University of Southampton

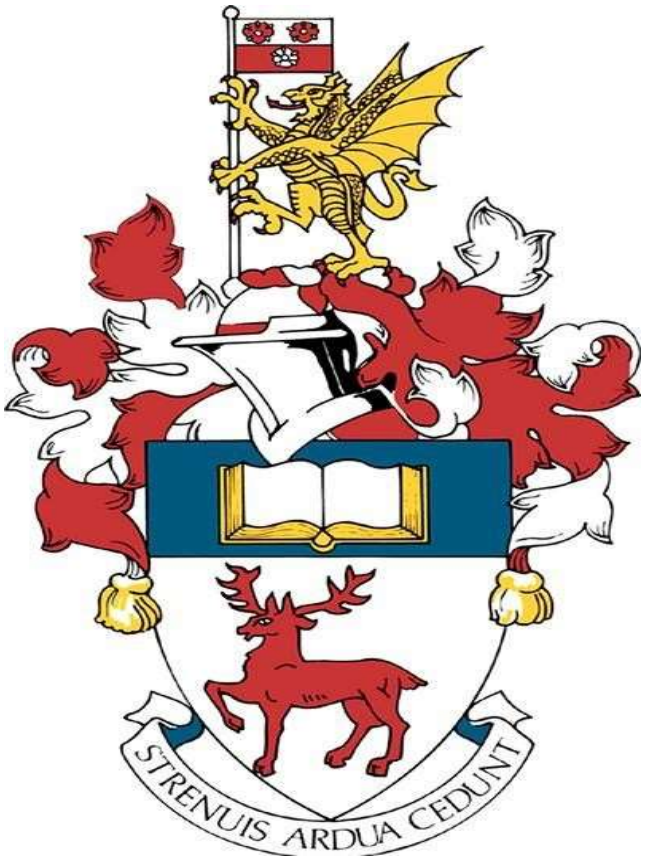
LEGAL RESEARCH AND DEVELOPMENT

SOUTHAMPTON STUDENT LAW REVIEW
2025 VOLUME 15, ISSUE 1

University of Southampton

SHOWCASING EXCELLENCE IN RESEARCH

SOUTHAMPTON STUDENT LAW REVIEW
2025 VOLUME 15, ISSUE 1



Southampton Student Law Review

Southampton Law School

Published in the United Kingdom
By the Southampton Student Law Review
Southampton Law School

University of Southampton
SO17 1BJ

In Affiliation with the University of Southampton, Southampton Law School

www.southampton.ac.uk/law/lawreview
ISSN 2047 – 1017

This volume should be cited (2025) 15 (1) S.S.L.R

Editorial Team

Acknowledgements

The Editors wish to thank William S. Hein & Co., Inc. and HeinOnline for publishing the Southampton Student Law Review on the world's largest image-based legal research collection. They also wish to thank vLex Justis for including the Southampton Student Law Review on their global collection of legal information.

The Editors also wish to thank all members of Southampton Law School who have aided in the creation of this volume.

Aminayanasam Tekena-Fubara and Kübra Yıldız
Editors-in-chief, Southampton Student Law Review
September 2025

Table of Content

Foreword

We don't always live in exciting times. Sometimes, we live in interesting ones; and as any seasoned Brit will tell you, that's rarely a good thing. Yet for lawyers, 'interesting times' are always exciting: when the rule of law stumbles, when international institutions creak under the strain of global challenges, and when calls to withdraw from regional and supranational systems grow louder, then legal minds are called into action. We are, undeniably, living through such a moment: a time of legal uncertainty, constitutional stress-testing, and a perma-crisis in international law. But it is in these moments that legal thinking and where fresh perspectives become necessary.

The students of Southampton rise to the occasion. With critical insight and a yet healthy dose of optimism, they looked at the failures of our legal systems offering not only diagnoses, but proposals for reform. Three contributions focus on contemporary challenges within domestic criminal law. Christian Badcock opens with a timely reflection on the future of the English and Welsh criminal justice system. Without adopting a wholly abolitionist position, he advocates for a less punitive and more rehabilitative model. Continuing the theme of reform, Erin Deane focuses on youth justice and evaluates the case for a 'minimum intervention, maximum diversion' approach, proposing a more developmentally appropriate response to youth offending. Nora Belkhiter turns our attention to miscarriages of justice, particularly where complex scientific evidence is concerned. She assesses the proposal for expert pre-trial panels to mitigate such risks, acknowledging their limitations but ultimately supporting the creation of a 'modified' type of panels for the monitoring of investigations. Sahira Xec's article addresses a different justice issue, namely the treatment of children's voices in judicial decision-making. In support of Lady Hale's concerns, she offers a critique of how courts weigh children's wishes, questioning whether the principle of autonomy is genuinely upheld in the face of welfare-driven paternalism. Erin Guizot shifts the focus to the commodification of the UK's housing system through neoliberal policies and 'financialisation', calling for rebalancing housing policy to protect tenants' rights as a social necessity rather than a market commodity. Three articles focus instead on contract and commercial law. Abigail Atkinson – quite differently from Guizot –challenges the rule against penalty clauses in contract law, arguing that it unjustifiably restricts the freedom of parties to allocate risk and enforce agreed-upon remedies. She defends contractual autonomy and critiques the 'traditionally paternalism' that underpins current doctrine in the UK. Xinyue Zhang examines the legal treatment of fraudulent claims in marine insurance, tracing a doctrinal shift from strict forfeiture rules to a more balanced approach, particularly following the Insurance Act 2015 and DC Merwestone's concept of 'collateral lies'. Vernon Smith addresses another judicial shift in the

UK's insurance doctrine. He contributes an analysis of causation in insurance law, questioning the continued reliance on the but-for test in multi-cause disputes, which has been judicially recognised as too narrow. Last but not least, bringing the analysis to a completely different jurisdictional level, Gregory Cooke addresses the role of international law, human rights law, and international criminal justice in conflict resolution, critiquing their limited practical utility of in volatile contexts such as Sudan.

Andrea Maria Pelliconi
Lecturer
University of Southampton Law School

The Shackles of Penalty Clauses on Contractual Freedom

Abigail Atkinson

Abstract

The unenforceability of penalty clauses, also known as the 'rule against penalty clauses' is an unjustifiable affront to the constitutional principle of contract: freedom. This research will explore the traditionally paternalistic approach the UK law takes when parties are at risk but ultimately argue that the imposition of unenforceability on penalty clauses prevents the protective intention of the parties in implementing penalty clauses. This research argues that whilst penalty clauses can be described as exorbitant at times, there is no reason to reign this in where it has been agreed freely. This paper will conclude that whilst the alternative option of liquidated damages poses a solution to the unenforceability of penalty clauses, the rule against penalty clauses is unjustifiable, and has created uncertainty and injustice.

Introduction

Freedom of contract is unjustifiably interfered with by the unenforceability of penalty clauses. Despite the efforts of Cavendish[1] and other sources to reason the imposition of penalty clauses as a protective manoeuvre, the restriction of freedom of contract is nonetheless unjustifiable. It is submitted that liquidated damages pose a more reasonable solution for the protection of parties whilst supporting the freedom of contractual parties to determine their 'agreed' remedies.

1. What are Penalty Clauses and How Do They Interfere with Freedom of Contract?

Penalty clauses are a type of agreed damages which penalise the non-performer by imposing an exorbitant level of damages. They are different to liquidated damages which represent a genuine attempt to pre-estimate the loss in line with the compensatory principle.[2] Penalty clauses directly interfere with freedom of contract, the cornerstone in contract law which enables parties to consensually enter agreements and tailor the agreement to their mutual understanding. Freedom of contract is relevant to the formation of the contract, the prevention of undue influence or duress, and certainty of term, permitting the enforcement

[1] *Cavendish Square Holdings BV v Makdessi* [2015] UKSC 67.
[2] *Dunlop Pneumatic Tyre Co Ltd v New Garage and Motor co Ltd* [1914] UKHL 1.

of the parties' intentions. The freedom of parties, emphasised in *Hadley v Baxendale*[3], manifested in stipulating what the damages shall be.[4] Not delivering the expectations which the parties have specifically agreed goes against their intentions and the principle of freedom of contract. Why is this freedom being directly inhibited by the unenforceability of penalty clauses?

2. Justifications of Interfering with Freedom of Contract

The justifications of interfering with freedom of contract stem from protecting the public interest, for instance, invalidating a contract which violates law or policy, voiding a criminal contract, or preventing a contract which cannot be safely or reasonably enforceable. The meaning of the good faith has been inferred broadly in law, for instance, in *Bates v Post Office Ltd (No 3)* [5], good faith was inferred as duties of mutual trust, loyalty, and performance with integrity. Additionally, according to the Consumer Rights Act[6], several implied terms were imposed to protect consumers' rights, sometimes restricting freedom of contract prioritising contractual fairness. When there is a contract between a consumer and a business[7], it must account for the potential power imbalance between the two parties. The law clearly exerts paternalistic protection for security and safety, but this cannot be the reason for penalty clauses to be limited so oppressively.

3. The Governance of Penalty Clauses and Their Interference – Can it Be Justified?

The enforceability of penalty clauses is paternalistic in preventing severe consequences where performance fails, therefore it will always be directly in conflict with liberalism and freedom of contract. However, the restriction over penalties creates several illogical consequences from doing so. Penalty clauses were first determined to be unenforceable in *Dunlop Pneumatic Tyre Co Ltd v New Garage and Motor co Ltd* [8] due to their extravagant and unconscionable[9] level in comparison with the greatest conceivable loss from the contract.

[3] *Hadley v Baxendale* [1854] EWHC J70.
[4] ibid.
[5] [2019] EWHC 606 (QB).
[6] Consumer Rights Act 2015.
[7] Consumer Rights Act 2015, s 9-11, s13, s 14.
[8] 1914] UKHL 1.
[9] ibid.

The leading authority, *Cavendish Square Holdings BV v Makdessi*[10], imposes a test to determine whether the legitimate interest[11] of the innocent party goes beyond financial compensation, but it is impossible to see why this is relevant when the parties have exercised their freedom of contract in crafting these terms. It is argued that the unenforceability of penalty clauses is irreconcilable with freedom of contract[12] and it is difficult to disagree considering *Cavendish Square Holdings BV v Makdessi* [13] prevents parties from achieving certainty and predictability. An innocent party cannot surely determine what they will receive upon the failure of the other party, all to protect the breaching party from the burden of a penalty they would have freely agreed to upon concluding the contract. Why is the law protecting the breaching party and subjecting the innocent party to financial loss?

There appears to be no logical basis to reign in unreasonable terms where both parties have consented to becoming subject to them according to freedom of contract. The penalty rule from *Cavendish*[14] unfortunately allows breaching parties to escape from the liability they have voluntarily assumed, undermining the certainty of contractual agreements. The threat of an unenforceable penalty discourages the provisions incentivising performance and leads to ambiguous agreements. As a result, parties may expend time and money on litigation, overwhelming an already inundated court system. The rule that intends to protect these innocent parties may actually expose them to the unyielding expense of litigation.

The unenforceability of penalty clauses results in uncertainty and injustice for innocent parties[15] as disproportionate protection is extended over a breaching party from incurring damages which they freely agreed to. Focus is placed on protecting them from a financial burden they knew they would incur, rather than on upholding the sanctity of the contract and respecting the freedom of the parties to create their own terms. Even the court in *Cavendish*[16] claimed that parties should be allowed to fix their own compensation[17], yet this was illogically undermined by preventing competent parties from certainly doing so with the penalty doctrine.

[10] [2015] UKSC 67.
[11] ibid.
[12] W Day, 'Penalty clauses revisited' (2014) JBL 512. See also Patrick Atiyah, *The Rise and Fall of Freedom of Contract* (OUP 1979).
[13] [2015] UKSC 67.
[14] *ibid.*
[15] Edwin Peel, *Trietel: The Law of Contract* (15th edn, OUP 2020).
[16] *Cavendish Square Holdings BV v Makdessi* [2015] UKSC 67.
[17] ibid.

In *Cavendish Square Holdings BV v Makdessi*[18], the court attempts to justify the unenforceability of penalty clauses by claiming it will be limited, and that such is a long-standing principle in UK law. The court are correct in that the application of penalties is limited, however the imposition of this rule, even on those few unlucky innocent parties, undermines freedom of contract. The law cannot justify exerting control over party freedom by claiming it only affects a small number of contracts.

Additionally, the court in *Cavendish*[19] claim that the legislation regarding unfair contract terms[20] fails to regulate all contractual scenarios and there continues to be power imbalances between contracting parties which warrant intervention. However, the statutes provide for these situations by imposing protective implied terms for the weaker party where a power disparity exists. It is argued that this is sufficient to provide equity to consumers[21], and it is unnecessary for penalty clauses to be powerless.

4. Why are Liquidated Damages Enforceable but Penalties Not?

Liquidated damages aim to be a genuine pre-estimation of the actual loss likely to be suffered due to breach.[22] This sits more comfortably with the courts as they are aligned with the compensation principle, whereas penalties are interpreted as punitive and derived from an excessive sum often unrelated to the loss truly suffered. For this reason, liquidated damages are seen as fair as they genuinely attempt to compensate the innocent party rather than operate a punitive response to breach.

The existence of liquidated damages can act as a justification for the unenforceability of penalties. From 2015[23] parties have been aware of the distinction between the two clauses, and the potential risk of inputting penalty clauses into contracts. To avoid the potential risk, a more sensible approach could be to impose a liquidated damages clause to ensure the innocent party will be compensated in the event of breach. There is no reason for breach to result in punishment, but to recognise and ameliorate the loss of the innocent party.[24]

[18] ibid.
[19] ibid.
[20] Unfair Contract Terms Act 1977; Sale of Goods Act 1979; Consumer Rights Act 2015.
[21] W Day, 'Penalty clauses revisited' (2014) JBL 512.
[22] *Cavendish Square Holdings BV v Makdessi* [2015] UKSC 67.
[23] ibid.
[24] *Morris-Garner v One Step (Support) Ltd* [2018] UKSC 20.

Returning to freedom of contract, how can the law allow for equity to undermine it so severely when it comes to penalty clauses when such clauses have been freely and consensually agreed? Freedom of contract is an established and constitutional principle of contract law. A more appropriate response from the judiciary would be to apply liquidated damages and penalties as enforceable where there is no issue of consent, duress, or undue influence, where intervention is genuinely necessary. It should be admitted that parties are often the best at determining and facilitating their own commercial interests, and there is no reason for the law to restrict this, even if attempting to serve a protective role.

Conclusion

In summary, the unenforceability of penalty clauses imposes an overbearing restriction on freedom of contract, preventing parties from enforcing terms they have mutually agreed on and confounding legal certainty and predictability. Whilst public interest and policy arguments are relevant and the enforceability of liquidated damages mitigates the vastness of the issue, it is both unnecessary and illogical to simply make penalty clauses unenforceable, given the existence of protective statute law and the authority of the parties to a contract to freely determine their terms.

Was Lady Hale Right to Worry About The Weight Courts Give Children in Their Decision Making?

