected in society.[195] We have begun to see these trends in countries like China, with great controversy surrounding the Muslim Uighur communities in the Xinjiang province. This region has been labelled as the "Frontline Laboratory for Surveillance".[196] Specifically, Human Rights Watch reports that facial recognition and other similar technologies are being used by Chinese government to "generate lists of individuals to be rounded-up by the police".[197] The result is to: "bolster its repression of the Muslim minorities […] by tracking virtually their every move, subjecting them to mass arbitrary

[190] *Catt* (n 95) 29.
[191] ibid 31.
[192] *Bridges* (n 3) 84.
[193] Bowcott (n 18). Court of Appeal Trial, R (*Bridges) v CC South Wales & ors* (C1/2019/2679) N.p. 2020. Web. 23-25 June 2020.
[194] Wang (n 33) 43.
[195] Carlo (n 10) 41.
[196] ibid 36.
[197] Human Rights Watch, 'China: Big Data Fuels Crackdown in Minority Region: Predictive Policing Program Flags Individuals for Investigations, Detentions' (*Human Rights Watch*, 26 February 2018) <https://www.hrw.org/news/2018/02/26/china-big-data-fuels-crackdown-minority-region> accessed 28 July 2021.

detention, forced political indoctrination, restrictions on movement, and religious oppression".[198]

It is important to draw attention to the situation in China because even though some might repeat the mantra: *if you have nothing to hide, you have nothing to fear*, the value of understanding the breadth of this technology will ultimately allow the individual, and therefore the collective, to "recognise the pervasive aspect of […] regimes of power"[199] – perhaps like Ed Bridges, and his legal team, did in the UK.

VI. Conclusion

This article began by illustrating the stories of two individuals: Ed Bridges and Robert Williams – both impacted by facial recognition in their own way. These are not unique stories, nor do they tell us everything that we need to know about this technology, but they can help us understand why the individual has a problematic relationship with a tool like facial recognition. However, the Court of Appeal's decision also shows that an individual, in this case Ed Bridges, can act to resist AFR.

In light of the continual expansion of AFR as the technology improves, it is useful to theorise on the mechanisms of power that sit beneath the notion of the individual within the context of facial recognition. Using Deleuze, in particular, we traced the idea that digital technologies no longer sustain the individual as the product of self-identity, as shown under Foucault's description of disciplinary modes of power. Instead, the individual is being replaced with the notion of the "*dividual*" – he/she is reduced merely to observable data.

We must remember that *Bridges* is the first case in the world adjudicating on AFR, which means the debate is still taking shape, and therefore, often appears heavily contradictory. As Celis described, facial recognition technologies have this dual inconsistency: on the one hand the position of the individual is weakened, but on the other, we also must recognise that AFR also means the ever-growing centrality of the face as a mechanism of individualisation.[200]

This article considered the contradictory duality in Celis' argument through a legal context. First, we explored how the law responds to the way AFR has increased the "potential" for mass surveillance by the state, and why this leads to a problematic or "weakening" role for the individual. We found that the courts will use a procedural approach to respond to surveillance methods to allow the state to determine what might be "necessary in democratic society". This is beneficial

[198] Maya Wang, 'The Robots Are Watching Us' (*Human Rights Watch*, 6 April 2020). <https://www.hrw.org/news/2020/04/06/robots-are-watching-us> accessed 28 July 2021.

[199] Celis Bueno (n 8) 88.

[200] Celis (n 58) 81.

to the courts given the risk of adjudicating on seemingly politicised decisions. However, procedural requirements for facial recognition simply "set the conditions for surveillance to occur, which will normalise tracking and identification, as well as reorganise and entrench organisational structures and practices"[201] – all of which intensify the "weakening" of the individual. Second, we reviewed why facial recognition has simultaneously created a situation whereby algorithms develop a "securitised identity" for the individual through their faces – a process that has in some ways become legitimised through the parameters of the watchlist. In fact, a unique characteristic of facial recognition technology is the watchlist, which carries with it a capacity to both categorise and amplify structural issues in society.

The global Covid-19 pandemic and the Black Lives Matter protests in 2020 (and beyond) have shown how quickly debates on facial recognition can be reignited, even in favour of a complete ban. Interestingly, various civil liberty groups in the UK, including Liberty or Big Brother Watch, campaign to ban AFR.[202] Arguably, this debate has become more pertinent since various conglomerates, including Amazon or Microsoft, declared a moratorium on the sale of AFR technologies to law enforcement bodies in the US in the wake of George Floyd's death.[203] With this in mind, it might be interesting to consider what legal arguments exist that that could enable a ban on this technology, particularly as Robert Williams, the man wrongfully arrested because of a fault in the system, seeks a ban in the US.[204]

Having said that, it is interesting to note that Liberty, in advocating for Ed Bridges at trial, falls short in asking the Court to issue a ban. In their submissions before the courts they write: "The Claimant does not suggest that the Defendant or other police forces could never lawfully deploy AFR".[205] In fact, the legal arguments in the courts often evolve around the need to develop legal regimes to

[201] Evan Selinger and Woodrow Hartzog, 'The Inconsentability of Facial Surveillance' (2020) 66 Loyola Law Review 101, 117.

[202] Liberty, 'I Resist Facial Recognition' (*Liberty Human Rights*, 2020). <https://www.libertyhumanrights.org.uk/campaign/resist-facial-recognition/> accessed 28 July 2021. See also, Big Brother Watch, 'Stop Facial Recognition' (*Big Brother Watch*, 2020) <https://bigbrotherwatch.org.uk/campaigns/stop-facial-recognition/>.

[203] Paul Kari, 'Amazon to ban police use of facial recognition software for a year' (*The Guardian*, 11 June 2020) <https://www.theguardian.com/technology/2020/jun/10/amazon-rekognition-software-police-black-lives-matter> accessed 28 July 2021; BBC, 'IMB abandons "biased" facial recognition tech' (*BBC News*, 9 June 2020) <https://www.bbc.co.uk/news/technology-52978191> accessed 28 July 2021; Louise Matsakis, 'Amazon Won't Let Police Use Its Facial-Recognition Tech for One Year' (*Wired*, 10 June 2020) <https://www.wired.com/story/amazon-facial-recognition-police-one-year-ban-rekognition/> accessed 28 July 2021; Karen Weise 'Amazon indefinitely extends a moratorium on the police use of its facial recognition software' (*The New York Times*, 18 May 2021) <https://www.nytimes.com/2021/05/18/business/amazon-police-facial-recognition.html> accessed 28 July 2021.

[204] ibid (n 6).

[205] Squires QC (n 85) 2.

protect privacy and data protection rights. In this way, both the political and social struggles for privacy and data protection are to a certain degree "losing strategies:" they tend to favour a highly procedural method instead of asking why collecting mass data is a necessity in the first instance, an approach that perhaps is better suited to understanding the position of the individual within the facial recognition debate.

For these reasons, we should recognise that the Court of Appeal's judgment paves the way for new regulation and Parliamentary scrutiny, which may indicate why police forces did not want to dispute the judgment further.[206] In fact, at the time of writing, the Police, Crime, and Sentencing Bill is making its way through the Commons, and although there is no specific law on facial recognition, Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS), a body responsible for assessing effective policing (amongst other things), has published a report outlining a '"need to develop" covert intelligence gathering methods and an expectation of increased use of facial recognition technology.[207]

Shortly after the Court of Appeal's judgement was published, the SWP noted that they would remain "completely committed to its careful […] deployment" stating they were "proud of the fact there has never been an unlawful arrest"[208] – and labelled the final decision as "a judgment that we can work with".[209] By contrast, the London Metropolitan Police publicly stated that whilst the judgment will be taken into consideration, London's policing needs are "different" from the issues raised in the appeal against the SWP. To their mind, so long as their use of the technology is "intelligence-led" it can be used as a tool to fight "serious crime" in the capital, an approach that could also bring about future legal challenges.[210] And yet, for all its weaknesses, the *Bridges* judgment still demonstrates the paramount

[206] Keenan (n 144) 19; Dan Sabbagh, 'South Wales Police lose landmark facial recognition case' (*The Guardian,* 11 August 2020) <https://www.theguardian.com/technology/2020/aug/11/south-wales-police-lose-landmark-facial-recognition-case> accessed 28 July 2021.

[207] Haroon Siddique, 'Civil liberties groups call police plans for demos an "assault" on right to protest' (*The Guardian*, 11 March 2021) <https://www.theguardian.com/law/2021/mar/11/civil-liberties-groups-call-police-plans-for-demos-an-assault-on-right-to-protest> accessed 28 July 2021.

[208] Jenny Rees, 'Facial Recognition use by South Wales Police ruled unlawful' (*BBC*, 11 August 2020) <https://www.bbc.co.uk/news/uk-wales-53734716> accessed 28 July 2021.

[209] ibid.

[210] Metropolitan Police, 'Live Facial Recognition' (*Metropolitan Police*, 2020) <https://www.met.police.uk/advice/advice-and-information/facial-recognition/live-facial-recognition/> (accessed 28 July 2021).

importance of an independent judiciary – particularly at a time where there are signs it may come under attack from the state.[211]

In many ways, facial recognition serves as a metaphor for the potential of increased surveillance – given that it forms part of our "invisible visual culture".[212] Interestingly, through the coronavirus pandemic we have witnessed quick changes to the monitoring of our everyday lives: the need to "track and trace" or "vaccine passports" was largely unthinkable prior to the pandemic. And yet, as Deleuze eloquently reminds us: "There is nothing to fear or hope, but only to look for new weapons".[213] To that end, we need to reflect on how the law can work to achieve much greater transparency to support the individual, rather than allow AFR – and technology in general – to become a tool that develops a perpetual state of surveillance.

[211] Tom Clark and Alex Dean, 'Judges in the dock: the inside story of the battle for Britain's Courts' (*Prospect*, 24 January 2020) <https://www.prospectmagazine.co.uk/magazine/judges-in-the-dock-battle-britain-courts-boris-johnson-prorogation-supreme-court-hale-miller-constitution> accessed 28 July 2021. See also, Editorial. 'The Guardian view on Boris Johnson in court: Brexit's war on the law' (*The Guardian*, 12 February 2020); Haroon Siddique, 'Judicial review changes will make government "untouchable", warns Law Society' (The Guardian, 30 April 2021) <https://www.theguardian.com/law/2021/apr/30/judicial-review-changes-will-make-government-untouchable-warns-law-society> accessed 28 July 2021.

[212] Paglen (n 186).

[213] Deleuze (n 32) 4.