Sahira Xec

Abstract

In law, a child's autonomy refers to their capabilities and capacities to form decisions that may affect their personal lives, yet it remains a controversial and unpredictable principle applied in practice. This article evaluates Lady Hale's disquiet around the value placed on children's wishes during judicial decision-making. With reference to the welfare principle, the Gillick competency principle, and the UN Convention on the Rights of a Child, this article aims to effectively examine whether the autonomy of a child is satisfactorily respected. Much of the case law surrounding this area aids in presenting the court's frequent prioritisation of welfare over voiced wishes of the child, particularly when the two intersect. With this, overarching academic arguments made by Ferguson, Eekelaar and Herring are incorporated in order to dissect the tension between protection of the child and active involvement from the child and others. Ultimately, this article finalises the assumption that Lady Hale's concerns are in fact valid and rational.

Introduction

According to Ferguson, our legal system does not have a 'child-centred theory of children's rights' that could positively impact or improve the outcomes in legal practice.[25] Initially, Lady Hale voiced the development of the courts listening to the voice of the child in family law cases, expressing how this evolution positively prioritises the wishes of the child in *Re D*.[26] However, Hale's later speech appears to be more in line with the opinion of many critics, including Ferguson, arguing that children's autonomy and wishes in matters of upbringing are too overlooked by the courts.

This essay will be discussing how Lady Hale is correct to show concern regarding the weight courts place on the wishes of children. This will be achieved through discussion of concepts such as the 'welfare principle', the 'Gillick competency principle', and reference to the UNCRC. These key concepts, alongside case law which concern the 'upbringing' of

[25] L. Ferguson, 'Not Merely Rights for Children but Children's Rights: The Theory Gap and the Assumption of the Importance of Children's Rights' (2013), 21 International Journal of Children's Rights, p.177–208.
[26] Re D (Abduction: Rights of Custody) [2006] UKHL 51, p.57.

children, whether medical or familial, will be reviewed thoroughly to conclude whether Lady Hale's apprehension for children's rights and autonomy is justified.

1. The Welfare Principle

The Children Act 1989[27] sets out the principle that a child's welfare should always be the court's paramount consideration when determining any question of a child's upbringing or administration of property or application of income. The act also establishes a 'welfare checklist', which outlines how the courts shall have regard to 'the ascertainable wishes and feelings of the child concerned, in light of his age and understanding'.[28] Cases such as *Re P*[29] show the courts paying strong attention to the child's welfare, whilst also granting the child's wishes.

Yet, there are many occasions where what the child wants and what is best for their welfare simply do not align, meaning their wishes are not always listened to by courts. For example, in *Re B*[30], despite wanting a change of surname due to their poor relationship with their father, Justice Wilson held that it would be in the children's best interests to keep the surname and reconnect with their father.

Undoubtedly, it is evident how crucial this welfare principle is in children's law, as seen in cases such as *J v C*.[31] Lord MacDermott stated here that the child's welfare can outweigh the rights and wishes of even unimpeachable parents, following assessment of the issue.[32]

However, there is the possibility that a child's own wishes may be imposed onto them by their parent, which Federle acknowledges by addressing how 'rather than empowering children through rights, we empower ourselves [adults] to intervene in their lives'.[33] Objectively using the welfare principle avoids feeding into any manipulation that may have shaped the child's decision.

An overarching issue which is often discussed by academics is how the viewpoints of a parent should not be entwined with the views of the child. Supporting Federle's claim,

[27] S.1(1), Children Act 1989.
[28] S.1(3), ibid.
[29] Re P (Residence Order: Child's Welfare) (1999) CA.
[30] Re B (Minors) (Change of Surname) [1996] 1 FLR 791.
[31] J v C [1970] AC 668.
[32] Ibid, p.715.
[33] K. H. Federle, 'Rights Flow Downhill' (1994), 2 International Journal of Children's Rights 343, p.365.

Eekelaar proposes that we should take a 'detached view' of the welfare principle, where the child's views are seen outside of their carer's interests to avoid them getting lost.[34] This would encourage the child's interests to be heard at a sufficient weight. Although, Eekelaar does acknowledge the weakness in this proposal, that there is no clear way to appropriately balance the interests of each party. Considering even a parent's right can be overlooked, children's interests are also likely to be outweighed.

There is also additional criticism from Herring, who claims that by isolating the child's interests from the caregivers', it could remove what is also valued by the child (referring to the child's need for their parent to also be rightly satisfied).[35] This could lead to the child's wishes being overlooked based on the weight placed on each interest. *Re P* gives sufficient support for this claim, where Butler-Sloss LJ stated that the welfare of N was 'crucially dependent' on the welfare of her carers.[36] It is vital to establish a balance between the manipulation a carer may have had on a child's views and the interests of the child and carer on a case-by-case basis.

A further criticism of the welfare principle is that it is indeterminate, allowing for much judicial discretion and, therefore, uneven application of children's wishes, in contrast with respecting the paramountcy principle.

Alongside this, judges are thought to hold very subjective powers, where they can claim that anything could have a potential effect on a child's welfare, regardless of if it truly does or not.[37]. A Canadian Judge once stated when determining a child's best interests that 'judges are tied by invisible threads of their own convictions'.[38] This suggests that judges` personal views are likely to be reflected in their decisions, including whether a child's wishes are listened to or not. After all, King and Piper point out there is no way of objectively measuring what would be best for a child's welfare, except when scientific evidence is presented.[39] However, it can be argued that it is impossible for the judiciary to fully void their decisions of opinion; their conclusions and thought processes must be justified. Though Reece's view is that it is ironic that judges try to adhere so strongly to the paramountcy principle despite

[34] J. Eekelaar, 'Beyond the Welfare Principle', (2002), Child and Family Law Quarterly, Vol.14, p.243.
[35] J. Herring, 'Farewell Welfare?', (2005), Journal of Social Welfare and Family Law, Vol.27, p.167.
[36] Ibid, p.31.
[37] J. Eekelaar, 'Regulating Divorce' (1991), Clarendon Press, Oxford, p.248.
[38] Rockwell v Rockwell (1998), 43 R.F.L.
[39] M. King & C. Piper, 'How the Law Thinks About Children' (1995), Arena, Aldershot.

their prejudice, whilst suggesting we could consider a more determinate principle, such as the primary caretaker rule, as it would serve the child's best interests well.[40]

An argument in favour of the welfare principle is that children require protection, as they are the most vulnerable in the eyes of the law, and thus require their welfare to be protected.[41] The consensus is although it is argued for children's wishes to be heard, their welfare should still remain paramount, and as long as they are receiving this legal protection, it should not be protested against. The Law Commission similarly stated that the paramountcy principle is very important and so modifying it could put children's welfare at risk; this concept needs to be strengthened rather than replaced.[42]

Furthermore, there are a host of conventions which promote children's rights and interests and also follow the welfare principle, such as the UNCRC. [43] Section 1(1) of the Children Act[44] aligns perfectly with Article 3 of the UNCRC[45] by prioritising the child's best interests. To abandon the welfare principle would send a wrong message, in which fundamental rights would be downplayed.[46]

However, Article 12 of the convention highlights a child's rights to express their views in matters which affect them. The courts not listening to a child's wishes could be seen as breaching this article. This has happened in previous cases, for example *Re W*,[47] which concerned a sixteen-year-old girl who refused medical treatment. Despite her maturity, the courts held that refusal of treatment by the child could certainly be overridden by judges if they reason it as appropriate. Though, this was subject to a limitation of being too far-reaching and essentially making the jurisdiction of the court limitless, surpassing the powers of a parent.

2. Gillick Competency Principle

The usage of the Gillick principle from the courts reaffirmed children's autonomy in the legal sphere, with the correct assessment of a child's circumstances leading to opportunities

[40] H. Reece, 'The Paramountcy Principle: Consensus or Construct?', (1996), Current Legal Problems, Vol.49, p.274.
[41] J. Herring (N 9), p.168.
[42] Law Commission Working Paper No.96 (1986), p.6.17.
[43] United Convention on the Rights of the Child (1989).
[44] Ibid, (N 3)
[45] Ibid, (N 19), Article 3.
[46] J. Herring (N 17).
[47] Re W [1992] 4 All ER 627.

open for those under eighteen to make their own choices and decisions. *Gillick*[48] was a landmark case about a child's access to contraception without the consent of a parent, in which judges established a test which could determine 'Gillick competency'. This essentially describes a child having sufficient understanding and intelligence, which can allow them to make decisions on the matter, independent of their parents.[49] Despite being a medical case, the decision in Gillick extended to all child upbringing cases.

However, Lord Fraser once explained that the rights of a parent regarding their child's upbringing end once the child turns eighteen and until then, the courts will be hesitant to enforce the wishes of the child[50]; this will decrease as the child matures, as the right of control turns to advice.[51] This statement demonstrates that the courts do have an overarching prejudice and resistance to listening to the wishes of younger children due to their age, thus limiting their autonomy and supporting Lady Hale's original worry. Lord Fraser's judgement and approach in Gillick seems to prioritise the child's best interests rather than their personal wishes. Weaknesses of Lord Fraser's approach lie in the fact that it is very limited and can only be used for cases of contraception, unlike Lord Scarman's which can be used generally. It is also not holistic in the sense of looking to further a child's autonomy and focuses solely on parental rights and welfare. Though, it is a very clear, straightforward and easily interpreted test which means it allows for little gaps in the law.

Lord Scarman, on the other hand, contrasts with the opinion of Lord Fraser by establishing a view which focuses more on the child's capacities as well as their autonomy. Eekelaar noted that Lord Scarman's opinion rests on the fact that once the child is at full capacity, a parent cannot impose their own view or decision on them, regardless of whether it aligns more with the child's best interests or not; the parent's rights are terminated once the child reaches capacity.[52] However, it can be argued that the first condition in Lord Fraser's test is that the child has an understanding, which aligns with Lord Scarman's approach.

Initially, courts seemed to be utilising the decision laid out in Gillick, by respecting a child's wishes and autonomy provided they were Gillick competent. Shortly after this, Lord Donaldson introduced controversial obiter remarks in *Re R*, distinguishing between determination and consent and claiming that even if a Gillick competent child were to refuse

[48] Gillick v West Norfolk AHA [1985] 3 All ER 402.
[49] [1986] 1 AC 112, p.186.
[50] Ibid.
[51] Hewer v Bryant [1970] 1 QB 357, p.369.
[52] J. Eekelaar, 'The Emergence of Children's Rights', (1986), 6 Oxford Journal of Legal Studies 161, p.181.

treatment, the treatment should be lawfully administered if an individual with parental responsibility consents.[53]

Re W[54] furthered the retreat from Gillick, effectively following Lord Donaldson's remarks and ruling that regardless of Gillick's competence, the right of consent does not allow for rejection of treatment in addition to the power of accepting it. The court decided in a way which went against W's wishes, regardless of her showing Gillick competence. Nolan LJ was guided by the welfare principle in his judgment, considering the welfare of the child over her wishes.[55]

Ferguson defends this in a way, claiming that not all children possess the capacity to make decisions and so a caregiver should rightly be the capable exerciser of their rights on the child's behalf.[56] So, although children should be given some autonomy in their decisions, not all children can decide alone, and so subjecting all children to autonomy may lead to wrong decisions being made for the child's welfare; a parent exercising their rights makes it less likely to end badly. Also, if a child is wrongly decided as Gillick competent and proceeds to make the wrong decision, it could have negative implications; therefore, having a parent as a decision-maker may be the safer choice.

Supporting this view, Eekelaar adds that subjecting all children to having their decisions made for them may be the safer option to protect their fundamental and developmental interests. After all, it is unclear to us whether a child may later look back and regret having made their own independent decisions rather than being given the appropriate guidance from their caregivers.[57]

Despite this, he does observe that the risk that comes with allowing children to have more independence in their decision-making is a valuable right which allows them to make and learn from mistakes; a view which aligns with Lady Hale's concern for children's wishes.[58]

Ferguson insinuates another view that capacity is difficult to be measured, therefore using it to determine whether a child should have their wishes heard or not is nonsensical.

[53] Re R (a minor) (wardship: consent to treatment) [1991] 4 All ER 177, p.129.
[54] Ibid, (N 23).
[55] Ibid, p.156.
[56] Ibid (N 1), p.17.
[57] Ibid, (N 28), p.181.
[58] Ibid, (N 28), p.182.

Additionally, she claims capacity varies depending on the complexity of the decision to be made, meaning application of the procedure could be uneven.[59]

Avoiding the allowance of children's wishes can also be said to be justifiable due to the protection of children's lives. In cases such as *Re E*[60]and *NHS Trust v CX*[61], both children were Gillick competent and refusing blood transfusions due to religious teachings. Treatment was still allowed despite this as the preservation of life overpowered personal autonomy; the courts could simply not ignore the risk that decision would pose to the children's lives.

Conclusion

The intention of this essay was to examine whether Lady Hale was correct to be concerned about the consideration courts give to the wishes of children, as the assessment of the welfare principle alongside the Gillick principle has suggested.

Whilst Lady Hale's concerns are generally justified, the uneven usage of both the welfare principle as well as the Gillick competency assessment makes it difficult to recognise whether children's wishes are well-heard or not. The courts do attempt to remain in line with ECHR principles, which empower children's autonomy, but they do still have a larger focus on the welfare of the child. It is important to note that ensuring the welfare of the child means that sometimes the child's views are not heard, but this can be acknowledged as being the most favourable option among other undesirable outcomes.

[59] Ibid, p.16.
[60] Re E (Children: Blood Transfusion) [2021] EWCA Civ 188.
[61] An NHS Trust v CX & Others [2019] EWHC 3033 (Fam).

BIBLIOGRAPHY

Eekelaar J, 'Beyond the Welfare Principle', (2002), Child and Family Law Quarterly, Vol.14

Eekelaar J, 'Regulating Divorce' (1991), Clarendon Press, Oxford

Eekelaar J, 'The Emergence of Children's Rights', (1986), 6 Oxford Journal of Legal Studies 161

Ferguson L, 'Not Merely Rights for Children but Children's Rights: The Theory Gap and the Assumption of the Importance of Children's Rights' (2013), 21 International Journal of Children's Rights

Federle K H, 'Rights Flow Downhill' (1994), 2 International Journal of Children's Rights 343

Herring J, 'Farewell Welfare?', (2005), Journal of Social Welfare and Family Law, Vol.27

King M & Piper C, 'How the Law Thinks About Children' (1995), Arena, Aldershot

Law Commission Working Paper No.96 (1986)

Reece H, 'The Paramountcy Principle: Consensus or Construct?', (1996), Current Legal Problems, Vol.49

Table Of Cases

<u>UK Cases</u>

An NHS Trust v CX & Others [2019] EWHC 3033 (Fam)

Gillick v West Norfolk AHA [1985] 3 All ER 402

Gillick v West Norfolk AHA [1986] 1 AC 112

Hewer v Bryant [1970] 1 QB 357

J v C [1970] AC 668

Re B (Minors) (Change of Surname) [1996] 1 FLR 791

Re D (Abduction: Rights of Custody) [2006] UKHL 51, p.57

Re E (Children: Blood Transfusion) [2021] EWCA Civ 188

Re P (Residence Order: Child's Welfare) (1999) CA

Re R (a minor) (wardship: consent to treatment) [1991] 4 All ER 177

Re W [1992] 4 All ER 627

Canadian Cases

Rockwell v Rockwell (1998), 43 R.F.L

Table of Legislation

Children Act 1989

United Convention on the Rights of the Child (1989)

The Commodification of Housing: Neoliberalism, Financialisation, and The Crisis of The Private Rented Sector

Erin Guizot

Abstract:

The aim of this article is to critically analyse how housing in the private rented sector has been treated as a commodity, leaving tenants in an inferior position within the housing system. Historically, the emphasis on deregulation illustrates how housing has shifted from a home to a profit-motivated business. The commodification of housing emphasises a need for tenants to play a greater role in housing policymaking to undermine the current prioritisation of landlords and investors interests. This will be achieved through assessing key conceptual ideas that flow from commodification. The continued influence of neoliberalism on policies fuels the view of housing as an investment rather than a home. Therefore, it will be suggested that policies need to prioritise the needs of tenants more to create a more balanced housing system that is less commodified.

Introduction

The commodification of housing, particularly within the private rented sector (PRS), has increasingly treated homes as profit generating assets rather than as essential social goods. The perception of housing has transformed from a basic human need to an investment. This ideological shift is particularly prevalent in the PRS where deregulatory policies, limited social housing stock and emphasis on investment returns have contributed to a housing environment which is characterised by unaffordability, insecurity and poor standards. Consequently, the PRS serves as a prime example of a housing market which is driven by profit rather than shelter. This shift has been attributed to economic and political policies which have driven the commodification of housing such as neoliberalism and financialisation.

This article argues that neoliberalism and financialisation have played central roles in this shift, leaving tenants in a vulnerable position as market dynamics dominate social needs. Beginning with an analysis of the contested concept of neoliberalism, the article explores its influence on housing policy and its role in shaping the PRS. It then explores how financialisation has allowed housing to be viewed as an asset, using examples such as

deregulated mortgage markets and platforms like Airbnb to show how this shift displaces residents and increases inequality. The discussion then turns to the broader process of commodification, revealing how the prioritisation of investment interests over tenant wellbeing results in insecurity, poor housing standards, and rising unaffordability. Finally, the paper draws on legal and policy perspectives to highlight the need for decommodification through stronger tenant protections and a renewed focus on the housing system's social function. By situating tenants at the centre of housing policy, the article concludes that the law must play an active role in addressing the housing market to serve its fundamental purpose as a home, not merely an investment.

1. Neoliberalism As a Term:

To understand the commodification of housing, this article begins by unpacking the contested but significant concept of neoliberalism, which informs the backdrop for modern housing policy. Neoliberalism has been defined as 'a political project involving reduced state intervention in social and economic affairs and the assertion of the superiority of market processes.'[62] Although Forrest and Hirayama have attempted to create a precise definition, 'the neoliberal phenomenon is notoriously hard to define'.[63]

In addition, the origins of neoliberalism are contested. According to Harvey, following the failure of global capitalism in the 1960s, 'the world stumbled towards neoliberalism as the answer through a series of gyrations and chaotic motions.'[64] Harvey therefore implies neoliberalism began in the 1960s, which contradicts the popular viewpoint that Margaret Thatcher introduced neoliberalism to the UK through the Right to Buy (RTB) in 1980.[65]

Firstly, neoliberalism has been attributed to many social issues such as stagnant living standards and inequality.[66] The term has attracted heavy criticism, yet it remains elusive and ill-defined. Well known figures like Friedrich Hayek advocated for neoliberalism[67] while

[62] Ray Forrest and Yosuke Hirayama, 'The Uneven Impact of Neoliberalism on Housing Opportunities' (2009) 33 International Journal of Urban and Regional Research 998, 998.
[63] Tibor Rutar, 'What Is Neoliberalism Really? A Global Analysis of Its Real-World Consequences for Development, Inequality, and Democracy' (2023) 62 Social Science Information 295, 296.
[64] David Harvey, 'Neoliberalism as Creative Destruction' (2007) 610 *The Annals of the American Academy of Political and Social Science* 22, 27.
[65] Aled Davies, James Freeman and Hugh Pemberton, 'Thatcher's Policy Unit and the "Neoliberal Vision"' (2023) 62 *Journal of British Studies* 77, 78.
[66] Rutar (n 2) 296.
[67] Stephen Metcalf, 'Neoliberalism: The Idea That Swallowed the World' *The Guardian* (18 August 2017) Neoliberalism: the idea that swallowed the world | Economic policy | The Guardian (accessed 7th February 2024).

some authors like Harvey labelled neoliberalism as 'creative destruction'.[68] These conflicting opinions about the meaning and origins of neoliberalism pose problems for dependency on the term as an explanation of our 'broken housing market'.[69] Venugopal argues that due to the ambiguity surrounding neoliberalism, only 'its descriptive shell'[70] can be usefully preserved. However, despite the contested nature, neoliberalism remains a marker of broader economic trends, particularly in the housing market.

Despite the efforts of numerous scholars, no clear and universally accepted definition of neoliberalism has been agreed upon. Rutar identifies this as a particular problem in studies where neoliberalism is the independent or dependent variable in their empirical studies as a clear understanding of neoliberalism is often missing.[71] This ambiguity is inevitably problematic with different scholars conflicting with one another's interpretation. Recognising this issue, Rutar has intervened to try and create a broad and empirically minded working definition of neoliberalism. In order to do this, Rutar views neoliberalism as a 'cluster of phenomena that has been associated with liberalisation, privatisation, deregulation or globalisation'.[72]

By adopting a multifaceted perspective on neoliberalism, the term is less prone to being misused without substantive meaning. Rutar's approach suggests that a society's degree of neoliberalism can be identified by examining components like liberalised markets with few price controls and trade barriers, modest government regulation, and tight money supply.[73] This comprehensive understanding of neoliberalism enables the term to offer more utility as a description of a country's economic state or policy objectives.

Rutar argues that the use of a cluster definition can alleviate some of the issues scholars face when describing neoliberal policies. This is because 'when the effort is made to provide a concrete definition of the term, there are multiple distinct definitions on offer. '[74] Therefore, Rutar was careful to avoid single and precise definitions, hence the recognition of neoliberalism as a cluster of components.

---

[68] Harvey (n3) 33.
[69] Ministry of Housing, Communities and Local Government, 'Fixing Our Broken Housing Market' (Cm 9352, 2017) https://www.gov.uk/government/publications/fixing-our-broken-housing-market (accessed 27th January 2024).
[70] Rajesh Venugopal, 'Neoliberalism as Concept' (2015) 44 *Economy and Society* 165, 182.
[71] Rutar (n 2) 299.
[72] ibid, 300.
[73] ibid, 302.
[74] ibid, 317.

This approach suggests that neoliberalism may not be a specific and rigid ideology, but rather a range of components. Accordingly, the primary issue may not lie in defining neoliberalism, but in the process of attempting to define the term. Many scholars sought to provide a concrete definition, but this has failed to adequately represent the complexity of neoliberalism.

2. Neoliberalism in The Private Rented Sector:

In the literature regarding the commodification of housing, neoliberal policies have often received the blame, either wholly or partially, for being responsible for this process. Neoliberalism has been perceived as an important factor in the transformation of housing into a business and investment opportunity.

Harvey suggests that neoliberalism incorporates four elements: privatisation, financialisation, management and manipulation of crises and state redistributions,[75] which can be applied to the housing sector. This article suggests that these four elements have been instrumental in turning housing into a commodity.

Firstly, privatisation has been exemplified through policies like the Right to Buy Scheme which has shifted social welfare provision like housing from the public to the private domain. The implementation of the Housing Act 1980[76], which allowed council tenants to purchase their homes at a discount, increased home ownership in the UK, but also reduced the available social housing stock. After the housing market's description as 'sclerotic, state-dependent and anti-enterprise regulatory culture'[77], the state reduced paternalistic policies which had a significant effect on social housing with 1,992,799 homes being sold through RTB between 1980 and 2021.[78] The erosion of state intervention, which has been prominent in Conservative housing policy, has led to the decline of social housing and a rise in the PRS. Consequently, those who are unable to afford their own homes are left with minimal options. This illustrates how although the RTB was implemented in the 1980s, the consequences are still seen in our current housing market in the context of the housing crisis.

[75] Harvey (n3) 35.
[76] Housing Act 1980.
[77] John Bone, 'Neoliberal Nomads: Housing Insecurity and the Revival of Private Renting in the UK' (2014) 19 *Sociological Research Online* 1, 2.
[78] Department for Levelling Up, Housing and Communities, 'Social Housing Sales and Demolitions 2020-21: Right to Buy Sales' (2022) https://www.gov.uk/government/statistics/social-housing-sales-and-demolitions-2020-21-england/social-housing-sales-and-demolitions-2020-21-right-to-buy-sales (accessed 4th April 2024) 1.

Secondly, regarding financialisation, deregulation in particular has allowed the financial system to have significant influence over redistributive activity.[79] Due to the financial institutions having greater involvement in housing policies like the RTB and deregulation of the financial sector,[80] housing is becoming more of an asset than a home. Land has become financialised, with the planning process geared towards facilitating investment rather than attending to socio-economic needs.[81] Whilst this serves the interests of landlords and investors, this fails to serve the interests of people using the housing market, highlighting the continuing notion of home as a business.

Thirdly, Harvey identifies neoliberalism as having a 'deeper process' of crisis creation, management, and manipulation to redistribute wealth.[82] Once the state has transformed into a neoliberal body, it can implement redistributive policies via a reduction in the welfare state and privatisation.[83] For instance, Harvey analysed Thatcher's election, deemed symbolic of neoliberal ideas in public and social policy initiatives,[84] and argued that although the RTB 'appeared as a gift', upon deeper analysis, the loss of affordable housing produced increased homelessness.[85] The erosion of the welfare state has moved the 'housing stock towards a more fully fledged neoliberal agenda'[86] which leaves tenants vulnerable to market conditions and economic crises.

The future impact of neoliberalism on housing remains vague and unclear. However, the existing deregulation and lack of state intervention illustrate how housing has been left to market forces. Consequently, neoliberalism still influences housing policies and could still play a future role depending on the regulatory decisions of successive governments. Therefore, if neoliberalism persists, the prospect of decommodification diminishes. These shifts have created a market optimal for the financialisation of housing, where homes increasingly function as financial assets.

[79] Harvey (n3) 36.
[80] Keith Jacobs and Tony Manzi, 'Conceptualising "Financialisation": Governance, Organisational Behaviour and Social Interaction in UK Housing' (2020) 20 *International Journal of Housing* Policy 184, 187.
[81] ibid, 190.
[82] Harvey (n3) 37.
[83] ibid, 38.
[84] Forrest and Hirayama (n1) 1002.
[85] Harvey (n3) 38.
[86] Forrest and Hirayama (n1) 1003.

3. Financialisation in The Private Rented Sector

Having outlined neoliberalism's policy impacts on the PRS, this section turns to financialisation, which embeds the neoliberal ideology. Scholars define financialisation as 'the increasing dominance of financial actors, markets, practices, measurements, and narratives, at various scales, resulting in a structural transformation of economies, firms, states and households.'[87] Despite this definition, the term is viewed as elusive and vague which creates hardship when explaining financialisation's impact on housing. Nonetheless, Aalbers argues that the multiple definitions of financialisation is 'hardly proof of the problematic nature of the concept'[88] and therefore it is still useful.

Financialisation's association with housing primarily stems from the deregulation of mortgage markets in the 1980s, which led to a mortgage explosion.[89] Financial deregulation increased private borrowing, which made credit more available alongside policies like the RTB.[90] Liberalising the mortgage markets allowed individuals to take out more unsecured loans which increased debts and house prices in the UK.[91] Therefore, financialisation is linked to 'individual households made vulnerable to predatory lending practices and the volatility of markets, the result of which is unprecedented housing precarity.'[92]

However, 'financialisation is not a one-size-fits-all process…rather, it is fundamentally fragmented, path-dependent and variegated.'[93] Therefore, the impact on housing is not uniform, but its link to deregulation helps explain the inconsistent pathway of regulation in the UK. Despite the non-linear trajectory, housing is still a central aspect of financialisation.[94]

The UN has recognised the negative impact of financialisation on housing, emphasising that referring to housing as capital investment disconnects housing from its social function of

[87] Rodrigo Fernandez and Manuel B Aalbers, 'Financialization and Housing: Between Globalization and Varieties of Capitalism' (2016) 20 Competition & Change 71, 71.
[88] Manuel B Aalbers, 'The Variegated Financialization of Housing' (2017) 41 International Journal of Urban and Regional Research 542, 543.
[89] Sebastian Kohl, 'Too Much Mortgage Debt? The Effect of Housing Financialization on Housing Supply and Residential Capital Formation' (2020) 19 *Socio-Economic Review* 413, 414.
[90] Grace Blakeley, 'Financialization, Real Estate and COVID-19 in the UK' (2021) *56 Community Development Journal* 79, 80.
[91] Kohl (n27) 416.
[92] UNHCR, *Report of the Special Rapporteur on adequate housing as a component of the right to an adequate standard of living, and on the right to non-discrimination in this context* (18 January 2017) UN Doc A/HRC/34/51, 5.
[93] Aalbers (n27) 544.
[94] ibid.

providing a secure place to live.[95] Financialisation has exacerbated housing inequalities through benefitting homeowners and investors, whilst younger people or those on lower incomes are most disadvantaged.[96] This leads to a growing polarisation of housing opportunities as houses shift from a place to live to a place to grow capital.[97]

The growing emergence of platforms such as Airbnb illustrates financialisation of housing in practice. Financial deregulation in the 1980s created rising house prices, homeownership, and the creation of real estate bubbles.[98] It is said that 'real estate platforms have produced an infrastructure that makes it easy to switch the uses and markets of the properties.'[99]

Rozena and Lees looked at the impact of 'Airbnbification' on the everyday residents in a specific building on Kensington High Street in London's Kensington district.[100] The authors pointed out that profit is being made from gentrified cities and is facilitating processes like capital investment in housing.[101] This increased housing investment is indirectly displacing residents through changes in the unique character and social identity of the area they live in.[102]

Having said this, Rozena and Lees point out that the law has attempted to regulate Airbnb listings in London through the Deregulation Act 2015. Sections 44 and 45 of the Deregulation Act 2015 state that 'temporary sleeping accommodation of any residential premises in Greater London' cannot 'exceed ninety nights per year.'[103] However, these regulations are easily surpassed with Camden having 104 estimated nights used per year.[104] Another way of surpassing the ninety-night limit is by providing multiple listings for individual rooms in the same building.[105] As a result, Airbnb has turned residential accommodation into a tourist hotspot. The lure of profit and short-term rentals has facilitated

[95] UNCHR (n31) 1.
[96] Dallas Rogers and Emma Power, 'Explainer: The Financialisation of Housing and What Can Be Done about It' (*The Conversation*, 23 March 2017) http://theconversation.com/explainer-the-financialisation-of-housing-and-what-can-be-done-about-it-73767 (accessed 16th February 2024).
[97] ibid.
[98] Javier Gil, Pablo Martínez and Jorge Sequera, 'The Neoliberal Tenant Dystopia: Digital Polyplatform Rentierism, the Hybridization of Platform-Based Rental Markets and Financialization of Housing' (2023) 137 Cities 1, 2.
[99]ibid, 3.
[100] Sharda Rozena and Loretta Lees, 'The Everyday Lived Experiences of Airbnbification in London' (2023) 24 *Social & Cultural Geography* 253, 253.
[101] ibid, 254.
[102] ibid.
[103] ibid, 257.
[104] ibid, 258.
[105] ibid.

commodification by removing valuable housing supply from the market, which emphasises how residents' needs are overlooked when profit is at stake.