Cambridge Law Review (2021) Vol VI, Issue ii, 119–152

Factortame-like Judicial Statute Disapplication and Dicey's Constitutional Orthodoxy: A Case for their Mutual Compatibility

Vincent LaFortune*

Abstract

In the second *Factortame* case, the Appellate Committee of the House of Lords disapplied part of an Act of Parliament that infringed Community law. In 1996, HWR Wade famously argued that that decision had engendered a revolution in the traditional doctrine of Parliamentary sovereignty: such judicial intervention, according to the constitutional scholar, was incompatible with the orthodox constitutional arrangement of the United Kingdom as authoritatively captured by AV Dicey. In this article, I argue that if Wade's analysis of the case still holds such sway today, it is because it has not been met with the rebuttal it, in fact, deserves. This article contends that, with all the necessary ingredients, *Factortame*-like judicial statute disapplication in virtue of an earlier statute is well within the boundaries of an orthodox Diceyan conception of Parliamentary Sovereignty. In other words, that provided all the necessary conditions are met, a court, in disapplying an Act of Parliament or part thereof, does not commit either of the two offences that immediately spring to the reader's mind when reading the *Factortame (No. 2)* decision, namely (i) that of subverting Parliament's legislative sovereignty by force of common law encroachment or (ii) that of allowing an earlier Parliament to bind its successor through manner-and-form hurdles. To achieve this objective, I formulate

* LLB candidate (University College London). My heartfelt gratitude goes to Professor Ronan McCrea (University College London) for his precious guidance in the writing of this article, and to Professor Philip Schofield for his constructive comments on early drafts. Insightful comments from anonymous reviewers were also instrumental in strengthening the piece, for which I am grateful. vincent.lafortune.19@ucl.ac.uk.

a new definition of "constitutional statute" and argue for a reconceptualization of Parliament's temporality. These two arguments, which I respectively label the *technical* and *constitutional* arguments, fuse together to show that a pristine Diceyan conception of Parliamentary sovereignty enjoys more expansive bounds than previously thought, so as to even encompass judicial disapplication of an Act of Parliament or part thereof in virtue of an earlier statute, when a particular set of conditions obtain.

Keywords: Factortame, statute disapplication, constitutional statute, constitutional orthodoxy, parliamentary sovereignty

I. Introduction

In his 1996 article "Sovereignty: Revolution or Evolution?",[1] HWR Wade characterized the *Factortame (No. 2)*[2] decision as a revolution in the "traditional [Diceyan] doctrine of Parliamentary sovereignty" because, by disapplying a provision in the Merchant Shipping Act 1988 in virtue of a provision in the earlier European Communities Act 1972, "the House of Lords elected to allow the Parliament of 1972 to fetter the Parliament of 1988".[3] Wade's claim is a bold but sensible one because judicial disapplication of an Act of Parliament seems *prima facie* anathema to the constitutional framework of the United Kingdom as authoritatively captured by Dicey.[4] It is also an appealing claim, not least because it has not yet been met with the rebuttal it in fact deserves. Indeed, all attempts to rationalize the decision systematically cede, however unwittingly, the most crucial terrain to Wade, namely the decision's incompatibility with Dicey's idea of the workings of Parliamentary Sovereignty in this jurisdiction.

This article contends that with all the necessary ingredients, *Factortame*-like judicial statute disapplication in virtue of an earlier statute is well within the boundaries of an orthodox Diceyan conception of Parliamentary Sovereignty. In other words, that provided all the necessary conditions are met, a court, in disapplying an Act of Parliament or part thereof, *does not* commit either of the two offences that immediately spring to the reader's mind when reading the *Factortame (No. 2)*[5] decision, namely (i) that of subverting Parliament's legislative sovereignty by force of common law encroachment or (ii) that of allowing an earlier Parliament

1 HWR Wade, 'Sovereignty: evolution or revolution?' (1996) 112 LQR 568.

2 *R. v Secretary of State for Transport Ex p. Factortame Ltd (No. 2)* [1990] UKHL 7, [1991] 1 AC 603.

3 Wade (n 1) 574.

4 AV Dicey, *Introduction to the Study of the Law of the Constitution* (first published 1885, Macmillan 1915), 3-4.

5 *Factortame* (n 2).

to bind its successor through manner-and-form hurdles. If these contentions are accepted, hitherto controversial case law such as the decision in *Factortame (No. 2)*,[6] can be safely reframed as perfectly Dicey-compliant. Moreover, incidental conclusions of almost equal importance will be arrived at on the way to this main one. A *technical argument* and a *constitutional argument* fuse together to support this thesis.

The *technical argument* puts forth a brand-new definition of the key ingredient for judicial statute disapplication to obtain, namely a constitutional statute. Necessary therefore, in the first place, is a review of the many different accounts currently existing on the nature of these somewhat perplexing special statutes. Secondly, in rejecting all currently available rationalizations of constitutional statutes, my definition signals a clear break from the current paradigm where constitutional statutes are said to be shielded from implied repeal *because* they boast some especial normative worth entitling them to constitutional status. I show why this irremediably puts the cart before the horse and argue that, instead, some statutes are of a constitutional nature *because* they are linguistically shielded from implied repeal. This enables me to show, *contra* prior explanations, that there is nothing controversial, from a Diceyan perspective, in a constitutional statute investing the courts with disapplication powers. With this redefinition of constitutional statutes, I am equipped to deflect contentions that *Factortame (No. 2)*[7] marked a fundamental shift in constitutional orthodoxy, something which other attempted counters to Wade's revolution thesis have failed to accomplish.

Next, the *constitutional argument* erects fortifications around the *technical argument* in anticipation of jurisprudential siege. It argues for a radical departure from the *idée reçue* that "a Parliament" is one unique sovereign unit established at every general election (in following the conventional characterisation, it is assumed here that when the composition of the House of Commons changes, a new Parliament is formed). Rather, I argue that Parliament, as the supreme and untrammelled legislator under a Diceyan conception of sovereignty, is *one temporally continuous entity*, and has been so since its inception. This argument acts as a prophylactic against the most obvious attack to which the *technical argument* exposes itself, namely a "constitutional-orthodoxy-flavoured" rebuke of the premises upon which it (the *technical argument*) stands. Furthermore, the *constitutional argument*'s importance is signified by the light it shines on a core aspect of the United Kingdom's constitution, namely the foundational principle of the separation of powers between the legislature and the executive.

Thus, the *technical* and *constitutional* arguments come together to show that a pristine Diceyan conception of Parliamentary sovereignty enjoys more expansive

[6] ibid.

[7] ibid.

bounds than previously thought, so as to even encompass judicial disapplication of an Act of Parliament or part thereof in virtue of an earlier statute, when a particular set of conditions obtain.

An important caveat must be borne in mind throughout this article: my thesis is not to be read as an attempt to salvage, or make the apologia of, Diceyan orthodoxy. It might well be the case that Parliamentary Sovereignty cannot today realistically be captured by Dicey's characterisation. Be that as it may, the aim here is to show that, however true it is that Diceyan orthodoxy is nowadays untenable in light of the constitutional landscape of the 21st century,[8] cases like *Factortame (No. 2)*[9] ought not to be taken as having contributed to this constitutional paradigm shift, for what occurred in that case is well within Diceyan comfort zone. A full-fledged defence of Diceyan orthodoxy and of its suitability to modern constitutional realities is a matter for much more voluminous endeavours. This is why this article assumes, from the get-go, that Diceyan orthodoxy is the undisputed norm.

The inaugural section of this paper does not set out to demonstrate anything new and yet it is important because it sets the scene for what comes next. It asserts a trite Diceyan proposition, namely that Parliament's sovereignty also entails its only weakness: that it may not, whatever linguistic brio is conjured up by the parliamentary draftswoman, bind "successive Parliaments" (or as I will prefer later under the *constitutional argument*: "bind itself"). It is this fundamental corollary of a pristine Diceyan conception of sovereignty that critiques of *Factortame (No. 2)*[10] would ultimately (that is, even if they accepted the *technical argument*) contend the case flouted and it is those contentions that the *constitutional argument* will meet. Once that "unfetterability" principle has been reiterated, and the Diceyan bedrock of this paper thus settled, I then proceed with the *technical argument*, which rebuts contentions of common law encroachment. The article concludes with the *constitutional argument*, whose function is to shield the *technical argument* from contentions that it flouts the unfetterability principle.

A final word of recapitulation. It will hopefully be apparent to the reader that I am embarking on two distinct albeit interconnected missions. The first mission is simply to rescue constitutional statutes from the brink. They currently sit in both a precarious and a contentious position because they are not rationalised as products of Parliament but as spawn of the common law, thus exposing their flanks to two charges. The first charge is that constitutional statutes are the result

[8] See, for example, Lord Steyn in *Jackson and Other v Her Majesty's Attorney General* [2005] UKHL 56, [2006] 1 AC 262, at [102]: "The classic account given by Dicey of the doctrine of the supremacy of Parliament, pure and absolute as it was, can now be seen to be out of place in the modern United Kingdom".

[9] *Factortame* (n 2).

[10] ibid.

of courts imbuing some statutes with a superior status according to the judge's own normative inclinations. This in turn makes the judiciary prone to all sort of criticisms, such as that of usurping the legislative function of Parliament. The second charge is that it follows from the common law source of constitutional statutes (that is, under the current rationale) that judges with different sympathies towards the notion of constitutional statutes could do away with them to "restore" orthodox Parliamentary Sovereignty. By repositioning constitutional statutes within the traditional Diceyan framework, I prophylactically address these critical dangers facing constitutional statutes. Indeed, constitutional statutes can be salvaged and rebranded as Parliament's own doing. The second mission is to shield those freshly rationalized constitutional statutes, especially those which grant disapplication power to the courts, from the contention that they fall foul of the rule, which flows from the traditional Diceyan notion of Parliamentary sovereignty, that Parliament cannot legislatively shackle itself. At the end, I hope the reader will accept *Factortame (No. 2)*[11] as a Dicey-compliant decision and that she will adopt the various incidental conclusions arrived at on the way there.

II. Why Is This Important?

At this point, the reader might query about the relevance of the extensive discussion that unfolds below. The decision of the House of Lords in *Factortame (No. 2)*, one might argue, belongs to truly bygone political and legal epochs. Nevertheless, let me make the case for the relevance of fostering the discussion to which this article seeks to contribute. Firstly, a remote, yet not totally irrelevant, reason to have this discussion might be that re-joining the EU in the (however distant) future is not so outlandish a prospect that the constitutional implication of such a choice should simply be relegated to oblivion. Secondly, and of more tangible import, there is no denying that the United Kingdom's constitution is to some extent an elusive concept which benefits greatly from spirited debate. The following discussion strikes at the heart of this debate by proposing new ideas about the nature of the entity that is Parliament as well as reconsidering and reframing the flexibility that it enjoys in its enactments. Thirdly, this article delves at length on constitutional statutes, a topic of great importance whose relevance is undeniably contemporary. Fourthly, and finally, there are important conclusions to be drawn

[11] ibid.

from this discussion on the fundamental principle of the separation of powers between the legislative and the executive.