The abundance of Airbnb as a form of housing investment is not limited to London, or even the UK. Similar findings in Madrid highlight the global nature of the issue. The ease of moving properties across markets accelerates capital turnover, allowing housing to become a financial asset.[106] This creates a platform hybridisation whereby housing is extracted from the residential market to generate higher rent which amplifies the exchange value of housing and future profit expectations.[107] 'For tenants, the consolidation of digital poly platform rentierism and the hybridisation of platform-based rental markets produces a neoliberal tenant dystopia.'[108]

To address the problems with financialisation, 'steps must be taken to insulate the housing market from financial markets and consumer markets.'[109] This includes controlling rents and increasing the stock of social housing to de-financialise and de-commodify the housing system.[110] The turn toward housing as a financial asset intensifies the commodification process, reducing homes to vehicles for profit maximisation. The next section explores commodification in more detail, especially its consequences for tenants.

4. Commodification of Housing

Madden and Marcuse define commodification as 'the name for the general process by which the economic value of a thing comes to dominate its other uses'.[111] In the housing context, investment prospects are therefore prioritised over housing's role as a lived space. This shift has led to a disconnect between the social value a home holds and its treatment as a profit opportunity. As a result, tenants are disadvantaged in the PRS whose needs are often overlooked when the availability of profit is more attractive. This reflects the neoliberal and financialised treatment of housing, highlighting how commodification acts as both a consequence and a reinforcement of these processes.

One major factor contributing to the prioritisation of housing as real estate is the influx of investment, especially foreign investment, into the housing market. This has led to a volatile

[106] Gil, Martínez and Sequera (n37) 8.
[107] ibid, 9.
[108] ibid, 10.
[109] Blakeley (n29) 82.
[110] ibid, 94.
[111] David Madden and Peter Marcuse, *In Defense of Housing* (Verso, 2016) 17.

housing market where tenants face the threat of rent increases and evictions as housing is used for financial gain rather than a long-term housing solution. Watt and Minton stress that cities like London are 'experiencing a vicious intertwining between housing as a speculative investment vehicle and housing as an agent of social insecurity'.[112] This implies a trade-off exists between investors interests and tenants' interests, but as this article has revealed, tenants do not possess the same bargaining power as investors.

In particular, foreign investment has led to a surge in demand for luxury real estate in cities which has signalled to developers to extend their portfolio of high-end developments. However, with 42% of properties owned by foreign nationals being vacant,[113] these developments worsen the housing crisis, which forces policymakers to start debates on how to tame the financialisation of housing.[114]

With investment demand increasing, house prices are rising which culminates in a lack of affordable housing. The Affordable Homes Programme (AHP) aims to address this lack of affordable housing by delivering up to 180,000 homes,[115] however this is of limited success with only 59,175 being built between 2021 and 2022.[116] This could be partly attributed to developers' reluctance to pursue projects that are not profitable to them. This exposes the detrimental effects of commodification to renters as basic needs like shelter become unattainable due to the money-orientated nature of housing.

5. Meaning of Home:

The shift from viewing housing as a home to an investment has led to a failure to recognise the social meaning of home. Policies favour the economic value of a home over social functions, thus neglects the needs of many PRS renters. Fox-O'Mahony believes the social value is important in housing policies as 'it is this relationship which marks home out as

[112] Paul Watt and Anna Minton, 'London's Housing Crisis and Its Activisms: Introduction' (2016) 20 City 204, 206.
[113] Rowland Atkinson 'London's Extraordinary Surplus of Empty Luxury Apartments Revealed' *The Conversation* (26 October 2018) London's extraordinary surplus of empty luxury apartments revealed (London's extraordinary surplus of empty luxury apartments revealed) (accessed 18th November 2023).
[114] Michelle Norris and Julie Lawson, 'Tools to Tame the Financialisation of Housing' (2023) 28 *New Political Economy* 363, 375.
[115] Wendy Wilson and Cassie Barton, 'What Is Affordable Housing?' (House of Commons Library 2022) https://researchbriefings.files.parliament.uk/documents/CDP-2023-0091/CDP-2023-0091.pdf, 6.
[116] Department for Levelling Up, Housing and Communities, 'Affordable Housing Supply in England: 2021 to 2022' (2022) https://www.gov.uk/government/statistics/affordable-housing-supply-in-england-2021-to-2022/affordable-housing-supply-in-england-2021-to-2022 (accessed 18th November 2023) 1.

different from other types of property*'*.[117] Therefore, it is essential for policies to recognise this relationship and address housing insecurity.

At the base of Maslow's hierarchy of needs are physiological requirements, which Maslow describes as 'the most proponent of needs.[118] Physiological needs would include basic necessities like shelter, which underscores the importance of home. While Maslow's pyramid is a useful tool to explain the significance of a home, it oversimplifies the complex factors that interact with housing. Housing involves various considerations such as the lack of supply and affordability issues that are not reflected in the hierarchy.

Following Fox-O'Mahony, the concept of home goes beyond mere bricks and mortar and is something much more complex. Bingham supports this notion stating that 'there are few things more central to the enjoyment of human life than having somewhere to live*'*.[119] Therefore, if the judges in the case law can recognise the intrinsic meaning of a home, then statutes should follow the same approach. This suggests that a clearer legal conceptualisation of home would potentially mitigate the issues associated with commodification. If the law better understood the function and value of a home beyond economic ones, then a more nuanced legal framework could be provided to better protect tenants.

Failure to recognise housing as a lived space has contributed to the notorious reputation of the PRS for having poor housing standards. According to the English Housing Survey, 23% of PRS dwellings are thought to fail the Decent Homes Standard.[120] This percentage appears even more unjust when learning that 33% of PRS tenants' incomes are spent on rent, which is the highest percentage compared to social renters and owner occupiers.[121] The contradiction in tenants paying the largest proportion of their income for the worst quality homes epitomises the negative effects of commodification and profit fixation. This reveals that in a commodified market, tenants suffer in poor housing conditions as attention is diverted to maximising profits rather than improving the living standards.

---

[117] Lorna Fox O'Mahony, 'The Meaning of Home: From Theory to Practice' (2013) *5 International Journal of Law in the Built Environment* 156, 157.
[118] Abraham Maslow*, Motivation and Personality* (Harper, 1954) 82.
[119] *Qazi v Harrow London Borough Council* (2003) UKHL 43 (2004) 1 AC 983 [8].
[120] Department for Levelling Up, Housing and Communities, 'English Housing Survey 2021 to 2022: Private Rented Sector' (2023) https://www.gov.uk/government/statistics/english-housing-survey-2021-to-2022-private-rented-sector/english-housing-survey-2021-to-2022-private-rented-sector (accessed 16th November 2023) 4.
[121] ibid, 2.

This section has illustrated the dangers of viewing housing as a commodity as it ignores tenants in housing policymaking. Madden and Marcuse emphasise this point by saying, 'housing policy has consistently been designed to meet the economic needs of the real estate industry and the political needs of those running the state'.[122] There is an urgent need for decommodification in the housing sector to reposition residents at the centre of the housing agenda and ensure the sector is fit for purpose. As it currently stands, there is a desperate need for reform in the housing market to move away from business ideologies to those that recognise the intrinsic value of the home.

Conclusion

This article has argued that commodification of housing, fuelled by neoliberalism, financialisation and deregulation have shaped PRS housing into something which is 'produced and distributed as a commodity to enrich the few'.[123] Through an exploration of key concepts such as neoliberalism, financialisation, and the social meaning of home, this article has demonstrated how housing has been transformed from a basic human necessity into an investment.

The analysis demonstrates that policies such as the RTB, the deregulation of financial markets, and the expansion of short-term rental platforms like Airbnb have collectively reinforced a housing system that prioritises market efficiency over tenant welfare. Particularly, financialisation has increased rent burdens and restricted access to affordable accommodation, thus increasing housing insecurity. This shift has been accompanied by a broader failure to recognise housing as a social good, contributing to rising inequality and a decline in living standards across the PRS.

These findings have highlighted the need to re-centre the social value of housing. If housing continues to be viewed in the current framework, the prospect of effective reform and decommodification diminishes. To achieve such reform, legal and policy interventions which focus on affordability, security and tenant wellbeing are needed.

Future reforms need to move beyond market considerations and acknowledge housing as a home rather than an investment vehicle. Improving tenant protections, expanding social

[122] Madden andMarcuse (n50) 139.
[123] ibid, 10.

housing, and developing legal frameworks that prioritise security and affordability are essential steps toward addressing the structural imbalances identified in this article.

What Would A Desirable Future in The English & Welsh Criminal Justice System Look Like?

Christian Badcock

Abstract

The criminal justice system encompasses multiple institutions, agencies, and organisations including the police, courts, and prison services. It can loosely be described as a system in England & Wales which seeks to achieve justice, protect the public, and rehabilitate. Despite the importance of the criminal justice systems' aims, objectives, and principles the current systems' effectiveness is up for debate. Subsequently, this article will explore what a desirable criminal justice system could look like in England & Wales, with acknowledgment for how this could function, be achieved, and the role of victims. The article will primarily concern itself with England & Wales, however, there will also be consideration for the Scandinavian jurisdiction and how this could provide potential inspiration for what a less punitive system could look like. Overall, the article offers an analysis into the criminal justice system through a social justice-based way of thinking that incorporates the guiding principles of fairness and due process. It is not proposing a wholly abolitionist approach, and it is acknowledged that punishment should remain a part of our criminal justice system, but with more emphasis on rehabilitation than there is at present.

Introduction

This article will put forward an argument for what a desirable future English criminal justice system (CJS) might look like. Such a broad question brings in elements from all parts of the CJS such as the police, courts, Crown Prosecution Service, prison facilities and support staff, Parole Board, and Probation Service. This question will be answered by examining which theories and principles should underpin a desirable system, giving a couple of specific examples of impactful changes that could be made, and looking to the Scandinavian model for inspiration. This is more of a principled exercise setting out a direction to travel in as opposed to a 'how-to' guide for building the separate elements of an effective justice system.

1. Meaning of 'Justice' and 'CJS'

Before you can put forward a vision of a desirable future criminal justice system, it is necessary to understand what 'justice' and justice in the context of a criminal justice system

means. Zedner sets out six different conceptions of how criminal justice might be understood.[124] It has theoretical elements and can be seen as a field of study or a normative theory. The normative theory questions why things such as incarceration and forced labour are considered wrong 'yet the state does all of these things in the name of criminal justice.'[125] There is also a more positive account of criminal justice. This is where 'criminal justice is properly the duty of the state'. Other conceptions include a form of governance which includes imposing social order,[126] a purely legal response to wrongdoing,[127] and the agents and institutional practices that respond to crime.[128] The latter two conceptions are more descriptive. Due to the theoretical nature of this article, theoretical conceptions of the CJS are the most relevant when considering what a desirable CJS would look like. It is still useful to recognise the truth in the other conceptions to enable a rounded discussion. For example, realising the CJS is also a form of governance invites a discussion in a political context.

There is also debate as to whether the CJS can be considered a 'system'. The criminal justice institutions can be considered 'interdependent and closely related' and thus a system.[129] However, 'it does not describe reality' as there is not coordination in a structured way and 'all of the agencies in the criminal justice process have different and, sometimes, competing objectives.'[130] Zedner notes how most CJS institutions and practices are the 'product of historical accident.'[131] This unplanned approach can also be considered a strength as components are added when they are necessary. When designing a desirable CJS it would be worth implementing a more deliberate design to facilitate cooperation between all agencies, but this should not be at the expense of discretion and flexibility.

2. Current CJS Problems

Understanding the issues in the current CJS is useful as a yardstick to measure a new system against and to see where change is most needed. England and Wales have the 'highest per capita prison population in Western Europe'[132] and nearly one-quarter of prisoners are held

[124] Lucia Zedner, *Criminal Justice* (OUP, 2004), 1-36.
[125] Ibid, 25.
[126] Ibid, 4.
[127] Ibid, 10-12.
[128] Ibid, 13-20.
[129] Anthea Hucklesby and Azrini Wahidin, *Criminal Justice* (OUP, 2013) 8.
[130] Ibid, 10.
[131] Lucia Zedner, *Criminal Justice* (OUP, 2004) 21.
[132] Jones C and Lally C, *Prison population growth: drivers, implications and policy considerations* (Parliamentary Office of Science and Technology, 2024) 4.

in overcrowded conditions.[133] There is a trend of longer custodial sentences, which have been increasing year-on-year since 2011, except in 2020, possibly due to the Covid pandemic.[134] Prisons generally have poor conditions which have been described as 'insanitary and unacceptable'[135] and there are questions around legitimacy, for example, due to cuts in staffing levels leading to more power sharing between staff and prisoners.[136] The CJS appears retributive and 'prison has become the defining tool of the punishment process.'[137]

Policing faces a loss of legitimacy,[138] there are concerns around misogyny (such as with the killing of Sarah Everard),[139] and 'institutional racism' (illustrated by the police response to the murder of Stephen Lawrence).[140] The Parole Board struggled with a lack of resources to support those with IPP (imprisonment for public protection) sentences.[141] Parole may also suffer from the fixation on retribution with 'signs that the aims and techniques of parole may be evolving once again to resemble something closer to a re-sentencing exercise that draws explicitly upon the punitive logics of fear, blame, vengeance and incapacitation.'[142] The Probation Service has been described as 'struggling'[143] and accommodation needs were only met in 43% of cases.[144] The current system has numerous issues which need addressing. A desirable future CJS will speak to how to remedy the current problems and be careful not to introduce any new, worse problems when proposing new techniques and structures.

3. Addressing Abolitionism

The first big question to ask is whether our desirable future CJS is built from the components of the existing CJS or whether we scrap our current 'decaying dustbins filled with those

[133] *Life in prison: Living conditions* (HM Inspectorate of Prisons, 2017) 6.
[134] *Criminal Justice Statistics quarterly: December 2021* (Ministry of Justice, 2022) 1.
[135] Jacqueline Beard, *The Prison Estate in England and Wales* (House of Commons Library, 2023), 5.
[136] Ben Crewe and Alison Liebling, 'Reconfiguring and reimagining penal power' in Alison Liebling and Shadd Maruna and Lesley McAra, *The Oxford Handbook of Criminology* (OUP, 7th edn, 2023), 846.
[137] *Do Better Do Less: The report of the Commission on English Prisons Today* (Commission on English Prisons Today, 2009), 11.
[138] Robert Reiner, *The Politics of the Police* (OUP, 4th edn, 2010) 94.
[139] Paul Bleakley, '"Would your level of disgust change?" Accounting for variant reactions to fatal violence against women on social media' (2022) 23 Criminology & Criminal Justice 845.
[140] Sir William Macpherson, *The Stephen Lawrence Enquiry* (1999), 41.
[141] Justice Committee, *IPP sentences* (HC 2022-23, 266 – III), paras 90-94.
[142] Harry Annison and Thomas Guiney, 'Populism, Conservatism and the Politics of Parole in England and Wales' (2022) 93 The Political Quarterly 416, 421.
[143] 'Chief Inspector calls for an independent review of the Probation Service, publishing his final annual report' (*HM Inspectorate of Probation*, September 2023) <https://www.justiceinspectorates.gov.uk/hmiprobation/media/press-releases/2023/09/annual-report-2022-2023/> accessed 31 December 2024.
[144] *HM Inspectorate of Probation Annual Report 2022/2023* (HM Inspectorate of Probation, 2023), 25.

pejoratively labelled as 'social trash''[145] and build something from the ground up. 'Defunding prisons… has been central to abolitionist praxis'[146] with the abolitionist stance being that prisons are fundamentally flawed and are places of 'terror and trauma'[147] with poor conditions. A key abolitionist argument is the link to social causes of crime, which leads to 'often-vulnerable, overwhelmingly destitute, and disproportionately black and minority ethnic people'[148] being 'churned' through the CJS. Many abolitionists want 'a justice system based on reparation and reconciliation rather than retribution.'[149] It could be contended that not all people in prison are there due to their vulnerability through structural inequality. Another criticism made of abolitionism in America, a point still relevant in England and Wales, is the question of whether the community could replace the state in fulfilling its responsibilities within a society as 'complex and diverse as ours.'[150] These measures still look like reforms, albeit significant ones, to the existing system rather than a recognisable 'abolition' of the current justice system. The point about reparation and being less retributive is a good one, however we can consider this a reorganising of priorities in the existing system rather than abolition of what we already have. Some of the ideas of how to reform problematic parts of our CJS are useful when considering what a desirable CJS should look like, but the framing of abolitionism is unhelpful.

4. Ideal Purpose of A CJS

Ascertaining the overall purpose of a CJS demands an examination of what the CJS is trying to achieve and what its core principles should be. This in turn will dictate how the 'system' operates at all levels. This is subjective and will vary according to factors such as political stance and cultural values. A good starting point is the normative models of criminal justice proposed by Case et al.[151] The five models proposed are crime control, due process, medical, rights, and victims' interests.[152] These only focus on one aspect of criminal justice at a time,

[145] Joe Sim, 'Confronting State Power: Dissenting voices and the demand for penal abolition' in Alison Liebling, Shadd Maruna, and Lesley McAra, *Oxford Handbook of Criminology* (OUP, 7th edn, 2023), 887.
[146] Ibid, 886.
[147] Ibid, 891.
[148] Ibid, 887.
[149] Angela Y. Davis, *Are Prisons Obsolete* (Turnaround Publisher Services Ltd, 2003,) 107.
[150] 'Emptying the Prisons: Rachel Barkow weighs whether arguments for prison abolition will bring needed change to the criminal justice system—or backfire' (NYU Law, 30 March 2023) <https://www.law.nyu.edu/news/ideas/rachel-barkow-prison-abolition-movement#:~:text=NYU%20LAW%20NEWS,Emptying%20the%20Prisons%3A%20Rachel%20Barkow%20weighs%20whether%20arguments%20for%20prison,residents%E2%80%94in%20prison%20or%20jail.> accessed 18 December 2024.
[151] Steve Case and others, *The Oxford Textbook on Criminology* (OUP, 2nd edn, 2021), 72-73.
[152] ibid.

which depicts the extremes. For example, the crime control model could facilitate abuses of individual rights and the victims' interests model could make our already retributive system[153] even more retributive. The ideal purpose of a desirable future CJS for England and Wales would be a combination of various aspects drawing on a few different models, with the primary focus on the rights based and due process models. This sets a boundary which the state should not cross while also setting principles around limiting arbitrary and unfair state power to avoid the state redefining this boundary.

The ancient principle of equality before the law[154] is ingrained in the English and Welsh legal system and must be a core principle of a desirable CJS. Another core quality is the idea of justice, which has been proposed as the main concern of social institutions.[155] Rawls imagined justice as fairness,[156] which could support fairness as the driving motivation behind the CJS. Currently Section 142, Criminal Justice Act[157] sets out the purposes of sentencing but does not give an order of priority. A desirable CJS would have a clear hierarchy of what its guiding principles were. In this article the main guiding principle is fairness, followed by the other principles discussed.

5. Role of Punishment

Punishment is ordinarily understood to mean hard treatment imposed on an offender by a recognised authority for an offence and always involves censure.[158] Christie writes that 'imposing punishment within the institution of law means the inflicting of pain, intended as pain'.[159] This dynamic is morally problematic[160] on a fundamental level. Had it not been the state administering (or at least allocating)[161] the punishment, this would 'not… normally be considered morally acceptable,'[162] which leads to the conclusion that punishment needs justifying. Durkheim argues that punishment of crime is a communicative good expressing

[153] Ian Loader, 'Playing with Fire? Democracy and the Emotions of Crime and Punishment' in Susanne Karstedt, Ian Loader and Heather Strang (eds), *Emotions, Crime and Justice* (Hart Publishing, 2011), 347.
[154] Magna Carta 1297, cl 40.
[155] John Rawls, *A Theory of Justice: Original Edition* (Harvard University Press, 1971), 3.
[156] ibid.
[157] Criminal Justice Act 2003, s 142.
[158] Rob Canton, *Why Punish? An Introduction to the Philosophy of Punishment* (Bloomsbury Academic, 2017), 3.
[159] Nils Christie, *Limits to Pain* (1981).
[160] Rob Canton, *Why Punish? An Introduction to the Philosophy of Punishment* (Bloomsbury Academic, 2017), 4.
[161] Lucia Zedner, *Criminal Justice* (OUP, 2004) 73.
[162] Rob Canton, *Why Punish? An Introduction to the Philosophy of Punishment* (Bloomsbury Academic, 2017), 8.

our values.[163] Punishment can be seen as essential in a healthy society to maintain this collective consciousness. However, this still leaves flexibility for what punishment should look like.

One justification of punishment is that 'individuals are free willed and therefore offenders can be held morally responsible for their actions.'[164] There is an issue with offenders who do not have free will, such as who are in a state of automatism or are committing crimes out of a need to survive. However, offenders should be punished when it is fair to hold them accountable for their actions. In this way it is clearly not fair to hold the person in a state of automatism accountable as they were not even aware of what they were doing, provided they did not induce such a state. Yet is it morally justifiable to punish somebody for stealing food if they were about to starve? The Supreme Court of Cassation in Italy ruled that this is not a crime.[165] This shows free will alone may not be sufficient justification.

Other justifications promote retribution and deterrence. These punitive approaches can be said to have a 'brutalizing' impact on the public.[166] Deterrence also relies on the assumption of a rational calculation which is not the case as offenders may feel that they will not be caught.[167] Therefore, these punitive approaches will mainly be avoided in a desirable CJS.

There is also a question as to the degree and type of punishment that is desirable. It should respect due process and be linked to fairness and proportionality. A justification of prison as a method of punishment is that it is an 'egalitarian punishment' in that 'liberty is a good that belongs to all in the same way.'[168] However, if used in a retributive way rather than just for public protection, prison cannot be seen as desirable. To conclude, punishment is justifiable and deserves a role in an ideal CJS. However, this does not mean that the retributive current system of mass incarceration is the best way to deliver this.

---

[163] Emile Durkheim, *Durkheim: the Division of Labour in Society* (Steven Lukes ed, Bloomsbury Publishing Plc, 2013), 66.
[164] Lucia Zedner, *Criminal Justice* (OUP, 2004), 86.
[165] Benjamin Soloway, 'Stealing Food if You're in Need Is Not a Crime, Italian Court Finds' (Foreign Policy, 3 May 2016) <https://foreignpolicy.com/2016/05/03/stealing-food-if-youre-in-need-is-not-a-crime-italian-court-finds/> accessed 25 December 2024.
[166] Lucia Zedner, *Criminal Justice* (OUP, 2004), 92.
[167] ibid.
[168] Michel Foucault, *Discipline and punish: the birth of the prison* (Vintage Books, 1977), 232.

6. Scandinavian Exceptionalism

Scandinavian exceptionalism is defined by Pratt as encompassing both the low levels of imprisonment in Scandinavian nations as well as the prison conditions which 'approximate to life outside.'[169] Pratt cites 'highly egalitarian cultural values and social structures of these societies'[170] as the root cause of this exceptionalism. The significance of this is that inmates are 'treated with respect for their human dignity.'[171] Building a desirable CJS requires being fair not only to victims, but also to offenders. The Scandinavian system has been able to maintain a 'strong focus on rehabilitation,'[172] which is likely aided by prisoners feeling their human dignity is being respected. When authority is exercised 'relationally', which can be understood as meaning as part of a reciprocal relationship, this 'generates commitment and comparatively high levels of legitimacy.'[173] This logically makes it easier to engage prisoners in the rehabilitation process.[174] However, some argue that these exceptional conditions resulted from prisoner resistance to welfare sanctions, as opposed to the sanction itself granting such conditions.[175] Consequently, the 'rehabilitative ideal' was unsustainable in some regions.[176] This flaw of the Scandinavian system can be explained by the fact that the rehabilitative model adopted treated prisoners as 'psychologically deficient subjects.'[177] This distinction helps us avoid this style of rehabilitation, reminiscent of the medical model,[178] but it does not preclude all rehabilitative efforts.

Scandinavian exceptionalism has come under criticism for the social costs it imposes, namely the pressures to conform to wider society[179] and a powerful state.[180] Whilst this gives scope to protect and look after the population, it also empowers human rights abuses such as Sweden's sterilization of the 'mentally retarded.'[181] Prison conditions that, on the face of

[169] John Pratt, 'Scandinavian Exceptionalism in an Era of Penal Excess: Part I: The Nature and Roots of Scandinavian Exceptionalism' (2008) 48 British Journal of Criminology 119, 119.
[170] Ibid, 120.
[171] Correctional Treatment in Institutions Act 1974 (Swedish Code of Statute 1974:203), 9.
[172] Peter Scharff Smith, 'A critical look at Scandinavian exceptionalism' in Thomas Ugelvik and Jane Dullum (eds), *Penal Exceptionalism?* (Routledge, 2012), 73.
[173] Ben Crewe and Alison Liebling, 'Reconfiguring and reimagining penal power' in Alison Liebling, Shadd Maruna, and Lesley McAra, *Oxford Handbook of Criminology* (OUP, 7th edn, 2023), 837.
[174] Jonathan Jackson and others, 'Legitimacy and procedural justice in prisons' (2010) *191 Prison Service Journal 4,* 8.
[175] John Pratt, 'Scandinavian Exceptionalism in an Era of Penal Excess: Part I: The Nature and Roots of Scandinavian Exceptionalism' (2008) 48 British Journal of Criminology 119, 131.
[176] ibid.
[177] ibid.
[178] Steve Case and others, *The Oxford Textbook on Criminology* (OUP, 2nd edn, 2021), 72-73.
[179] John Pratt and Anna Eriksson, 'In defence of Scandinavian exceptionalism' in Thomas Ugelvik and Jane Dullum (eds), *Penal Exceptionalism?* (Routledge, 2012), 373-374.
[180] Ibid, 374-376.
[181] Ibid, 375.

it, can look like a hotel, also carry different pains. One example is 'the feeling of being nearly free but not free yet' adding to prisoners' frustrations.[182] This means that the experience for the prisoner may not actually be affected by prison conditions as much as it might be expected. It is also worth pointing out that the social conditions are too different between England and Scandinavia for an emulation of the Scandinavian model to be effective, with England being more diverse than Scandinavia's 'homogenous societies.'[183] Overall, Scandinavian exceptionalism can be used as a useful demonstration of the strength of adopting egalitarian values, but is not a model to be copied exactly as shown by the flaws discussed.

7. Social Justice-Based Arguments

In this article, 'social justice' led reform refers to efforts to address the social and systemic causes of crime, rather than fixating on punishment.[184] With this in mind, McNeill proposes a multifaceted approach to rehabilitation, encompassing four distinct elements of integration.[185] Firstly, there is legal rehabilitation which involves the legal system supporting an offender to live a normal life.[186] For example, criminal records could be sealed to allow an offender to participate in the labour market, as engaging in the labour market is correlated with desistance from crime.[187] Moral rehabilitation is to meet moral demands and secure moral communication. McNeill feels reparation is best placed to do this.[188] Social rehabilitation is to do with the 'informal social recognition and acceptance'[189] of a reformed offender. Finally, psychological rehabilitation is most aligned to the classic understanding of rehabilitation where there is individual change in the offender developing their 'motivation, skills and capacities.'[190] McNeill also points to the idea of positive and negative forms of penal power and encourages the positive form,[191] distinguished between by Garland.[192] Retribution is a 'negative' form of penal power because it takes something from the offender.

[182] Ibid, 350.
[183] Ibid, 373.
[184] Rob Canton, *Why Punish? An Introduction to the Philosophy of Punishment* (Bloomsbury Academic, 2017), 156.
[185] Fergus McNeill, 'What Good is Punishment?' in Farrall and others (eds), *Justice and Penal Reform* (Taylor & Francis Group, 2016), 74-75.
[186] Ibid, 76.
[187] Paul McGuinness, Fergus McNeill and Sarah Armstrong, *The use and impact of the Rehabilitation of Offenders Act (1974): Final Report* (The Scottish Centre for Crime & Justice Research, 2013), 16.
[188] Fergus (n 62), 76.
[189] Ibid, 77.
[190] Ibid, 75.
[191] Ibid, 72-74.
[192] David Garland, 'Penality and the Penal State' (2013) 51 Criminology 475, 501-505.

Philosophies of punishment such as rehabilitation may be regarded as 'positive' because they aim at 'grafting the offender back into the social body rather than severing him or her from it.'[193]

This rehabilitation-based positive penal power approach focuses on reintegration in society as a metric of success.[194] This is effective because crime levels are arguably more to do with social factors.[195] If this is true, it would follow that it is not possible to 'reform' people by only focusing on the offender but not the underlying social conditions. The Scandinavian example also supports the idea that we should not be treating prisoners as though they have a mental disorder, but to support people to integrate into society. Ex-offenders who are integrated into society are better supported, which would logically lead to a reduction in crime rates.[196] This would alleviate some of our current issues, such as the overcrowding in prisons.

It could be argued that social justice-based rehabilitation could jeopardise public safety by mixing more offenders with communities. However, public protection can in fact be enhanced through rehabilitation, as it reduces recidivism. Some people may never be able to be rehabilitated, which requires a continuation of the current CJS and prison system alongside this social justice based approach. This is not advocating for an abolitionist perspective of social justice such as the one put forward by Scott and Codd,[197] and not everyone even commits crimes due to social conditions. Rather, it is aiming to rehabilitate as many people as possible while still protecting the public where necessary. Positive forms of penal power need the cooperation of everyone in society, which is challenging, but have the potential to dramatically improve the outcomes of the English CJS.