III. Parliament's 'Unfetterability'

As mentioned in Section I above, before embarking on my main task I must briefly look at some corollaries a Diceyan understanding of Parliamentary sovereignty entails. At the risk of repeating myself, the derivatives established below simply reflect the commitment of this paper to Diceyan orthodoxy: they are taken for the sake of argumentation, for as mentioned above, I do not commit to such constitutional view, but simply strive to show that Diceyan orthodoxy is not threatened by occurrences of judicial statute disapplication of which the *Factortame (No. 2)*[12] case is the most poignant example.

As is well-known, the core proposition under Dicey is that Parliament is sovereign and, the case being, its intentions, communicated in writing via Acts of Parliament, are necessarily beyond dispute. Necessarily then, this paper adheres to the view that "Parliament […] has the right to make or unmake any law whatever; and further, that no person or body is recognised by the law of England as having a right to override or set aside the legislation of Parliament".[13] From that proposition, many sub-propositions may be derived. Germane to my discussion is the sub-proposition that emerges from considering the merits of the claim that "the legislative authority of an existing Parliament may be limited by enactments of its predecessors."[14]

The proposition that Parliament may bind its "successors" and that consequently the sovereignty of future Parliaments is "qualified and precarious" clearly stands at odds with "all available English authority".[15] In his *Introduction*, Dicey provides two examples of the futility of attempting to bind "future Parliaments": The Union with Scotland Act 1706 and the Union with Ireland Act 1800.[16] Without going into too much detail, it is impossible for Parliament's intention, in 1706 and in 1800, to have been clearer: the conditions upon which the unification of the nations depended were to be entrenched and shielded from any future repeal. Nevertheless, each respective Act saw some of their sacrosanct provisions repealed in the years following their enactment.[17] These two examples alone, being clear yet unsuccessful attempts to shackle "future Parliaments", provide compelling grounds to assert that however forcefully and unequivocally

[12] ibid.
[13] Dicey (n 4) 3-4.
[14] ibid 21.
[15] Wade, 'The Basis of Legal Sovereignty' (1955) 13(2) CLJ 172, 185.
[16] Dicey (n 4) 65.
[17] ibid.

Parliament may convey its desire to lock a piece of legislation in an impenetrable vault and discard the key in the depths of the Great Blue Hole, the fact remains that this very key, through the mystical connivances of *Lady Sovereignty*, always remains at Parliament's disposal.

Yet, analogical reasoning is not entirely satisfactory, and Dicey was aware of this (although I wonder why he restricted himself to providing the logical necessity of the unfetterability of Parliament in the form of a footnote). Indeed, inductively derived conclusions leave us in want of the more reasoned and logical explanation which we would get via the deductive route, that is, an explanation which would readily discard any postulate to the effect that linguistic brio from clever draftswomen could eventually be successful at binding "future Parliaments". Fortunately, such logical explanation is available, and here is Dicey's version of the substantive rationale underpinning the unfetterability principle:

> "The logical reason why Parliament has failed in its endeavours to enact unchangeable enactments is that a sovereign power cannot, while retaining its sovereign character, restrict its own powers by any particular enactment. An Act, whatever its terms, passed by Parliament might be repealed in a subsequent, or indeed in the same, session, and there would be nothing to make the authority of the repealing Parliament less than the authority of the Parliament by which the statute, intended to be immutable, was enacted. 'Limited Sovereignty', in short, is in the case of a Parliament as of every other sovereign, a contradiction in terms".[18]

These are wise, true, and sufficient words indeed, but let us elaborate ever so slightly. Dicey is right in saying that as long as a sovereign is sovereign, she cannot restrict her own powers. This is not a normative proposition or a statement of what the sovereign should not do, but a corollary of what we mean by the noun substantive "sovereign". "Sovereignty" is, to borrow Jeremy Bentham's terminology, a "fictitious entity".[19] A fictitious entity is a noun substantive which, although grammatically talked as if existing in the real world ("sovereignty" is often talked of as "something" one "possesses") does not. Nonetheless, although "sovereignty" does not evoke any material image in the mind, it may be *paraphrased* in so many terms rooted in physicality which themselves elucidate the meaning of the noun. Thus, something or someone is talked of as being invested with sovereignty when that thing or person(s) enjoy omnipotence in a given sphere of activity. For example, Man is sovereign over his own decisions and may not

[18] Dicey (n 4) 24.

[19] Jeremy Bentham, *The Works of Jeremy Bentham, 11 Vols.* (J. Bowring (ed.), OLL 1838–43) viii 3.

"conclude himself":[20] if at this very moment I promise to myself that I will go out and purchase Zola's *Germinal* tomorrow morning, and further utter to myself that I cannot on any pretext come back on my promise, my legs do not automatically start walking towards the bookshop the next morning even though I have since changed my mind. Similarly, to say that Parliament is the legislative sovereign is to say that Parliament enjoys an absolute power to enact whatever law it so desires. If it today stipulates that henceforward apple-picking shall be forbidden on Tuesdays, and that this law shall be entrenched and immutable, and if tomorrow it feels awful about its enactment and admits of its absurdity by providing for its repeal, then by the terms of its sovereign status, it is the most recent will that prevails over the earlier one, however formidable in its formulation the latter is. To suggest the possibility of a "legislatively limited legislative sovereign" is therefore to trade in oxymorons since this phrase is, to recycle Dicey's words, a "contradiction in terms".[21] Such terms are mutually exclusive and cannot co-exist, and so either someone (or something) is sovereign or, however slight the limitation is, is not so.[22] Therefore, as long as Parliament "retain[s] its sovereign character", it must be free to legislate as it wishes. Notice that this leaves open the extremely stimulating question of whether Parliament *should* be able to abandon its sovereignty if it so desired, but this inquiry is moot to our present quest: we simply need to show that the Parliament of the day in 1988 was sovereign and remained sovereign even though it saw one of its enactments disapplied by the House of Lords in 1990. If I succeed in showing this, claims that *Factortame (No. 2)*[23] was a revolution, and more generally that statute disapplication under any circumstances is a violation of Parliament's sovereignty, are deflected. With the hope that my commitment to a full-fledged Diceyan orthodoxy is clear, I now proceed with the *technical argument.*

IV. The Technical Argument: Redefining Constitutional Statutes

This section surveys the prevailing definitions of, and subsequently formulates a new definition for, the main ingredient required for statute disapplication to obtain, namely a constitutional statute.

The term "constitutional statute" may at first introduce unease in someone who has always been taught that in the United Kingdom, Parliament is sovereign and free to legislate as it wishes. Indeed, that term instantly directs the mind to legal systems where written constitutions exist and are supreme, in stark contrast with the way things work in this jurisdiction. But the force of the adjective "constitutional"

[20] Dicey (n 4) 64.
[21] Dicey (n 18).
[22] Dicey (n 4) 68.
[23] *Factortame* (n 2).

here used is, in truth, much tamer. Indeed, constitutional statutes have this sole, but important, distinguishing characteristic that they are immune from implied repeal. To be sure, constitutional statutes are not *enshrined* or *entrenched* or *anchored* like the provisions in the Constitution of the United States are. Constitutional statutes are said to be immune from the operation of repeal by implication but may always be repealed by an express provision from Parliament through the ordinary Parliamentary process.

Considering the foregoing, it is easy to see that since judicial statute disapplication in virtue of an earlier statute necessarily entails a conflict between two statutes, the only way such an event can possibly take place is if the earlier statute has constitutional status. Indeed, were it not for the constitutional status of the earlier statute, the latter would be repealed by implication instead of not only surviving but also disapplying the later statute.

Constitutional statutes were first coined by Laws LJ in *Thoburn*,[24] and since then a substantial amount of ink has been spilt in the quest to uncover the rationale underpinning their special treatment. I begin this section by conducting a survey of the most authoritative explanations of constitutional statutes to date, focusing particularly on the important work of Ahmed and Perry on the topic.[25] After surveying the current landscape, I attempt to show that if any of the currently available theses tendered as providing the best rationalization of constitutional statutes is accepted, then whatever justification within that array of choice one ends up subscribing to, one is forced to drop one's commitment to Diceyan orthodoxy. Indeed, as I shall show, the special status enjoyed by constitutional statutes is currently owed to the constitutional worth of their subject matter i.e., a statute will be shielded from implied repeal if it has some special constitutional or normative weight. It is within the confines of those premises that judges and commentators have attempted to explain why some statutes ought to be immune from implied repeal and they have focused their efforts on establishing the minimum content that a statute should have to enjoy constitutional status. As we will see, at the heart of those explanatory endeavours are value judgments exercised by the courts which are fundamentally at odds with a pristine conception of Diceyan orthodoxy. Thus, if any of the current rationales is adopted, we inevitably concede that the common law has designated a special class of statutes which it shields from

[24] *Thoburn v Sunderland City Council* [2002] EWHC 195, [2002] 3 WLR 247.

[25] See, for example, Farrah Ahmed and Adam Perry, 'The Quasi-Entrenchment of Constitutional Statutes' (2014) 73 The Cambridge Law Journal 514; Farrah Ahmed and Adam Perry, 'Constitutional Statutes' (2017) 37 OJLS 461.

Parliament's intentions, a clear disruption of the constitutional hierarchy and an acknowledgment of the soundness of Wade's revolutionary thesis.

If this article is to prove that Diceyan constitutional orthodoxy is not endangered by judicial statute disapplication, the first step is to deny that constitutional statutes are granted their immunity because they are of some distinctive normative importance. In other words, I must challenge the fundamental premise currently shared by all justificatory arguments. I do so by encouraging a shift away from the current view that "because it is a constitutional statute it may not be impliedly repealed" to "because it may not be impliedly repealed it is a constitutional statute". I occasion this shift by challenging the common conception of the "doctrine of implied repeal" and by showing why it is that constitutional statutes are instead statutes worded in a particular way which immunises them from implied repeal, thereby unquestionably making them products of Parliament rather than of the common law. This is the *technical argument*.

A. Constitutional statutes: the traditional views

(i) Traditional Definitions of Constitutional Statutes

There is extensive literature devoted to coining a definition of constitutional statutes. I shall briefly assess such literature, guided by the precious work of Ahmed and Perry, before considering how the gap is traditionally bridged between the content of a constitutional statute and its distinctive property of being protected from implied repeal.