8. Importance of Procedural Justice

Procedural justice speaks to how 'the process by which decisions are made is of considerable importance; to the extent that perceptions of procedural fairness are more critical than outcome fairness, as drivers of perceived legitimacy.'[198] An approach that prioritises

[193] Fergus (n 62), 73.
[194] Ibid, 78.
[195] Michael Cavadino, *The Penal System: An Introduction* (Sage, 6th edn, 2020).
[196] United Nations, *Introductory Handbook on the Prevention of Recidivism and the Social Reintegration of Offenders* (United Nations Office on Drugs and Crime, 2018), 7.
[197] David Scott and Helen Codd, *Controversial Issues in Prisons* (McGraw-Hill Education (UK), 2010).
[198] Harry Annison, Nicola Carr and Thomas Guiney, *The Future(s) of Parole: Ideas, Institutions and Practices* (unpublished, 2025), 18.

procedural justice can leverage this sentiment to create a CJS that feels fairer. This means people are more likely to recognise the legitimacy of the system, likely resulting in greater compliance.[199] There is also evidence to suggest that a lack of perception of fairness may cause crime.[200] Such a critical function makes strong procedural justice an essential principle in a desirable future CJS. What this may look like in practice is avoiding unnecessarily harsh treatment, as Scandinavia demonstrated in its hotel-like prison conditions. The due process model[201] sets out an extreme version of what a system defined by checks and balances might look like. Although due process may be less effective where the processes themselves are changed, this can be mitigated by having strong principles of democracy and fairness.

9. Politics of Criminal Justice

A desirable CJS would not just be founded on robust principles but also be practically possible, including political practicality. Loader acknowledges that criminal justice has become more 'populist in style and punitive in substance.'[202] This would be a major block to implementing a more rehabilitation-based approach, because a political party implementing this may risk being accused of being soft on crime, which could hurt their chances of re-election. One approach to combatting this is the cognitive deficit model. This involves better educating the public as to how the CJS works,[203] as actual understanding is generally poor.[204] A drawback is the lack of focus on the role emotions play. However, emotions and cognitive processes are to some extent intertwined with each other.[205] This may not eliminate a punitive mindset but could reduce the opposition to rehabilitation-based approaches. Another approach is the insulation model. This proposes that politics should be separate from criminal justice.[206] It tries to avoid the politicisation of the CJS but may be counter-productive,[207] as removing crime control from public debate can weaken public trust and lead to exploitation by authoritarian leaders. These risks show that more open

[199] Kristina Murphy and Tom Tyler, 'Procedural justice and compliance behaviour: the mediating role of emotions' (2008) 38 European Journal of Social Psychology 652.
[200] Ben Bradford, Matteo Tiratelli and Paul Quinton, 'Does stop and search reduce crime?' (2019) 58 The British Journal of Criminology 1212.
[201] Steve Case and others, *The Oxford Textbook on Criminology* (OUP, 2nd edn, 2021), 72-73.
[202] Ian Loader, 'Playing with Fire? Democracy and the Emotions of Crime and Punishment' in Susanne Karstedt, Ian Loader and Heather Strang (eds), *Emotions, Crime and Justice* (Hart Publishing, 2011), 347.
[203] Ibid, 349.
[204] Nicola Marsh and others, *Public Knowledge of and Confidence in the Criminal Justice System and Sentencing* (Sentencing Council, 2019).
[205] Elizabeth A Phelps, 'Emotion and cognition: insights from studies of the human amygdala' (2006) 57 Annu Rev Psychol 27.
[206] Ian Loader, 'Playing with Fire? Democracy and the Emotions of Crime and Punishment' in Susanne Karstedt, Ian Loader and Heather Strang (eds), *Emotions, Crime and Justice* (Hart Publishing, 2011), 352.
[207] Ibid, 354-355.

conversations are needed but that insulationism could still be used where public debate is more likely to harm rather than help, such as with sentencing in contentious current cases.

A mix of both models, with greater emphasis on the cognitive deficit model, is the most useful for an ideal CJS. While politics can and should never be completely separate from criminal justice in England and Wales, separating them where possible reduces the risk of populism. For areas that cannot be separated, addressing the public's cognitive deficit should enable decisions to be based in logic, as people would be able to better understand the reasoning behind the decision. Another suggestion I would make is changing the metrics by which the success of the CJS is measured. McNeill proposes reintegration of offenders into society as a 'positive and challenging metric' by which to measure success against.[208] This would serve as a useful metric for the more rehabilitation-focused approach advocated within this article, given its focus on offender outcomes.

10. Victims of Crime

Victims also deserve to be considered by the system. Better understanding of who the victims are will enable a desirable CJS to put in place measures to support these groups of people. It could be argued that offenders can also be victims, due to the state creating 'policies that promote various forms of inequalities that are known to perpetuate criminogenic conditions.'[209] This might explain why 'punishment tends to disproportionately target minority groups'[210] and people who are 'the most marginalised and vulnerable in society.'[211] These are least able to advocate for themselves and resist negative influences. Truly tackling the root cause of these social conditions is arguably beyond the scope of the CJS. A desirable CJS can still support these especially vulnerable individuals by adopting the more social justice-based approach to rehabilitation which has been discussed previously in this article. This also has the benefit of supporting the families who are deeply affected by having a loved one imprisoned.[212]

It is worth briefly acknowledging feminist perspectives on the current conditions of the English & Welsh criminal justice system. There is a feeling that the CJS usually operates

[208] Fergus (n 62), 78.
[209] Marie Manikis, 'Recognising State Blame in Sentencing: A Communicative and Relational Framework' (2022) 81 The Cambridge Law Journal 294, 308.
[210] Henrique Carvalho and Anastasia Chamberlen, *Questioning Punishment* (Taylor & Francis 2023) 49-50.
[211] Ibid, 50.
[212] Cara Jardine, 'Families, imprisonment and penal power: a radical analysis' (2023) 6 Justice, Power and Resistance 278.

within a context of the 'male, white subjectivity of law.'[213] Hudson attempts to tackle this challenge by advocating a discursive approach to justice.[214] The idea is that everyone's voice should be heard. This links to the cognitive deficit model by addressing a lack of different perspectives being heard but is more about addressing the lack of knowledge of decision makers within the CJS. More discussion likely translates to greater representation which contributes towards building a fairer, more desirable CJS. The ideas discussed may seem to be focusing on the offender more than the victims of the crime themselves. While both interests are important, supporting offenders effectively should, in the long term, result in fewer victims.

In the current system victims are represented by the state and can be said to be 'pushed completely out of the arena.'[215] Christie argues that conflicts can be seen as property.[216] The victim becomes a 'double loser'[217] in that the victim loses control of what is 'one of the more important ritual encounters in life.'[218] A desirable CJS would involve victims more, but would balance other interests as well to avoid the flaw of victims wanting retribution or being overly lenient in a pure victim interests' model.

11. Considering Specific Plans for Reform

It is useful to ground the theoretical principles of a desirable justice system by considering a couple of tangible examples. An enhanced system has been proposed for prolific offenders.[219] This would ensure that those who most need support to reform receive it. A disproportionate benefit could be achieved as 'almost a tenth of offenders drive up to half of total crime,'[220] Another suggestion was a national hub for victim care.[221] A centralised system could provide all case updates and advice in one place. The benefit is making victims feel more ownership of their case, which goes back to the idea of conflicts being the property of the victim.[222] Punishment in the community could be achieved through an intensive

[213] Barbara Hudson, 'Beyond white man's justice: Race, gender and justice in late modernity' (2006) 10 Theoretical Criminology 29, 30.
[214] Ibid, 34-35.
[215] Nils Christie, 'Conflicts as Property' (1977) 17 The British Journal of Criminology 1, 3.
[216] Ibid, 1.
[217] Ibid, 3.
[218] Ibid, 3.
[219] Harvey Redgrave and Madeline Rolfe, *A Plan to Reform the Criminal-Justice System* (Tony Blair Institute for Global Change 2023) 31.
[220] Ibid.
[221] Ibid, 34.
[222] Nils Christie, 'Conflicts as Property' (1977) 17 The British Journal of Criminology 1, 1.

community work order.[223] This would encourage judges to use alternatives to imprisonment more, which could be a starting point to moving away from mass incarceration. However, it would need to be examined whether this type of work order helps an offender's reintegration or increases their frustration with the feeling of being almost free. It cannot be at the expense of public safety. This would require a guarantee of high-quality monitoring, which may be difficult with the current stretched resources.

Justice reinvestment is an interesting approach where local authorities use funds that would have been spent on incarceration to invest in the most impacted communities.[224] This contributes to addressing the root social causes of offending. However, there are also suggestions that local authorities use it to fund social services[225] such as drug rehabilitation. This loses the community focus by 'doing 'to' communities rather than 'through' communities.'[226] The preceding example shows the extent to which social justice relies on community participation.

Conclusion

To conclude, this article has considered the theoretical, and a few practical, dimensions of a desirable future English CJS. Core principles such as fairness, due process, and a rights-based approach have been advocated for. Drawing inspiration from the Scandinavian model, the importance of respecting human dignity and fostering rehabilitation have been emphasized throughout. The article has been constructed on a solid bedrock of an understanding of a political backdrop, abolitionist perspectives and the role that punishment should play. This all points towards moving beyond the current overreliance on retributive punishment in favour of social justice based arguments focusing on reintegration and addressing the root causes of crime, recognizing offenders' potential victimhood within social inequalities. Public education and a shift towards measuring success by offender reintegration were proposed methods of mitigating populist pressures. This article was not intended to definitively rank the factors identified as important in their order of priority but does lean towards procedural fairness and social justice-based approaches. Finally, it is

[223] Harvey Redgrave and Madeline Rolfe (n 96), 35.
[224] Susan B Tucker and Eric Cadora, *Ideas for an Open Society: Justice Reinvestment* (Open Society Institute, 2003), 5.
[225] Todd R Clear, 'A private-sector, incentives-based model for justice reinvestment' (2011) 10 Criminology & Public Policy 585, 599.
[226] Ruth Armstrong and Shadd Maruna, 'Examining Imprisonment Through a Social Justice Lens' in Farrall and others (eds), *Justice and Penal Reform* (Taylor & Francis Group 2016) 147.

critical to keep the CJS under ongoing review because the conception of a desirable CJS may change over time as societal values evolve.

Less is More? A Critical Appraisal into Minimum Intervention and Maximum Diversion in The Youth Justice System

Erin Deane

Abstract

The youth justice system is a part of the criminal justice system, specifically dedicated to children aged between 10 and 17 years of age. It incorporates various organisations, government institutions, and agencies, including youth custody, youth courts, and the police. It is a system in England & Wales which aims to rehabilitate, prevent offending, and achieve justice. Despite the importance of these aims, the youth justice systems' current approach towards youth offenders and crime is up for debate. Consequently, this paper offers a critical evaluation into whether an alternative approach to youth offending would be more appropriate than the current approach. Particularly, through the consideration of whether a minimum intervention and maximum diversion approach would be better suited. The paper will outline what a minimum intervention and maximum diversion approach could look like, with consideration for the current age of criminal responsibility, labelling theory, and the age crime curve. The article will primarily concern itself with England & Wales, however, there will also be acknowledgment for certain approaches taken in Europe.

Introduction

The youth justice system (hereinafter 'YJS') aims to primarily provide 'prevention, diversion, and a clear commitment to developing interventions to avoid unnecessary prosecutions and criminalisation of children.'[227] It is a system specifically targeting children and 'involves many agencies working together including the police, courts and prison service.'[228] Despite this, questions arise regarding the effectiveness of the YJS and specifically, whether it fails to address the competing interests between these identified objectives. A prominent perspective is that the YJS should reduce the risk of offending through punitive strategies to achieve public protection. However, as will be explored, this

[227] Crown Prosecution Service, Legal Guidance: Children as suspects and defendants [2023] < https://www.cps.gov.uk/legal-guidance/children-suspects-and-defendants > accessed 14th April 2025.
[228] Youth Justice Legal Centre, Youth Justice System [2021] < https://yjlc.uk/resources/legal-terms-z/youth-justice-system#fn_1 > accessed 14th April 2025.

perspective is contested amongst growing literature which acknowledges childhood vulnerability, innocence, and institutional flaws.

Subsequently, this paper will provide a critical evaluation into whether the YJS would be more successful if it were to take a minimum intervention and maximum diversion approach to offending, as opposed to a maximum intervention and custodial approach. The paper will first explore diversion and its development within the youth justice system. It will then explore potential strategies used to implement diversion, including consideration for the age of criminal responsibility and youth cautions. This will be supported through identification of the current issues within the system including labelling,[229] the age crime curve, and public expenditure. The paper will acknowledge the implications of these criticisms and suggest how the system could reconcile these issues if it were to implement minimum intervention and maximum diversion. Alongside this, the paper will recognise the barriers to implementation to include individual positivism, politics, and time. Nevertheless, the paper will conclude that despite some areas of concern, there should be implementation of minimal intervention and maximum diversion within the YJS. Subsequently, it would provide a more balanced approach to youth offending, in comparison to the current approach taken.

1. Diversion

The act of diversion refers to an informal mechanism of control. It is the act of diverting youth offenders away from the formalised criminal justice interventions and towards social, state, and community interventions.[230] This paper will focus on the act of diverting youth offenders away from custody; however, it will also include elements of diversion from the criminal justice system completely. The welfare approach reflects diversionary principles by prioritising the 'welfare of the child'[231] and acting in their 'best interests.'[232] It requires removing children from 'undesirable surroundings,' 'proper provision [are] made for his education and training,' and acting in a way which is best suited to the child in those circumstances.[233] Subsequently, these principles advocate diversion as an approach to youth

[229] Howard Becker, 'Outsiders: Studies in the Sociology of Deviance' (The Free Press 1963) 9.
[230] Youth Justice Legal Centre, Diversion [2021] < https://yjlc.uk/resources/legal-terms-z/diversion > accessed 14th April 2025.
[231] Child and Young Person's Act 1933, s 44.
[232] United Nations Convention on the Rights of the Child (adopted 20 November 1989, entered into force 2 September 1990), 1577 UNTS 3, art 3.
[233] Child and Young Person's Act 1933, s 44.

offending as they support rehabilitative interventions being utilised over custodial sentences thus, encouraging greater community interventions.

The developmental model is a broad criminogenic framework which further endorses a diversionary approach to youth offending as it suggests offending is a phase in adolescent life.[234] The implication of such perspective is that it suggests offending is a developmental stage which will be overcome naturally. This is significant as, following this theory, offending will naturally resolve itself overtime and therefore, no intervention is necessary. The focus on diversion within these models justifies and encourages minimum intervention, especially through demonstrating greater success of rehabilitation and societal re-integration through less invasive strategies. However, whilst these perspectives are supportive in theory, this paper argues they have limited application in reality. The current approach to youth offending places a greater focus on retribution over rehabilitation thus, highlighting their limitations. The difference between ideals and reality is exemplified through the development of the YJS overtime.