The only definition given in a judicial context is found in Laws LJ's *Thoburn*[26] judgment: "in [his] opinion a constitutional statute is one which (a) conditions the legal relationship between citizen and state in some general, overarching manner, or (b) enlarges or diminishes the scope of what we would now regard as fundamental constitutional rights".[27] "Ordinary statutes may be impliedly repealed […] constitutional statutes may not".[28] Laws LJ goes on to say that this "special status of constitutional statutes follows the special status of constitutional rights",[29] and that just like fundamental constitutional rights, constitutional statutes are a development of the common law.[30] Since this is the only judicial definition to date, I must give it due consideration and will do so by

[26] *Thoburn* (n 24).
[27] ibid 62.
[28] ibid.
[29] ibid.
[30] ibid 64.

following attentively the incisive work on the subject-matter conducted by Ahmed and Perry.

Ahmed and Perry, authors of the most recent comprehensive academic paper on constitutional statutes, [31] have identified a first flaw with Laws LJ's definition. They argue that it is underinclusive.[32] To illustrate their point, it is a sound practice to consider particular statutes which have come to be recognised as having constitutional status and see whether they fit easily under Laws LJ's definition. It is in conducting such matching exercise that Laws LJ's definition seems to fall short. For example, if the Parliament Acts of 1911 and 1949 are to be deemed statutes of constitutional status, as they usually are, then we must conclude that not all constitutional statutes are about conditioning the relationship between citizen and state, or are concerned with the codification of constitutional rights: sometimes, constitutional statutes shape the relationship between state institutions themselves.[33] From a descriptive point of view, Laws LJ's definition therefore inaccurately captures the breadth of constitutional statutes, which are more varied in nature than what his definition admits.

Since these "counterexamples […] cannot be easily smoothed over",[34] Ahmed and Perry take on the task of tweaking the definition so that it captures the full range of constitutional statutes. Their endeavour begins with an analysis and half-endorsement of David Feldman's proposed definition. The "institutional approach"[35] to the identification of constitutional statutes, according to Feldman, is preferable because it captures statutes that are in line with what is generally understood as the normal meaning and function of a "constitution" : that it should "constitute the state and its institutions and confer functions, powers and duties on them". [36] This definition clearly embraces the Parliament Acts of 1911 and 1949 and thereby seemingly solves the under-inclusiveness of Laws LJ's definition. It seems fair to assert, as Ahmed and Perry do, that Feldman is of the view that it *suffices*[37] for a statute to confer functions, powers, and duties on organs of the state to acquire a higher status and ultimately immunity from implied repeal. At the very least, Feldman does not qualify further the type of functions, power and duties which would be required to render the statute constitutional.

Ahmed and Perry accept that Feldman's definition solves Laws LJ's underinclusiveness problem, but they contend that it is guilty of the opposite effect. By showing how Feldman's institutional approach would achieve the weird

31 F. Ahmed and A. Perry, 'Constitutional Statutes' (2017) 37 OJLS 461.
32 ibid 466.
33 ibid.
34 ibid.
35 David Feldman, 'The Nature and Significance of Constitutional Legislation' (2013) Jul. LQR 129.
36 ibid 350.
37 Ahmed and Perry (n 31) 467.

result of including statutes such as the Coroners and Justice Act 2009 and Crown Prosecution Service Inspectorate Act 2000 ('CPSIA 2000') to the corpus of constitutional statutes, they argue that Feldman's definition is, unfortunately, over-inclusive. Although Ahmed and Perry "agree […] that all constitutional statutes create or regulate state institutions",[38] they qualify this subject matter common to all constitutional statutes as a necessary *but not sufficient* ingredient. This is the insight upon which they build to find the missing ingredient lacking in Feldman's definition.[39]

As anticipated, this missing ingredient is a certain *normative importance* which the statute must possess for it to qualify as a constitutional statute.[40] Be that as it may, "normativity" being the nebulous concept that it is, simply asserting that a constitutional statute is equal to Feldman's definition plus a certain normative weight is rather unhelpful. Ahmed and Perry suggest that the right question to ask in determining whether a statute passes the "normative notoriety" threshold is what would be the effect of that statute's repeal on the state's governing *abilities* or *disabilities*: "A statute might be of great direct influence because it confers wide-ranging powers, for example, or because it makes unlawful large swathes of state action".[41] Their definition of a constitutional statute thus runs as follows:

> "A constitutional statute is a statute at least a part of which (1) creates or regulates a state institution and (2) is among the most important elements of our government arrangements, in terms of (a) the influence it has on what state institutions can and may do, given our other governing norms, and (b) the influence it has on what state institutions can and may do through the difference it makes to our other norms".[42]

Although this definition shares similarities with Feldman's institutional approach, it additionally requires the statute to boast normative importance as defined in 2(a) and (b). With this latter criterion, Ahmed and Perry are able to suggest that statutes which condition the relationship between state institutions or that create such institutions, but that are not inherently of "great importance" like the CPSIA 2000 will not risk being included in the select category of implied-repeal-shielded constitutional statutes.[43] In that sense, they avoid Feldman's over-

[38] ibid.

[39] ibid.

[40] In that sense, Ahmed and Perry approve of Paul Craig's insight on the matter in Paul Craig, 'Constitutionalising Constitutional Law: *HS2*' (2014) PL 373.

[41] Ahmed and Perry (n 31) 469.

[42] Ahmed and Perry (n 31) 471.

[43] ibid 472.

inclusiveness problem. Importantly, this normative criterion is not grounded on some fundamental common law principle that the statute in question seeks to promote, but rather is determined through the analysis of the statute's "direct and indirect importance" regarding its influence on "governing norms".[44] That way, Ahmed and Perry's definition avoids Laws LJ's under-inclusiveness problem.

In as far as Ahmed and Perry's objective was to provide a definition that captured all currently recognized constitutional statutes – whether judicially, academically, or what may come to the same, intuitively identified[45] – they seem, at first glance, to have coined a working definition which best does justice to the variety of constitutional statutes routinely recognised. Nevertheless, one burning question remains: what is it exactly, in the broader context of the doctrine of Parliamentary sovereignty and the corollaries which flow from it, that justifies the higher status ascribed to these statutes? How do we get from the insightful but otherwise benign observation that some statutes boast superior normative qualities to justifying their special treatment with regards to implied repeal?

(ii) The Traditional Justifications for The Privileged Treatment of Constitutional Statutes

Coming up with a definition which aims at capturing all recognized constitutional statutes is one thing, but it is quite another to justify the immunity to repeal by implication granted to those statutes falling under said definition. As will be shown, because their justification arguably initiates a paradigm shift *away* from the orthodox legal justifications for the special treatment of constitutional statutes, it is at the stage of justifying the special treatment enjoyed by constitutional statutes that Ahmed and Perry's analysis is most compelling. Indeed, as Ahmed and Perry themselves state towards the end of the section concerned with formulating their definition of constitutional statutes (with which we were concerned in Section IV.A.(i)): "Perhaps most importantly, our definition accounts for how constitutional statutes ought to be treated for the purposes of implied repeal […]".[46] Although this seems to augur well for Diceyan orthodoxy, because Parliament has a central place in their justification, it will be submitted that, despite being more nuanced, their approach still boasts the same fundamental flaw inherent in all other commentators' works and is, if accepted, fatal to Diceyan orthodoxy. But first, let's

[44] Which distinguishes their approach to normative importance from that of David Jenkins: David Jenkins, 'Common Law Declarations of Unconstitutionality' (2009) 7 International Journal of Constitutional Law 183-201.

[45] Ahmed and Perry (n 31) 466: "The Parliament Acts 1911 and 1949, for example, *strike us (and others) as obvious examples of constitutional statutes* (emphasis added)".

[46] ibid 473.

briefly survey the traditional justifications for granting some statutes, once their "constitutional" worth identified, an immunity to implied repeal.

It is remembered that for Laws LJ, constitutional statutes are those that codify fundamental common law rights. Said rights are greatly valued by the common law and, necessarily, by common law judges whose task it is to protect them. Somewhat fortuitously, Laws LJ also deems the common law to be the originator of the rule of implied repeal (this view will be disputed in Section IV.B.(i) below). It follows that the courts are not only justified in, but entrusted with, tweaking this rule (i.e., shielding some statutes from an otherwise clear case of implied repeal) in their quest to protect and uphold common law constitutional rights as far as possible. While for Feldman constitutional statutes are not limited to those codifying common law constitutional rights, his justification for their special treatment by the courts is akin to that of Laws LJ's. Along with other reasons that are not central to our discussion, Feldman attributes the impossibility of impliedly repealing constitutional statutes to, inter alia, "the risk of such an important law being amended or repealed without proper consideration in Parliament".[47] Thus, in both Laws's and Feldman's justifications there is an appeal to the statute's importance to justify its immunity from the operation of implied repeal. To be sure, the courts, after having identified a particular statute under either Laws's or Feldman's definition, should shield them from implied repeal because those definitions identify, in their own ways, statutes of great importance.

Ahmed and Perry voice two powerful criticisms of these legal justifications for granting immunity to constitutional statutes from the operation of implied repeal. The first is that, as seen previously, "not all constitutional statutes have one of [the] two subject matters" to which Laws LJ refers in his definition.[48] Thus, what justification are we to give to current constitutional statutes outside of the scope of Laws's definition such as the Parliament Acts? Indeed, on Laws LJ's account, the common law should not be preoccupied with shielding the Parliament Acts and other statutes which regulate institutional matters. Ahmed and Perry thus contend that his Lordship's definition does not do justice to the breadth of constitutional statutes that are today recognized with the result that his justification for the common law's intervention in tweaking the rules of implied repeal only captures those that do conform to his definition. This is unsatisfactory for, surely, we need a uniform justification that applies to all constitutional statutes.

In my view though, it is the second flaw with Laws's justification as identified by Ahmed and Perry that truly highlights its unsatisfactoriness. And although Ahmed and Perry think it a "less obvious objection", I submit that this is in fact an

[47] Feldman (n 35) 352.

[48] Ahmed and Perry (n 31) 474.

obviously fatal objection which is of peculiar importance. The following passage has my full endorsement, and it is worth quoting it at length:

> "[…] there is no plausible understanding of parliamentary sovereignty that would allow judges to unilaterally impose new limits on Parliament's powers, or to assign a meaning to a statute at odds with the one Parliament intended. Laws LJ's argument that judges should modify the doctrine of implied repeal requires them to do both: it requires them to limit Parliament's power to make legal change by implication and to assign a meaning to statutes based partly on the desirability of preserving fundamental rights, rather than honouring parliamentary intent".[49]

Although Ahmed and Perry do not address it, the same could be said of Feldman's justification: it requires judges to limit Parliament's power to repeal by implication based on the reticence of impliedly repealing statutes that, due to their conferring powers and imposing duties on state institutions, are of some especial normative worth.

This criticism from Ahmed and Perry seems to be hinting at an incoming vindication of Diceyan orthodoxy. Indeed, it suggests that within the confines of the United Kingdom's orthodox constitutional framework as sketched by Dicey, there is simply no justification for courts to decide whether to give effect to a clear case of implied repeal on the basis of the statute's promotion of fundamental common law constitutional rights as Laws LJ would have it or, as a matter of fact, on any other ground which purports to distinguish the treatment of statutes according to their content. Such judicial behaviour could arguably be criticized as usurpative of Parliament's legislative sovereignty and is unacceptable if Diceyan orthodoxy is at the core of the United Kingdom's constitutional framework. But, again, I must stress that whether Diceyan orthodoxy still is at the core of the United Kingdom's constitutional framework is not the concern of this paper. What can be asserted, however, is that Laws LJ's justificatory argument does require an abandonment of Diceyan orthodoxy, as Ahmed and Perry also point out.

Having rejected Laws's justification as unsatisfactory, Ahmed and Perry propose a rationale which at first glance avoids the constitutional discomfort produced by those hitherto tendered. Their justification for the special status accorded to constitutional statutes relies on the presumption of consistency according to which, in times of political normality, Parliament may be presumed to legislate consistently with its prior enactments.[50] What is more, the "greater the

[49] ibid.

[50] ibid 475.

weight Parliament can be expected to accord a statute, and the greater the potential disruption of that statute's repeal, the stronger the presumption of consistency will be".[51] A constitutional statute, as it has been defined by Ahmed and Perry, is one to which Parliament may be presumed to have given considerable weight and also one whose "repeal […] would be highly disruptive", especially in times of political normality.[52] It follows that "stronger evidence should be required to show that a later ordinary statute and an earlier constitutional statute are inconsistent than that two ordinary statutes are inconsistent".[53] Were this stronger evidence to be lacking, it is presumed that the later ordinary statute should have its ambit narrowed so as to not impede on the earlier constitutional statute's jurisdiction.

Although a move in the right direction, it is submitted that Ahmed and Perry's justificatory argument for the special treatment of constitutional statutes boasts both different and similar shortcomings as those of Laws LJ and others. Thus, if their justification is adopted, Diceyan orthodoxy would have to be abandoned.

B. Where the traditional rationales all falter

(i) Problems specific to Ahmed and Perry's rationalisation

Firstly, it is unclear what kind of stronger evidence would do to rebut the presumption of consistency and thus to warrant implied repeal of a constitutional statute or part thereof. What is clear, however, is that such evidence must be sourced elsewhere than in the statutes themselves since we are already working under the assumption that there is clear linguistical incompatibility between two statutes. It is possible that what Ahmed and Perry have in mind is for the courts to rely on *Pepper v Hart*[54] to gather the necessary evidence. But *Pepper* allows reliance on Parliamentary material only where "legislation is ambiguous or obscure or leads to an absurdity".[55] Repeal of a constitutional statute is not "absurd" – would we say that the express repeal of a constitutional statute is "absurd"? – and would in any case only occur impliedly if the implication is unequivocal and clear, not "ambiguous or obscure". Thus, it is doubtful whether *Pepper* could at all be relied upon in a case where a constitutional statute is faced with the threat of implied repeal.

Be that as it may, let's be blind to the above criticism and assume that the statute should automatically be deemed "obscure" and "ambiguous" if it threatens

[51] ibid 476.
[52] ibid 477.
[53] ibid.
[54] *Pepper (Inspector of Taxes) v Hart* [1992] UKHL 3, [1993] AC 593.
[55] ibid 640.

a constitutional statute in "times of normality". Arguably, the courts would be looking for some mention of the constitutional statute in question and for some words conveying the intention for its repeal. Then, since we are dealing with evidence, what evidential threshold should be set: "balance of probabilities" or "beyond reasonable doubt"? Ahmed and Perry would most likely argue for the latter, given the status they ascribe to constitutional statutes in times of political normality. Now let's further assume that there is clear and express intention in Hansard that the later normal statute should revoke the earlier constitutional statute. In light of this analysis of Hansard, reasonable doubt seems to dissipate and implied repeal to be justified. This is how I interpret the nature of the evidence that Ahmed and Perry are suggesting would be required for implied repeal to take effect.

But the pressing question is why, if it was discussed and apparently decided, did Parliament not provide explicitly for the repeal of the earlier constitutional statute in the later statute itself? The discrepancy between Hansard and the enacted statute introduces doubt as to whether Parliament's intention was indeed to repeal the constitutional statute after all deliberations were concluded. That being the case, the burden of proof, which requires the absence of any reasonable doubt, is not satisfied. The corollary of this observation is that Ahmed and Perry's approach, on my understanding of it, means that constitutional statutes as identified against their definition may only be repealed expressly, something they earnestly deny. [56] It must be emphasised how important it is for Ahmed and Perry to allow for the possibility of constitutional statutes being impliedly repealed:

> "to say that a statute can only be repealed expressly is thus to say that a statute cannot be repealed despite Parliament's clear intention to do so. That amounts to a limit on Parliament's sovereignty – something which, we have said, judges cannot unilaterally impose".[57]

And yet, in my submission, their methodology leads to the very outcome they were at pains to disavow: it will never be possible for Parliament to repeal a constitutional statute other than expressly because the extra-statutory evidence which must contain express statements to the effect that the constitutional statute with which the new statute conflicts should be repealed, does not make its way into the final wording of the statute itself.

To conclude the above discussion, it is worth pondering some *obiter dictum* by Laws LJ in *Thoburn*. Maybe my gripe with Ahmed and Perry's evidential approach

[56] Ahmed and Perry (n 31) 477.
[57] ibid.

is the same Laws LJ had in mind when he stated that "in my judgment general words could not be supplemented, so as to effect a repeal or significant amendment to a constitutional statute, by reference to what was said in Parliament by the minister promoting the Bill pursuant to *Pepper v Hart*".[58]

(ii) The Overarching Flaw

But there is one overarching, fundamental flaw endemic in all the aforementioned rationales. I shall expose that flaw by showing how it manifests itself in Ahmed and Perry's justificatory argument because that argument poses the strongest resistance to my rebuttal and if it goes, so do all other justifications.

The gist of the fundamental problem is that the justifications studied so far appeal to concepts which do not flow from an orthodox conception of Parliamentary Sovereignty and thus are necessarily fatal to Diceyan orthodoxy. For just like Laws, Feldman and others, Ahmed and Perry rely on a *qualitative appreciation* of the statute's content to justify the bolstered presumption of consistency and to shield it as far as possible from implied repeal. In other words, the courts should use their (Ahmed and Perry's) definition to identify statutes which have an especial normative worth (defined as the statute's direct and indirect influence on governing norms) entitling them to constitutional status, which in turns shield said statutes from an otherwise clear case of implied repeal bar some extra-statutory evidence to suggest that it was indeed Parliament's intention to effectuate the repeal (the problems with this approach are highlighted in Section IV.B.(i) above). In doing so, they inevitably condone judicial questioning of the quality of Parliament's intentions, something which is alien to the very idea of sovereignty where there must be no justification for, nor qualitative assessment of, the legislative sovereign's volition as it is communicated through Acts of Parliament.

It must however be acknowledged that Ahmed and Perry's justificatory argument is, in a very important way, much subtler than the others. Indeed, *contra* Laws LJ, under Ahmed and Perry's view judges do not tweak the doctrine of implied repeal to preserve norms which are sacrosanct to the common law. Rather, the courts intervene on the purported grounds of giving effect to Parliament's intentions: this attempts to shift the rationale from self-interested to almost altruistic judicial intervention. Nevertheless, let's recall Ahmed and Perry's fundamental criticism of Laws LJ's justificatory argument: "[…] there is no plausible understanding of parliamentary sovereignty that would allow judges to *unilaterally* impose new limits on Parliament's powers, or to assign a meaning to a statute at odds with the one Parliament intended (emphasis added)". Does it follow from this that if Laws LJ's justificatory argument would have been to the tune that the common law tweaks

[58] *Thoburn* (n 24) 63.

the doctrine of implied repeal because Parliament "is likely to have appreciated the significance of, and deliberated particularly carefully about, constitutional statutes, because they"[59] give statutory effect to common law fundamental principles, then the only flaw in his Lordship's position would be its rather less fatal under-inclusiveness? Indeed, on that latter entirely plausible reformulation of the common law argument for preserving constitutional statutes from implied repeal, it seems that judges would not anymore be "unilaterally" imposing new limits on Parliament's powers, but rather doing so pursuant to the intentions they ascribe to Parliament itself and, as such, not fall foul of Ahmed and Perry's fundamental criticism. By allowing for such an easy counter, something seems to be amiss with Ahmed and Perry's argument. Indeed, a legal system premised on Parliamentary Sovereignty where the legal interpreters may potentially cloak their normative inclinations under the pretence of giving effect to the sovereign legislator's will is a recipe for trouble. To be sure, judges must of course often "fill in the gaps" in legislation and exercise their discretion as they best see fit when the rules alone fall short of providing the answer. That is indeed a consequence of the "irreducibly open-textured" natural language used by the legislator in shaping the laws.[60] It must however be recalled that we are working under the hypothesis of a clear case of repeal by implication where there is irresistible linguistic incompatibility between two statutes: there is no "penumbra of uncertainty"[61] as far as the operation under scrutiny is concerned and this makes judicial tergiversation unacceptable.

The thought that Parliament did not accord the same weight to an ordinary statute as it did to a constitutional statute however defined may be factually right. That it "appreciated the significance of, and deliberated particularly carefully about, constitutional statutes"[62] more so than other statutes, may or may not be true. But the essential privilege of a legislative sovereign is that the rationale behind its decisions is not to be questioned and that its most up-to-date intentions as gathered from Acts of Parliament are the law. To be sure, in shifting the focus to Parliament's intentions, Ahmed and Perry do contribute substantially towards the emancipation of constitutional statutes from the contentious and frankly controversial view that they are products "of the common law, for the common law", as Laws LJ would have it. Nevertheless, their approach still requires judicial

[59] Ahmed and Perry (n 31) 477.
[60] HLA Hart, *The Concept of Law* (first published 1961, Clarendon 2012) 128.
[61] ibid 134.
[62] Ahmed and Perry (n 31) 477.

interference and involvement of a degree which is at odds with what a Diceyan conceptualization of Parliamentary Sovereignty would be comfortable.

(iii) Recapitulation

In this subsection, I have assessed the most authoritative attempts at coining a definition and providing a justification for constitutional statutes, in particular through the important work conducted by Ahmed and Perry on the matter. The key takeaway from this survey is that, as things stand, constitutional statutes are granted immunity from implied repeal because judges and commentators imbue them with a certain normative worth. The main debate between the different views pertains to what definition best captures constitutional statutes, and how to best rationalize the special treatment they receive.