In the 1980's there was greater utilisation of a law-and-order approach.[235] This era saw Margaret Thatcher become the Prime Minister and introduce a 'tough on crime' and 'short, sharp, shock' approach to youth detention centres.[236] Such approach drove punitive aspects of punishment, custodial sentences, and reduced the implementation of rehabilitative ideals.[237] Despite this, there was also growing recognition that minimal intervention could reduce public expenditure. Subsequently, this provided positive endorsement for the use of cautioning and other informal disposals. It demonstrated an emerging alliance between politics, public expenditure, and crime control because the political climate influenced the Government's response to offending (penal popularism).[238] For example, the Government's diversionary response was largely driven by the political motivation to reduce tax, according to societal attitudes.[239] This point in history is significant because it exhibits recognition of the advantages a minimal intervention could achieve.

---

[234] Terrie Moffitt, 'Adolescence-limited and life-course-persistent antisocial behaviour: A developmental taxonomy' [1993] *100*(4) *Psychological Review* 675.
[235] Stephen P Savage and Lynton Robins, *Public Policies under Thatcher* (1st edn, Palgrave, 1990) ch 6.
[236] Ian Cummins, Thatcherism and its Legacy (1st edn, OUP, 2021) ch 1.
[237] Alan Norrie and Sammy Adelman, 'Consensual Authoritarianism and Criminal Justice in Thatcher's Britain' [1989] 16(1) Journal of Law and Society 112.
[238] Anthony Bottoms, 'The Philosophy and Politics of Punishment and Sentencing' in Clarkson & Morgan, *Politics of sentencing reform* (OUP, 1995), 17 – 49.
[239] Alan Norrie and Sammy Adelman (n 11).

Nonetheless, a key problem with this approach is that it did not offer a complete resolution to the 'penal crisis' because prison populations continued to rise.[240] Thus whilst the Thatcher era was characterised through greater diversionary techniques,[241] it was vigorously challenged by society who deemed it to be arbitrary, unacceptable, and wrong following Bulger.[242] The Labour Party echoed this public fear by rejecting the previous Conservative mantra – "we ran for office as New Labour, we will govern as New Labour."[243] Labour, instead, focused on the evidential basis of crime control and wanted to address 'what works' in punishment to reduce reoffending.[244] They implemented this through a balanced approach whereby punishment could serve both punitive and rehabilitative aims in order to be 'tough on crime, tough on the causes of crime.'[245] There was greater emphasis on rehabilitative interventions and it was operationalised through their introduction of targets, policing strategies, and alternative offending approaches. For example, the introduction of drug treatment and testing orders (DTTO).[246] Their approach emerged from the societal perception of diversion being a 'metaphorical slap on the wrist with no further consequences.'[247] Accordingly, it was distinguished from Thatcher's response.

The development of the YJS, therefore, demonstrates the large influence politics and societal attitudes has on responses to youth offending. It is important that this is recognised because the implication of such influence would likely determine the success of certain responses to youth offending, as seen with the Blair Government's 'tough on crime, tough on the causes of crime' approach.[248] In order for a minimal intervention and maximum diversion to be successful, the political climate needs to be acknowledged because it could act as a potential barrier to implementation. Consequently, without support from the public and political majority, such approach is likely to be unsuccessful. Nevertheless, the evolution of the YJS demonstrates diversion to be a historical concept which, in the past, has been successful. This indicates potential for a minimum intervention and maximum diversion approach to be

[240] Michael Cavadino, James Dignan, George Mair, & Jamie Bennett, *The Penal System* (6th edn, SAGE, 2020) ch 10.
[241] Alan Norrie and Sammy Adelman (n 11).
[242] *R v Secretary of State for the Home Department, Ex parte V. and R v Secretary of State for the Home Department, Ex parte T* [1997] 2 WLR 67.
[243] Tony Blair, 'General Election Victory Speech' (London, 5th May 1997) < http://www.britishpoliticalspeech.org/speech-archive.htm?speech=222> accessed 17th November 2024.
[244] Labour Party Manifesto 1997.
[245] Susan Ratcliffe, *Oxford Essential Quotations* (6th edn, OUP, 2018).
[246] Crime and Disorder Act 1998, s 61.
[247] Jon Robins, 'Youth Cautions and the Slap on the Wrist' [2012] The Justice Gap < https://www.thejusticegap.com/youth-cautions-and-the-slap-on-the-wrist/ > accessed 8th may 2025.
[248] Susan Ratcliffe (n 19).

implemented in modern society. This is providing there is greater consideration for public and political influence on its operationalisation and effectiveness. To be successful, it would be required to have a greater outline of what diversionary strategies could or would be used. It would also be required to acknowledge the tension which exists between responses to offending and public expenditure.

To summarise, society currently favour a retributive approach to punishment. A diversionary approach away from the formalised custodial system supports rehabilitative, community, and less invasive interventions being prioritised. As historically demonstrated, diversion has potential to be successfully implemented, however, whether it could successfully be achieved in todays, modern society is dependent on whether it would be supported amongst society and the political climate. It is for this reason, the paper will now go on to consider potential amendments, approaches, and strategies which could introduced in order to implement a minimum intervention and maximum diversion approach, and whether these would be positively received.

2. Implementing Minimum Intervention and Maximum Diversion

Following this account, a minimum intervention and maximum diversion approach could be implemented into the criminal justice system. However, how such approach could be achieved remains unseen. Accordingly, this paper will now explore different factors which could contribute to the implementation of a minimum intervention and maximum diversion approach. Notably, the influence the age of criminal responsibility and youth cautions could have.

2.1. Age of Criminal Responsibility

The age of criminal responsibility refers to the age in which an individual can be arrested and charged with a crime.[249] The age of criminal responsibility in the United Kingdom ('UK' hereinafter) is 10 years-old,[250] however, it differs across countries. For example, in Brazil it

[249] Houses of Parliament, Age of Criminal Responsibility, [2018] *Houses of Parliament Parliamentary Office of Science & Technology*, POSTNOTE 577.
[250] Children and Young Persons Act 1933, s 50.

is 12 years-old,[251] in Argentina it is 16 years-old,[252] and in Germany it is 14 years-old.[253] A key problem with the age of criminal responsibility in the UK is that empirical evidence exemplifies the damaging and long-lasting effects the YJS has on children. Children are physically, biologically, and emotionally different to adults.[254] They are shown to find formalised intervention difficult to comprehend, intense, and intimidating as formalised interventions facilitate adult capacities, which are greater than children's.[255] The current age of criminal responsibility fails to address these concerns by allowing young children aged between 10 and 17 years of age be subjected to the criminal justice process.

International legal framework echoes this concern, and casts further doubt on the appropriateness of the UK's current age of criminal responsibility. This is a result of the 'clearly developing consensus in Europe and internationally that the age of criminal responsibility should be set at 12 years at a minimum.'[256] Consequently, the UK has one of the lowest ages amongst Europe. Alongside the 'trend for countries around the world to raise their ages'[257] it is found that UK children are the most disadvantaged, in comparison to 21 other European organisations.[258] These statistics further challenge the UK's age of criminal responsibility and suggest that the current approach is unsatisfactory. The UK approach is failing to address these scientific and international concerns, despite presumptions that a higher criminal age would enable a more advantageous approach and keep children out of the system.[259] In order to align the UK with international standards, the criminal age of responsibility should be raised to 12 years of age at a minimum. The implications of raising the criminal age of responsibility would not only align greater with international standards, but it would also successfully implement a minimal intervention approach. This is because through raising the age of individuals who are capable of being arrested and charged with a crime there is a narrowing of the scope of children capable of being brought into the formal

[251] UNICEF, Legal Minimum Ages and The Realization of Adolescent's Rights <https://www.unicef.org/lac/media/2771/file/PDF%20Minimum%20age%20for%20criminal%20responsibility.pdf > accessed 8th May 2025.
[252] Ibid.
[253] Child Rights International Network, Minimum Ages of Criminal Responsibility in Europe < https://archive.crin.org/en/home/ages/europe.html > accessed 8th May 2025.
[254] Mark Telford, 'The Criminal Responsibility of Children and Young People: An Analysis of Compliance with International Human Rights Obligations in England and Wales' [2012] 5(2) IJPL 107 – 120.
[255] Children's Commissioner, Children's experience as victims of crime [2024].
[256] Thomas Crofts, 'Catching up with Europe: Taking the Age of Criminal Responsibility Seriously in England' [2009] 17 *European Journal of Crime, Criminal Law and Criminal Justice* 284.
[257] Neal Hazel, 'Cross-National Comparison of Youth Justice' (Youth Justice Board 2008) 32.
[258] UNICEF, 'The Well-Being of Children: How Does the UK Score?' (Report Card No 11, 2013) 3.
[259] Thomas Crofts (n 30), 275.

system.[260] This further demonstrates the positive associations with a minimal intervention and maximum diversion approach for youth offending.

Furthermore, such reform is supported by the research to date which suggests that encouraging rehabilitation and mitigating the damaging effects of formalised judicial intervention to be the most reasonable solution to youth offending.[261] Thus, further reiterating the need for a minimal intervention and maximum diversion approach. These benefits are internationally recognised because diversion enables the 'police to concentrate on more serious crime.'[262] For example, Sweden diverts 40% of youth offenders whilst Scotland 41% and the Netherlands prioritises diversion towards voluntary work.[263] These benefits justify and encourage minimal intervention and maximum diversion to be implemented through raising the age of criminal responsibility. A reasonable assumption would be that through the UK raising the age of criminal responsibility it would likely emulate the similar successes seen internationally. It would also likely encourage compliance through informal discretion by parents, teachers, youth workers, or police officers as seen in other jurisdictions.

2.2. Youth Cautions

Alongside the age of criminal responsibility, a minimal intervention approach could also be achieved through greater emphasis on youth cautions. Youth cautions are an out of court disposal (hereinafter 'OOCD') which divert offenders away from custodial sentences and criminal charges.[264] An approach which has greater utilisation of OOCD's would align with the recognition that children should not be criminalised when it could be dealt with through alternative measures,[265] whilst simultaneously tackling the presumption that diversion is 'a 'slap on the wrist.'[266] This is because, unlike other alternatives to custody, youth cautions are recorded as Code 2 on the Police Recorded Crime and Data Outcomes but considered immediately spent.[267] The implications of a youth caution therefore suggest that the latter

[260] UK Parliament House of Commons Justice Committee, Annex A: Government Response to the Justice Select Committee's Twelfth Report of Session 2019-20: Children and Young People in Custody (Part I) [2020] *House of Commons 306.*
[261] Mark Telford (n 28).
[262] Rob Allen, Less is more – the case for dealing with offences out of the court [2017] Transform Justice 2.
[263] Rodney Morgan, 'Summary Justice: Fast-But Fair?' [2008] *Centre for Crime and Justice Studies* 18.
[264] Legal Aid, Sentencing, and Punishment of Offenders Act 2015, s 135.
[265] National Police Chiefs Council, 'Charging and Out of Court Disposals: A National Strategy' [2017] NPCC Portfolio Lead for Charging and Out of Court Disposals.
[266] Jon Robins (n 21).
[267] Rehabilitation of Offenders Act 1974, s 1.

perception of being a slap on the wrist is too simplistic and inaccurate. This is because there are consequences associated with a youth caution as despite being spent immediately, they appear on criminal records or the sex offender register.[268] Subsequently, youth cautions represent a punishment being imposed and this is significant because it is likely to deter children away from offending due to the associations and stigmatisation of a criminal record.

Henceforth, youth cautions enable a punishment to be imposed whilst mitigating the damaging effects associated with formalised forms of intervention, such as a custodial sentence. This is because, for example, youth cautions would mitigate the labelling theory as these convictions are not always declared.[269] This is important because youth cautions represent a method which could be utilised in order to implement a successful minimal intervention and maximum diversion approach which simultaneously achieves a deterrent effect, reduced recidivism, and demonstrates the 'soft law' perception to be unsatisfactory. The 'soft law' interpretation of diversionary attempts likely derives from misconceptions dependent on how a question is being asked,[270] and not from the developed of practises utilised by the YJS. Therefore, the introduction of youth cautions could be another approach which successfully implements minimal intervention and maximum diversion from custodial sentences in youth offending. It would enable greater risk management, a deterrent effect, and resolve 'soft law' perceptions.

Nevertheless, despite potential for success, in order to implement such reform, there needs to be greater acknowledgement of potential due process concerns. The formal criminal justice process ensures due process through rights protection whereas, diversionary approaches do not always ensure legal advice, burden of proof, and the capacity to appeal is available.[271] This is despite such measures being essential to achieve fairness. These concerns are problematic because they indicate there is potential for discrimination within a minimal intervention and maximum diversion approach due to the lack of standardised principles and procedures. An implication of these concerns is that it reduced the potential effectiveness such approach could have. For instance, the disparities are apparent with the current implementation of youth cautions as there is a disproportionate number of suspects

[268] Sexual Offences Act 2003, s 80.
[269] Becker (n 3).
[270] Julian Roberts and Mike Hough, 'Sentencing Young Offenders: Public Opinion in England and Wales' [2004] 5 (3) British Society of Criminology, 211 – 228.
[271] Children and Young People's Centre for Justice, Children and Young People In Conflict With The Law: Policy, Practise and Legislation [2024], s 11.

from Black, Asian, and Ethnic Minority Groups (hereinafter 'BAME').[272] The due process concern presents a potential barrier to implementation of a minimal intervention and maximum diversion approach; however, it could be reconciled where there is greater acknowledgement of this risk and more standardised principles introduced into diversionary methods. For instance, a more guided implementation of youth cautions could resolve this racial bias and improve the effectiveness of minimal intervention and maximum diversion.

Thus, to summarise, the age of criminal responsibility offers one example of how a minimum intervention and maximum diversion approach could be achieved, to a slight extent. It would require the age of criminal responsibility to be raised in order to reduce the scope of children able to be brought into the formal criminal justice system. As asserted, it is supported amongst international legal framework and would encourage greater utilisation of informal interventions, such as voluntary work. Moreover, youth cautions provide a second example of how a minimum intervention and maximum diversion approach could be achieved. Greater utilisation of youth cautions would successfully implement a minimum intervention and maximum diversion approach, without the labelling and negative effects associated with the current approach, as it encourages youth crime to be approached through alternative measures. Providing the due process concerns are addressed, youth cautions have the capability to implement a minimum intervention and maximum diversion approach to a large extent. Following this, the paper will now critically evaluate the barriers to implementation. It will offer a critical appraisal into the current criminal justice system issues and why these could hinder the success of a minimum intervention and maximum diversion approach being implemented. The paper will particularly focus on the labelling theory,[273] the age crime curve, statutory guidelines concerning custodial sentences, the current resource crisis evident within custodial institutions, and finally the current sentencing approach.

3. Issues with The Current System

Despite potential for a minimum intervention and maximum diversion approach to be successfully implemented within the current criminal justice system, the current issues with the system need to be addressed. The current problems with the criminal justice system present barriers to implementation and, consequently, have the potential to hinder a successful minimum intervention and maximum diversion approach being achieved. It is for

[272] David Lammy, An independent review into the treatment of, and outcomes for, Black, Asian, and Minority Ethnic individuals in the Criminal Justice System [2017] *The Lammy Review*, ch 1.
[273] Becker (n 3).

this reason the paper will now critically evaluate the key issues, and therefore barriers, within the criminal justice system.