Yet, the jump from a statute having normative importance to it being shielded from implied repeal is simply unreconcilable with an orthodox Diceyan understanding of Parliamentary sovereignty. Despite attempted tweaks and modified appearances, the stance which had been initially developed by Laws LJ never fundamentally changed. In all the definitions explored above, the fact remains that some statutes are shielded from implied repeal *because* they are constitutional (a term which, as seen, attracts many definitions) and that it would, for various reasons, be harmful to allow implied repeal. In my view, Ahmed and Perry have kickstarted a paradigm shift in the narrative that had until then prevailed regarding constitutional statutes, but I believe that their own justification does not go far enough as it still demands that the courts form their own appreciation of what kind of statute Parliament must have intended to give some special weight to.

As Ahmed and Perry put it, with the rejection of all the above justificatory arguments "comes a choice": either we conclude that "the line of cases" which has systematically recognised new constitutional statutes is misguided "insofar as it suggests there is something special about constitutional statutes with respect to implied repeal", either, less drastically, we conclude that the problem lies with the current justifications, and we try to salvage constitutional statutes with our own.[63] We will of course opt for the latter option, but it is worth considering how precarious the state of affairs is as things currently stand.

First, claims that constitutional statutes currently undermine our traditional understanding of Parliamentary sovereignty are not ludicrous. Although when Wade wrote his critique[64] of *Factortame (No. 2)*,[65] the *Thoburn*[66] judgment (containing Laws LJ's definition of constitutional statutes) had yet to

[63] Ahmed and Perry (n 31) 474.
[64] Wade (n 1).
[65] *Factortame* (n 2).
[66] *Thoburn* (n 24).

be pronounced, it is clear that had he had the chance to respond, he would have pointed to the incompatibility of Laws LJ's rationalisation of constitutional statutes with an orthodox understanding of Parliamentary sovereignty, and he would have been right in doing so. Second, judges more recalcitrant to common law involvement in qualifying Parliamentary sovereignty could potentially abolish the special treatment enjoyed by constitutional statutes if the justificatory conundrum is not properly addressed.

Time has now come to propose my own definition and justification for the existence of constitutional statutes. It will be my mission, in the following paragraphs, to show that implied-repeal-shielded statutes are entirely explainable within the strict boundaries of orthodox Parliamentary sovereignty, and that reliance on extra-statutory considerations is not required to identify a class of statutes shielded from implied repeal.

C. Salvaging constitutional statutes

(i) On implied repeal

Let me for a moment return to the basics. The sole distinguishing property of a constitutional statute is that it is immune to implied repeal. Implied repeal, per Laws LJ, is the rule "that if Parliament has enacted successive statutes which on the true construction of each of them make irreducibly inconsistent provisions, the earlier statute is impliedly repealed by the later. The importance of the rule is, on the traditional view, that if it were otherwise, the earlier Parliament might bind the later, and this would be repugnant to the principle of Parliamentary sovereignty".[67] This definition clearly highlights the nexus between implied repeal and Parliamentary sovereignty. Hence, it comes (to me at least) as a surprise to read later on in the same judgment that His Lordship deems the doctrine of implied repeal to be a creature of the common law.[68] It is my submission that the doctrine of implied repeal is no such creature. In fact, implied repeal is not a "doctrine" at all: it is simply, and uncontroversially, a corollary of the doctrine of parliamentary sovereignty itself. Repeal occurs impliedly when a later statute is inconsistent with an earlier statute i.e., when one or more provisions in each statute is or are mutually exclusive. That is so because Parliament's volition is, logically, always assessed from the latest Act, in a way that were we to take a snapshot of the whole body of Acts of Parliament after the latest enactment, any irreconcilably inconsistent temporally antecedent provision would naturally be automatically discarded. In other words, it would not be representing Parliament's most up to date volition if repeal of the earlier provision did not take place. It is hardly conceivable that this is a doctrine

[67] ibid 37.

[68] ibid 60.

of the common law: it is not as if the courts had a choice to give effect or not to Parliament's wishes. The implied repeal of a statute simply is the result of the courts doing their job under the orthodox constitutional framework, nothing more.

But this is not merely a semantic argument. Indeed, it shows that *nothing in the rule about implied repeal stipulates that every statute shall be impliedly repealable*. Following from the very essence of implied repeal, *it not being a doctrine of its own, but rather a derivative of the doctrine of Parliamentary sovereignty, and therefore not amenable to being interpreted from a scope which favours one's normative ideals*, whether a statute is impliedly repealable or not is entirely and solely dependent upon the wording that is used to convey Parliament's intention.

(ii) My definition

To be clear, this paper's aim is not to discredit the importance attributed to statutes hitherto labelled as "constitutional" by judges and commentators. Rather, it seeks to find a new way to justify their immunity to implied repeal which fits within a traditional Diceyan understanding of Parliamentary sovereignty.

In light of the above discussion, it will hopefully have become clear to the reader that I am not going to tender a definition which attempts to capture statutes currently recognized as constitutional. It is not the business of the judiciary to confer preferential treatment to a statute in the face of implied repeal in virtue of a certain normative importance it boasts in the eyes of the judge. Rather, the better view is to leave it to the statute itself to tell us how it should be handled.

I wish to suggest that some statutes are immune from implied repeal not *because* they boast normative importance (although they often do), but *because* such immunity is linguistically provided for in the statute itself. In other words, some statutes are constitutional *because* they are immune from implied repeal as opposed to them being shielded from implied repeal *because* they fit under one definition of "constitutional statute". The latter view puts the cart before the horse. This draws directly from the abovementioned proposition that there is nothing in the principle of implied repeal to say that every statute ought to be impliedly repealable. My definition thus reads as follows:

"A constitutional statute is a statute which *expressly conditions* the substance of past and future laws by tasking the courts to *supervise* the content of the entire body of statutes and, *when provided for*, remedy enactments that fall short of the *standard* required by the statute. "Constitutional statute" here takes the meaning of a statute which, although always vulnerable to express repeal, for the time being shapes (to

varying degrees) the constitution of (as in the composition of) the entire body of laws of the United Kingdom".

Since this definition relies exclusively on the language of the constitutional statute, it avoids the need to undertake a normative analysis of the statute under scrutiny the result of which is amenable to vary according to the judge or the state of the common law at any point in time. But perhaps most importantly, it means that constitutional statutes are not only completely immune from implied repeal but are also immune from claims that they should not be. Linguistically speaking, and Parliament's intentions are only assessable from language, it is impossible to contemplate the repeal by implication of statutes that fit within the above definition. Indeed, the express intention for a statute to supervise the content of other statutes logically exempts that supervising statute from the application of implied repeal. Finally, since constitutional statutes are now solely products of Parliament's intentions, (almost)[69] nothing seems to be in the way of such statutes investing the courts with disapplication powers.

(iii) HRA 1998 and ECA 1972

To illustrate the above proposition, let's discuss two statutes widely accepted as constitutional even under the currently prevailing narratives: The Human Rights Act 1998 (HRA 1998) and the European Communities Act 1972 (ECA 1972) the latter of which, for the purposes of this discussion, will be assumed to still be in force.

Let us start with the HRA 1998. Section 3(1) of the HRA 1998 requires the courts to, "so far as it is possible to do so", read "primary legislation and subordinate legislation […] in a way which is compatible with the Convention rights".[70] The supervisory function of this provision is clear: all past and future legislation, whether primary or secondary, given two competing but equally eligible interpretations attributable to the statute, where one of which is in conflict with Convention rights, should be given the construction which aligns the piece of legislation with Convention rights, notwithstanding that the interpretation chosen is somewhat less eligible than the other from a conventional interpretative approach. In other words, because a later Act of Parliament could be interpreted in a way which infringed Convention rights does not mean that it impliedly repeals the earlier HRA 1998 or that it is shielded against it. That is so because the HRA expressly conditions the law so far as interpretatively possible. But the HRA goes further in its role as a constitutional statute. It also tasks the courts, in case of an interpretive impasse (i.e., where an interpretation compatible with the ECHR would be plainly *contra legem*), with notifying Parliament that its enactment is

[69] See Section V below.

[70] Human Rights Act 1998, section 3(1).

incompatible with the ECHR. Such is the function of section 4 and its "declaration of incompatibility" which allows the incompatibility between the HRA and a later statute to subsist without repeal taking place.[71] In that sense, the constitutional reach of the HRA is whole : either the law under scrutiny is already interpretively compatible with Convention rights and passes the checkpoint, either it is able to bear two meanings (and maybe one interpretation is to be preferred to the other on a literal approach) in which case the meaning compatible with Convention rights is selected, either it is irresistibly incompatible and Parliament is notified and a special procedure becomes available should the Government wish to comply with the court's findings.[72] The HRA is not shielded from implied repeal because it is of normative import (although it undeniably is). Nor because Parliament must have given the HRA some especial considerations (although it surely did). The statute and its provisions are shielded from implied repeal because Parliament has linguistically tasked the courts to supervise its own enactments and to interpret them as far as possible with ECHR rights, or, if that is impossible, sound the alarm while providing for the mutual incompatibility to stand.

Let's turn now to the ECA 1972, another Act of Parliament widely reputed to enjoy the status of constitutional statute.[73] Section 2(4) of the ECA 1972 provides that "any enactment passed or to be passed […] shall be construed and *have effect subject to the foregoing provisions of this section*" (emphasis provided).[74] The supervisory nature of this section is clear: the words "have effect subject to" conveys as much. What is more, section 2(4) is an *active supervisory provision*, as distinct from the *passive supervisory* nature of the HRA 1998. In addition to being a constitutional statute of the kind the HRA 1998 is, the ECA 1972 tasks the courts to remedy any provision in other statutes which run contrary to Community law: United Kingdom laws may only stand ("have effect") if they comply with EU law. That the ECA 1972 was a statute of constitutional magnitude in conventional terms is not in question here: the contention that it could not be impliedly repealed because of its normative importance is. Not only does this reasoning depart from Diceyan orthodoxy, but as shown, it is not even necessary to go to such contentious lengths. Indeed, the answer is found within the statute itself. It is Parliament's wish that its doings be supervised by the courts and kept in line with EU law so long as the ECA 1972 is not expressly repealed, or a statute expressly shielded from the bite of section

[71] ibid section 4(6).
[72] ibid section 10.
[73] See, for example, *Thoburn* (n 24) at 62 and Ahmed and Perry (n 31) 462.
[74] European Communities Act 1972, section 2(4).

2(4). It is Parliament which precludes, by its own wording, the operation of implied repeal.

(iv) Anticipating objections

The view I am proposing in this paper is a drastic departure from the current rationales underpinning constitutional statutes. I should therefore attempt to meet some possible counters.

The first obvious objection to this view is that it entails that Parliament can literally make any law constitutional by inserting a section equivalent to section 2(4) of the ECA 1972. This indeed follows from the view I have proposed above, and I accept this ramification. Nonetheless, there are two things to keep in mind which, I hope, abates the concern. The first, which should never be forgotten, is that however much a statute shields itself from implied repeal, that is as far as it can go to safeguard its perennity. A later statute which directly addresses the repeal of the earlier constitutional statute proceeds untrammelled in its deed, however strongly worded the constitutional clause in the earlier statute may be. It would therefore seem that an appeal to the perhaps disconcerting consequence which our view entails (that Parliament can make *any* statute constitutional) would seem to rely too heavily on a meaning of "constitutional" that constitutional statutes in our jurisdiction simply do not share with their counterparts elsewhere. The second point in answer to such criticism would be that the chances that a constitutional clause would be attached to a trivial statute are slim. Indeed, it is not inconsistent at all with our view that most statute which will be shielded from implied repeal and thus be constitutional under our definition will also meet the criteria of the definitions tendered by Ahmed and Perry.[75] This will indeed stem from the normative importance that those statutes will usually carry which will have made it apposite for the legislature to have annexed the supervisory constitutional clause. Correlation, though, is not causation.

An example of a constitutional statute captured by my definition which might take readers by surprise is the statute that was in question in *Ellen Street Estates Ltd v Minister of Health*.[76] Section 7(1) of the Acquisition of Land (Assessment of Compensation) Act 1919 provided that "The provisions of the Act or order by which the land is authorised to be acquired, or of any Act incorporated therewith, shall, in relation to the matters dealt with in this Act, *have effect subject to this Act*, and so far as inconsistent with this Act those provisions shall cease to have or shall not have effect." This is a paradigmatic case of constitutional statute under the newly coined definition: the statute tasked the courts to supervise past and future enactments. It was also an *active supervisory statute*. But in the *Ellen Street Estates* case,

[75] See above.

[76] *Ellen Street Estates Ltd v Minister of Health* [1934] 1 K.B. 590.

Section 7(1) was repealed. It may seem from reading the judgment that Scrutton LJ thought he was giving effect to repeal by implication: "Parliament can alter an Act previously passed […] by enacting a provision which is clearly inconsistent with the previous Act".[77] But if Section 7(1) rendered the Act a constitutional one, on our definition even a clearly incompatible later provision would not entail its repeal. Have I got it all wrong? I do not, for clearly the better view is that Section 46(2) of the Housing Act 1925 (which contained provisions, in s46(1), which were at odds with the 1919 Act) provided for the express shielding of that statute from the bite of s7(1), the supervisory clause. Section 46(2) provided as follows: "Subject as aforesaid, the compensation to be paid for such land shall be assessed in accordance with the Acquisition of Land (Assessment of Compensation) Act 1919." The effect of this sub-section is blatantly clear: by providing the words "subject as aforesaid", Parliament expressly shields section 46(1) of the 1925 Act from the operation of the 1919 active constitutional provision. If it had not said as much, Parliament's will would have had to be construed by the courts as wishing for the disapplication of section 46(1), but section 46(2) demands that the supervisory jurisdiction of the courts, as empowered by section 7(1), not apply to section 46(1).

A final criticism which it is incumbent on me to address is that I have simply cherry-picked currently recognized constitutional statutes that do contain the all-important constitutional provision. This is, indeed, a critique to which I am legitimately exposed but to which answers exist. Firstly, the statutes I have so far discussed are the only statutes whose constitutional status have had actual practical judicial effect. Other statutes have been recurrently labelled as constitutional by the courts,[78] but this has not yet been followed by any concrete application. At the very least, what my analysis strove to demonstrate is that where repeal by implication has been refused and where the court's active supervision was exercised as in *Factortame (No. 2)*, there was no need for judges to rely on anything but the statute itself to refuse the operation of implied repeal.

D. Conclusion

It is to be hoped that the benefits of this *technical argument* and of this new definition of constitutional statutes are apparent if we are to parry claims of constitutional upheaval pursuant to the *Factortame (No. 2)* judgment. The starting point to my justification was that implied repeal is not in fact a doctrine properly so-called. Rather, implied repeal derives from the need to give effect to Parliament's most recent intentions: if such intentions clash with chronologically earlier ones,

[77] ibid 595.

[78] *R (on the application of HS2 Action Alliance Limited) (Appellant) v Secretary of State for Transport and another (Respondents)* [2014] UKSC 3, [2014] 1 WLR 324, 207.

they take precedence. But by wording a statute so as to confer it supervisory functions, Parliament pre-emptively tells the courts that implied repeal does not apply to that statute. Since this justification for immunity to implied repeal is now sourced from Parliament's own, express volition, there is no dichotomy between Parliament's intentions and the rationale for the impossibility of implied repeal: the impossibility is provided by the statute itself. Gone is the need to ascribe higher-order status based on certain qualitative criteria. Furthermore, now that a way has been provided for implied-repeal-shielded statutes to be understood solely as products of Parliament's intention, there is no risk that their legitimacy be questioned. And what is more, if a constitutional statute grants strike down powers to the courts, then to do so, if the circumstances arise, is simply to give effect to Parliament's intentions, as the House of Lords duly did in *Factortame (No. 2)*.

V. The Constitutional Argument: Parliament's 'Atemporality'

The *technical argument* sought to demonstrate that there is no disconnect between Parliament's intentions and the immunity to implied repeal enjoyed by constitutional statutes. This new definition meets and deflects claims that constitutional statutes are in fact instruments of the common law designed to stymie Parliament's sovereignty in favour of some normative ideals by putting hurdles in the normally necessarily free legislative path a sovereign entity properly so-called ought to enjoy. Yet, it would be rash to assume the matter settled. Indeed, there is a more subtle yet entirely tangible conceptual hurdle that needs to be addressed: the idea that Parliaments succeed one another. If the widespread assumption that at every general election a "new" Parliament comes into existence is not overcome, then my defence of constitutional statutes and of the outcome of cases such as *Factortame (No. 2)* as being entirely Dicey-compliant fails, and Wade's revolution thesis prevails.

The problem is best illustrated by the very probable retort with which the *technical argument* is prone to be met: 'it is all very well that it is Parliament's express intention that some statutes be shielded from implied repeal and that some of these statutes empower courts to disapply incompatible legislation, but this fails to distinguish between what are intentions from an *old and obsolete* Parliament (e.g., Parliament of 1972), and those from the *newly empowered* Parliament (e.g., Parliament of 1988) whose intentions are, naturally, supreme. In enacting a statute, the later sovereign Parliament clearly does not *intend* for it to be disapplied by a provision found in a statute enacted by a previous Parliament, even though that provision, under the *technical argument*, seemingly requires courts to quash the new Parliament's statute unless expressly shielded. To require such express provision instead of allowing it to take place by implication is to foray into the territory of

manner-and-form requirements which is foreign to an orthodox understanding of Parliamentary sovereignty. Therefore, the *technical argument* is flawed, because it asserts that an earlier Parliament may sneakily bind a later one. In short, an Act of Parliament should never be under threat from a temporally antecedent one'.

This is essentially the rebuttal Wade had in mind when, commenting on Lord Bridge's construction approach to the effect that section 2(4) of the ECA 1972 applied to Part II of the MSA 1988, he said that his (Lord Bridge's) "hypothetical section would take effect by authority of the Parliament of 1988, not the Parliament of 1972."[79] A few lines later, Wade further submits that "to hold that its terms are putatively incorporated in the Act of 1988 is merely another way of saying that the Parliament has imposed a restriction upon the Parliament of 1988."[80] If this argument stands, my *technical argument* is rendered worthless because I will have violated the necessary derivative of Parliamentary sovereignty discussed in Section III above, namely that Parliament may never be legislatively restricted whatsoever, including by manner-and-form hurdles.

With my *constitutional argument*, I make good my commitment to Diceyan orthodoxy and show that the *technical argument* is concordant with the fundamental principle of absolute legislative freedom.

A. When semantics distort the conceptual

(i) "Parliament and its successors […]"

The phrases "Parliament and its successors", "Successive Parliaments", "Future Parliaments", "The Parliament *of* [insert any date]" are all widely used across the legal literature. In HWR Wade's piece, I count ten occurrences where the author makes use of phrases conveying that Parliaments succeed one another.[81] In *Thoburn*, the fact that "Parliament cannot bind its successors" is repeated multiple times.[82] It is ubiquitous terminology.

Surely, this terminology can be useful. It allows commentators and readers to situate themselves in time. For example, referring to "the Parliament of 2020" will be an easy way to rapidly put the ensuing commentary in the context of the coronavirus crisis. As a contextualisation tool, then, it is harmless. Perhaps some commentators also employ this phraseology as a linguistic shortcut to emphasize

[79] Wade (n 1) 570.
[80] ibid.
[81] Wade (n 1).
[82] *Thoburn* (n 24) 51, 58–59.

the fact that from time-to-time Parliament has tried, and inevitably failed, to handcuff its future *self*.

Unfortunately, it is not just some innocuous linguistic shortcut or mere contextualizing device. It is not, in other words, only about semantics. When we commonly talk of past and successive Parliaments, we mean it. Indeed, under this view, Parliaments succeed one another e.g., the Parliament *of* 1988 is distinct from the Parliament *of* 1972: they are both completely independent sovereign entities. To talk of "successors" is to necessarily imply that every time there is a general election and hence a newly composed House of Commons, there is a new sovereign Parliament. For example, the "Parliament of 1972" became sovereign in 1970 until the following general election in 1974 which saw it lose its sovereignty in favour of the Parliament of 1974, so on and so forth.

The implications of such a view must be carefully scrutinised. Assuming that this is how we should understand the nature of Parliament, key features of our constitutional framework seem unexplainable. It is striking, for example, that despite the steady stream of new sovereigns since 1689, some "Acts of Obsolete Parliaments" carry on without *any word of acquiescence having been uttered by the legitimate, current Parliament*. For example, under this understanding of Parliamentary sovereignty, all new sovereign Parliaments since the end of the 1861 Parliament have had the Offences Against the Person Act 1861 imposed upon them at the outset of their new sovereignty. This surely flouts any sound definition of sovereignty: the investiture of a new sovereign demands, by force of logical necessity, *tabula rasa*.

Yet, if we take a quick glance at our current constitutional arrangement, statutes enacted by earlier legislative assemblies are not commonly deemed to be forcibly imposed upon "later" Parliaments. Acts of Parliament seamlessly ride the turbulent waves of general elections and progress unscathed through the passage of time unless they are repealed. This simple observation points to a different conceptual understanding of Parliament and casts serious doubts on the "successive Parliaments" hypothesis.

I should stress once again that my contention is not merely semantic in nature. It would be semantic if I was simply displeased with the way the concept of Parliamentary sovereignty, with which I otherwise agreed, was formulated. Thus, to come back to what was said above, if the idea that Parliaments succeed one another was merely a linguistic shortcut or a stylistic device, I would have no gripe with it, or perhaps only the one that it fuels the obfuscation of the true picture of the constitutional framework. But this is not what is happening here. Rather, what I am forced to conclude is that the *semantics inform the conceptual* so much so that, in

my submission, the true and sound conception of Parliamentary sovereignty has been blurred irredeemably.

(ii) Parliament is one continuous entity

It is submitted that there is only, and has ever only been, one Parliament. Since its inception in 1689,[83] Parliament has never been stripped of its sovereignty nor replaced by a new Parliament after each general election. Parliament is a temporally continuous entity: this is the *constitutional argument*.

The source of the misapprehension that Parliaments succeed one another lies in a failure to make the distinction between the legal and the political: between, that is, Parliament and Parliamentarians. The former is the entity into which legislative supremacy is vested by the rule of recognition of this jurisdiction.[84] Parliamentarians, on the other hand, are politicians devoid of any legislative powers as such. Parliament, the legal entity, is completely independent from all the political processes that run in the foreground of the political sphere to legitimise the supremacy of its enactments. The most flagrant example of such political processes are general elections which serve as the ultimate accountability mechanism for Parliament and Government. General elections see Parliamentarians succeed one another and yet Parliament, *qua* the lawgiver, is unaffected by these internal changes.

Proof of this is readily available. The exclusive conduit through which the supreme legislator communicates its volition, and the only instrument with legal weight, is the Act of Parliament. Not Acts of "Conservative" Parliament; not Acts of "Labour" Parliament. The political leanings and manifestos which Parliamentarians carry in and out of the House of Commons have no legal impact whatsoever on the source of laws. Although political promises and manifestos do eventually expire and are succeeded by new ones, Acts of Parliament are always contemporary, continuous; and so is Parliament.

Once again, this is not an account of how things should be, but how things actually are. If one accepts that the validity of Acts of Parliament is unchanged despite the multiple general elections that fill the political landscape in the years following the enactment of an Act of Parliament – and a general election is the event which, according to the "successive Parliaments" thesis, discriminates between an earlier and a later Parliament – then one must also accept that the constitutional picture is otherwise than one where the legislative sovereign is whatever body of

[83] This date was also taken by Mark Elliott in: Mark Elliott, 'The United Kingdom's Constitution and Brexit: A "Constitutional Moment?"' (2020) Horitsu Jiho 15-22, University of Cambridge Faculty of Law Research Paper No. 22/2020.

[84] Hart (n 60) 149.

Parliamentarians sit at one point in time and therefore that as soon as there is a change in the composition of such body, there is a new sovereign Parliament.

The rule of recognition of the United Kingdom is that whatever Parliament duly enacts is added, or *superimposed*, to the conversation it is continually having with *itself*, and it is the sum of all the additions *and* subtractions which take place within that internal monologue that give us the law of the jurisdiction. The courts, in grappling with this complex soliloquy, ought to remain loyal to the letter of Parliament's textually communicated volition and refrain from interpreting legislation and the relationship between different legislative texts according to the nature of the political body who enacted the respective statutes.

B. The implications of the constitutional argument

(i) A new relationship between legislature and judiciary

The implications of this reconceptualization of Parliament as a temporally continuous entity are momentous for this article's thesis. Again, an Act of Parliament must be deemed to be always speaking and always wished for by Parliament unless Parliament conveys, impliedly or expressly, that it does not intend the Act to have further effect. This flows from the idea of a temporally continuous sovereign legislator which, since 1689, has had but one omnipotent voice. If an Act of Parliament is disapplied by the courts as a result of its not conforming to boundaries set by an earlier supervisory constitutional statute as defined in Section IV.C above, the correct interpretation to give to the disapplication is that Parliament is *perfectly happy and has intended for* the statute to be disapplied because it has expressly tasked the courts, and has not expressly revoked such task, to be wary of and make right any incompatibility in previous legislation and, what is more, remedy any oversight in all future Acts of Parliament.

The opposite approach, informed by the view that "Parliaments" literally, conceptually, succeed one another, would be to assume, in light of intentions imputed to Parliamentarians belonging to the majority party, be it through what is said by the minister spearheading the Bill or by inference from the party's political inclinations, that the later Act was clearly intended to override the earlier one. Since the "successive Parliaments" thesis is predicated on the premise that general elections entail a "new" Parliament, there creeps inevitably in such reasoning notions which are anathema to sound constitutional reasoning, namely judicial assumptions of a statute's effect based on the political composition and leaning of the House of Commons. To approach statutory construction in such a way is to dangerously blur the fine line between the legal and political realms. Yet, it is what the "successive Parliaments" approach invites one to do, consciously or not.

It is absolutely vital, if the separation of powers stands as a core principle of the United Kingdom's constitution, as it clearly and necessarily does, that Parliament is characterized, in the legal constitutional context, as the atemporal, apolitical body that it is whose laws should not be interpreted with any political gloss.

(ii) In practice: Shielding Factortame (No. 2) from Wade's critique

In 1990, the House of Lords disapplied a provision of the Merchant Shipping Act 1988 (MSA 1988) by virtue of section 2(4) of the ECA 1972 which provided in essence that "European Community law was to prevail over Acts of Parliament 'passed or to be passed'".[85]

HWR Wade was categoric in his critique of the judgment: "the Parliament of 1972 has imposed a restriction upon the Parliament of 1988",[86] something which obviously is at odds with the fact that Parliament cannot bind itself and, accordingly, with the principle of Parliamentary sovereignty. Thus, for Wade, when "the House of Lords elected to allow the Parliament of 1972 to fetter the Parliament of 1988 in order that Community law might be given the primacy"[87] by tweaking the rule "that an Act of Parliament in proper form had absolutely overriding effect, except that it could not fetter the corresponding power of future Parliaments",[88] it amounted to a "technical revolution",[89] but a revolution no less.

Naturally, Wade's criticism was not left unanswered. And swiftly did the response come in the form of *Thoburn*.[90] As seen previously, Laws LJ's judicial definition of constitutional statutes seemed at first to quell Wade's concerns: it is not that the House of Lords tweaked the well-established rule of recognition, it is that the ECA 1972 is a constitutional statute in the eyes of the common law and "[the ordinary rule of implied repeal] has no application to constitutional statutes".[91] It is to be hoped that the section of this paper concerned with the *technical argument* demonstrated why this is a precarious justification. Although it certainly aimed to refine *Factortame*, *Thoburn*'s own rationale for constitutional statutes and their interaction with implied repeal had gaping holes. While the outcome it leads to is right, the justification afforded by *Thoburn* is unsatisfactory and precarious as it does not stand on solid constitutional ground.

Yet, Laws LJ was right: the ECA 1972 may not be impliedly repealed and is a supervisory constitutional statute which empowers the courts to disapply incompatible statutes. It is so because on its proper construction and nothing

[85] Wade (n 1) 568.
[86] ibid 570.
[87] ibid 574.
[88] ibid.
[89] ibid.
[90] *Thoburn* (n 24).
[91] ibid 63.

more (*technical argument*), it provides for every past and future Acts of Parliament to be supervised by and subjected to EU law. In 1972, Parliament, in accordance with the political conditions of the UK's accession to the European Community, decided to empower the courts to disapply Acts of Parliament at odds with EU law.

In 1988, at the time of the passing of the Merchant Shipping Act 1988, Parliament maintained such wish since nothing in the 1988 Act expressly disavowed s2(4) of the ECA 1972. It follows that Parliament in 1988 was *happy* for the MSA 1988 to be regulated by s2(4) of the ECA 1972 and repealed in its name. Buttressed by the fact that Parliament is one temporally continuous sovereign entity which can add layer of supervision to its own enactments (*constitutional argument*), it is easy to see how, far from having successfully been permitted to bind itself, Parliament in fact had its intentions afforded the respect they should.

C. Conclusion and broader implications of the above view

The importance of avoiding the language of "successive Parliaments" is twofold. First, it enables us to give new dimensions to an orthodox understanding of Parliamentary sovereignty. Under this reconceptualization, there is nothing contentious in the idea of a dialogue between Parliament and the courts where the former asks the latter to supervise and make good any oversight in its legislative functions. Such a dialogue can only be adequately understood when the idea that multiple Parliaments continuously succeed one another is jettisoned. Again, this is not mere semantics, as the way we conceptualize Parliament has important ramifications on the constitutional validity of judicial actions under orthodox principles.

Second, it avoids the dissemination of distorted constitutional narratives. Indeed, in the *Factortame (No. 2)*[92] case, it could have been tempting for the courts to make the leap of taking the minister's wish (of the Act being intended to supersede the ECA 1972) for Parliament's (which was, at that time, composed of a majority of MPs from the minister's Conservative party). But the separation of powers at the heart of the United Kingdom's constitution will not allow such callous shortcut. Whatever its majority in the House of Commons, the Government is not Parliament. It was always possible for Parliament to expressly shield the MSA 1988 from the bite of section 2(4), but without express provision to that effect,

[92] *Factortame* (n 2).

Parliament's intentions (albeit surely not the Government's) were for section 2(4) to disapply the infringing provision in the Merchant Shipping Act 1998.

VI. Conclusion

The aim of this paper was to show that *Factortame*-like judicial statute disapplication is not abhorrent to an unqualified Diceyan doctrine of Parliamentary sovereignty. To achieve this objective, this paper challenged two prevailing ideas in modern constitutional discourse.

Firstly, the *technical argument* submitted that constitutional statutes are not shielded from implied repeal *because* they are of especial normative worth (although they often are), but because it is Parliament's express intention that they supervise past and future legislation, with the result that implied repeal is simply unavailable. The fact that implied repeal is not a doctrine of its own but a corollary of the doctrine of parliamentary sovereignty itself paved the way for this purely construction-based approach. This hopefully avoids the precarious position which resulted from a normative-quality-inspired rationale for constitutional statutes' immunity to implied repeal. Moreover, once a statute is vested with a supervisory function, it may further provide on what should happen with past and future statutes that infringe its provisions. When the constitutional statute provides that they should be disapplied, then the proper move from a court is to do precisely just that, unless expressly told otherwise.

Secondly, the *constitutional argument* showed that Parliament must be viewed as a continuous entity whose apolitical legislative identity is not influenced by its political composition. This was a crucial step to properly attach the *technical argument* to the orthodoxy train. Indeed, the "successive sovereigns" view would have legitimized claims that this seemingly construction-based approach was simply manner-and-form hiding in plain sight.

Statute disapplication, when all ingredients obtain, is therefore not the transgression of Parliament's sovereignty under an orthodox Diceyan understanding, nor even its qualification: it is its full, unqualified expression.