3.1. Labelling

The labelling theory suggests the 'offender' or 'deviant' label on a child provokes the self-fulfilling prophecy,[274] which encourages compliance with these deviant behavioural patterns.[275] If you label a child to be a deviant, a criminal, or an offender they are more likely to live up to this 'label' and be deviant or commit a crime. The stereotypical image of 'troubled youths' carries stigma, which places children in the spotlight and further criminalises them.[276] The current punitive strategies fail to address the negative impacts labelling children has,[277] especially through custody and policing.[278] An issue with the current response is that it is unintentionally encouraging youth offending through labelling and unsuccessfully achieving the YJS aim to reduce offending.

The labelling theory is further exemplified through the media's negative portrayal of children as threatening, violent, and problematic for social order.[279] The demonisation of youths is achieved through the primary focus on rare cases, such as Bulger,[280] which are highly publicised and accelerate inaccurate representations of children. Bulger[281] illustrated children to be labelled negatively[282] and 'demonic,' despite the rarity of child-on-child murder. It provoked punitive responses to youth offending to emerge in the 1980's. The media's representation of youths links back to public opinion driving government responses to offending because it is shown that the media's portrayal fuels popular punitiveness in society, despite inaccuracies.[283] This illustrates why the current response is problematic. The current responses to youth offending are failing to address public misconceptions[284] and, consequently, fuelling the negative labels associated with the younger generation. The media

[274] ibid.
[275] Edwin Lemert, '*Human Deviance, Social Problems, and Social Control'* (2nd edn, Prentice Hill 1972), 50 – 51.
[276] Monica Barry and Fergus McNeil, '*Youth Offending and Youth Justice'* (eds, Jessica Kingsley 2009), part 2.
[277] David Farrington, 'The Cambridge Study in Delinquent Development: A Long-Term Follow-Up of 411 London Males' [1977] Institute of Criminology.
[278] Lesley McAra and Susan McVie, 'Youth Justice? The Impact of System Contact on Patterns of Desistance from Offending' [2007] 4 (3) *European Journal of Criminology* 330.
[279] Phil Cohen, '*Rethinking the Youth Question'* (Palgrave 1997).
[280] *Ex parte V* (n 12).
[281] ibid.
[282] Porteous and Colston, 'How Adolescents are Reported in the British Press' [1980] Journal of Adolescence 203.
[283] Julian Roberts and Mike Hough (n 44), 211 – 228.
[284] Marie Gillespie and Eugene McLaughlin, 'Media and the Shaping of Public Attitudes' (The Centre for Crime and Justice Studies, no 49, 2003), 8 – 23.

heightens punitive responses to crime because its representation of youth offending magnifies it to be an issue and accelerates public hysteria in order to drive for harsher punishments and adultification,[285] despite their negative impact on children.

The current YJS failure to acknowledge the detrimental impacts labelling has on youths suggests an alternative approach is required. The implementation of minimum intervention and maximum diversion could be an advantageous alternative which mitigates the stigmatisation of youths. Such approach would address the labelling theory by advocating for less invasive strategies, which do not have negative connotations and, subsequently, prevent a label being placed on children. It would reduce the self-fulfilling prophecy occurring. A minimum intervention and maximum diversion approach would be introduced through informal control mechanisms and facilitated through greater support networks and rehabilitation in the community, in order to resolve the criminal career development and high re-offending rates. It is a minimalistic approach which ensures the child is held responsible, whilst simultaneously addressing labelling concerns and public misconceptions. Unlike the current system, it would provide strategies which have greater alignment to the YJS principles, objectives, and rehabilitative methods. These could include youth cautions, absolute discharges, and reparation orders which are immediately spent and not declared. All of which prevent the labels associated with criminal backgrounds emerging. This is an important difference to the current system because the mitigation of labelling provides children the opportunity to reintegrate into society without stigmatisation;[286] it improves their life chances through employment and support networks which, subsequently, results in reduced recidivism due to a reduction of discriminatory behaviours associated with the criminal label.[287]

To summarise, children are currently being demonised in the media which initiates their criminal career. It is because of this there is a need for an alternative approach, which would likely be satisfied through a minimum intervention and maximum diversion approach. Such approach would be advantageous at achieving the YJS aims, imposing a punishment, and reducing the negative associations with labelling.

[285] Stanley Cohen, '*Folk Devils and Moral Panics: The Creation of the Mods and Rockers'* (eds, Paladin 1972), 12 – 20.
[286] Becker (n 3).
[287] Centre for Justice Innovation, Youth diversion evidence and practise briefing: minimising labelling [2021], 1- 4.

3.2. Age Crime Curve

The age crime curve identifies offending to peak between ages 16 and 24 and decline with age.[288] The current youth justice approach is unsatisfactory because it focuses on the offence commission and does not address the reasons behind offending, such as age. The statistical data provokes the question of whether punitive punishments are required when, in theory, children will outgrow offending.[289] Following such interpretation of the age crime curve it reiterates the developmental model and suggests a minimal intervention and maximum diversion approach would be appropriate[290] because children will stop offending.[291] As previously explored, the informal responses to youth offending offered through minimal intervention and maximum diversion would offer a more suitable alternative which not only addresses a reason behind offending to be age, but would also protect the public, mitigate labelling, and the detrimental effects of custodial sentences. Aligning the YJS approach to a minimum intervention and maximum diversion approach would thus, have evidential support that children grow out of crime regardless of the intervention taken. Therefore, it suggests minimal formalised intervention is suitable. In order to implement such approach, it would require greater support from developmental institutions such as schools, families, and the community. In doing so, the YJS would become more effective, efficient, and favourable because it would have greater acknowledgment for age as a predeterminer of criminality and recognise the evidential explanations of offending.[292]

3.3. Custody

Custodial sentences should be used as a last resort[293] where absolutely necessary.[294] Despite these statutory guidelines, such principle is inconsistent with the YJS practise. Currently, 900 children are in custody[295] for less serious crimes including theft, common assault, cannabis possession, and handling stolen goods.[296] A key problem with these statistics is that

[288] HM Inspectorate of Probation, 'Young Adults' [2021] < https://www.justiceinspectorates.gov.uk/hmiprobation/research/the-evidence-base-probation/specific-sub-groups/young-adults/ > accessed 10th April 2023.
[289] Ibid.
[290] Edwin Schur, *"Radical Non-Intervention: Rethinking the Delinquency Problem"* (eds, Prentice Hall 1973).
[291] Andrew Rutherford, *"Growing out of Crime: The New Era"* (eds, Waterside Press 1986), ch 8.
[292] HM Inspectorate of Probation (n 54).
[293] Sentencing Council, 'Sentencing Children and Young People' [2017] Sentencing Guidelines < https://www.sentencingcouncil.org.uk/overarching-guides/magistrates-court/item/sentencing-children-and-young-people/> accessed 27th April 2023.
[294] Police and Criminal Evidence Act 2019, Code C.
[295] Charlie Taylor, 'Review of the Youth Justice System in England and Wales' [2016] Ministry of Justice and Youth Justice Board 2.
[296] Tim Bateman, 'The State of Youth Justice 2020' [2020] NAYJ Briefing, ch 6.

it casts doubt on the operationalisation of statutory guidelines. It has resulted in the custodial approach being strongly contested for 'damage[ing] children'[297] whereby the environment 'is not equipped to meet their needs.'[298] This is evidenced through 29% of children feeling low, 26% disrespected, and 53% denied daily exercise.[299] Critics have further argued custody to be institutional child 'abuse' which overlooks welfare,[300] as there is evidence of violence, emotional, physical, and sexual abuse;[301] children in custody are being subjected to isolation, humiliation, and assault,[302] with 28 deaths being recorded between 1990 and 2005.[303] A widely held view is that custodial sentences take children "away from their parents […] can be seriously disruptive to their education. […] it carries a social stigma and risks the child being labelled."[304]

Accordingly, custodial sentences are strongly contested for being inadequate, unsatisfactory, and ineffective.[305] The negative associations with custodial sentences are only further accelerated by the limited resources, finances, and staffing. The limited capability of the custodial system is significant because it resulted in a decommissioning of youth institutions, despite the same offending population.[306] Consequently, the decommissioning prevented meaningful engagement, rehabilitation, and accelerated prison overcrowding.[307] Evidently, custodial institutions are not suitable for children. The custodial environment is compounding alienation and fails to address how the institutional failures further subject children to harm. For example, boys are 18 times more likely to commit suicide whilst in custody when compared to those in the community.[308] These statistics and attitudes highlight the current YJS response to offending to be inadequately suited for children. It suggests an

[297] Barry Goldson, 'How Prison Damages Children [2002] <https://www.theguardian.com/society/2002/jul/28/youthjustice.crime > accessed 8th May 2025.
[298] Her Majesty's Inspectorate of Prisons [2000] in Barry Goldson, 'Child Imprisonment: A Case for Abolition.' [2017] 5 (2) Youth Justice.
[299] Tracey Gyateng, Alessandro Moretti, Tiggey May and Paul J. Turnbull, "Young People and the Secure Estate: Needs and Interventions" [2013] Youth Justice Board, Institute for Criminal Policy Research 20 – 36.
[300] Barry Goldson and Deborah Coles, "In the Care of the State? Child Deaths in Penal Custody in England and Wales" [2005] Inquest: London 142
[301] Drake and Butler [2007] in Jo Staines, *"Youth Justice"* (Palgrave Macmillan 2015), ch 3.
[302] Barry Goldson (n 71).
[303] Barry Goldson and Deborah Coles (n 74), 141.
[304] End Child Imprisonment [2019] in Tim Bateman, 'The State of Youth Justice 2020' [2020] NAYJ Briefing, ch 6.
[305] End Child Imprisonment, Why child imprisonment is beyond reform: A review of the evidence [2024], 3 – 27.
[306] Youth Justice Board. Ministry of Justice, Youth Justice Statistics 2015/16 England and Wales; Youth Justice Board. Ministry of Justice, Youth Justice Statistics 2015/16 England and Wales 2017/18.
[307] Charlie Taylor (n 69).
[308] Submission from the Prison Reform Trust < https://committees.parliament.uk/writtenevidence/70373/pdf/ > accessed May 2025.

alternative approach is required thus, supports the need for a minimalistic approach,[309] which could be achieved through abolishing youth custody.[310] If the youth custodial system were to be abolished children would be alleviated from the negative effects[311] and traumatic conditions[312] of custody. Such reform would be academically supported, with many 'advocating for phasing out prison department custody for juvenile offenders.'[313]

The current YJS response to offending appears to be over-ambitious with its aims, because it fails to achieve public protection, prevention of crime, and the rehabilitation of the offender simultaneously. Rather, it is strongly contested for accelerating offending and being ineffective thus, there is a need for an alternative approach which primarily focuses on early intervention and prevention.[314] This could be achieved by diversion out of the YJS entirely and towards parenting classes, youth inclusion, and safer schools programmes which utilise minimal formalised intervention, unless necessary.[315]

3.4. Resources

As asserted, the YJS is characterised through a lack of resources, funding, and staffing. Despite this, this paper argues that the research to date has primarily focused on the detrimental impacts of custody rather than public expenditure. The research into the YJS would be more useful if they included greater considerations for public expenditure, which currently is not treated in much detail. The Youth Justice Board appear to be over-ambitious in their achievements because, on average, they spend £197 million annually implementing custodial sentences yet this expense is not reflected in the impacts of custody.[316] A reasonable assumption would be that such a high expense would be associated with beneficial outcomes yet, this is shown to not be the case. This is an unrealistic ideal. The failing of penal custody draws favouritism to minimal intervention and encourages the abolition of custody.[317] If custody were to be abolished public money could be funded into more effective divisionary

[309] Charlie Taylor (n 69).
[310] Ibid.
[311] Barry Goldson (n 71).
[312] Howard League, 'Children in Prison' [2019] Howard League for Penal Reform < https://howardleague.org/blog/children-in-prison/ > accessed 2nd April 2023.
[313] National Association for the Care and Resettlement of Offenders, 'Phasing Out Prison Department Custody for Juvenile Offenders.' [1989] Juvenile Crime Committee Policy Paper 1.
[314] Sarah Bunn and Penelope Brown, 'Age of Criminal Responsibility' [2018] House of Parliament < https://post.parliament.uk/research-briefings/post-pn-0577/ > accessed 4th April 2023.
[315] Ibid.
[316] Inspectorate of Prisons, Children's Custody: a decade of missed opportunities and decline [2024] < https://hmiprisons.justiceinspectorates.gov.uk/news/childrens-custody-a-decade-of-missed-opportunities-and-decline/ > accessed 8th May 2025.
[317] Barry Goldson (n 63).

methods, such as educational and parental programmes. It would result in reduced public expenditure, which reiterates previous political arguments made by the Conservatives to deliver services for money at reduced public costs.[318] It is a favourable approach for public expenditure which reflects the corporatism model for the most cost-effective justice system.[319] This alternative approach to tackling offending incentivises policymakers and lowers achievement expectations because it advocates for diversion outside of the justice system and towards education, health, and social work. Minimum intervention and maximum diversion being achieved at reduced expenditure would be politically favoured, therefore, tackling the previously outlined political climate barrier.

3.5. Sentencing

The current sentencing approach offers another issue with the YJS as statistics suggests it accelerates youth offending and recidivism. The youth justice system aims to 'prevent offending,' however, such exposition is unsatisfactory because 21% of all individuals cautioned or convicted for indictable offences were under 18, and 22% of 10-to-25-year-olds reported offending in the previous year.[320] These figures suggest the aim to prevent offending is inaccurate because, evidently, it is not being achieved. The issues associated with sentencing could contribute to recidivism, because judicial discretion enables inconsistent treatment, with many youth sentences being questioned for harsher treatment than adults.

In one perspective, sentencing discretion is essential to facilitate bespoke responses to cases before the court. Judicial discretion allows for there to be 'humanity in sentencing' as there is acknowledgment of the individual circumstances which contribute to criminality,[321] whilst simultaneously protecting the public by enforcing necessary precautions, such as harsh sentences.[322] However, an alternative interpretation of judicial discretion raises concern on its consistency. An alternative perspective suggests that judicial discretion risks up-tariffing and net-widening because there is an endless scope of potential interference.[323] This concern is further accelerated by the low age of criminal responsibility which enables children as

[318] Alan Norrie and Sammy Adelman (n 11).
[319] John Clarke and Janet Newman, *"The Managerial State: Power, Politics and Ideology in the Remaking of Social Welfare"* (eds, SAGE Publications Ltd 1997), ch 4.
[320] Stephen Roe and Jane Ashe, 'Young People and Crime: Findings from the 2006 Offending, Crime and Justice Survey' [2008] Home Office Statistical Bulletin, 6 – 7.
[321] Jessica Jacobson & Mike Hough, Mitigation: the role of personal factors in sentencing [2007] Prison Reform Trust 56.
[322] *Ex parte V* (n 12).
[323] Stanley Cohen, *'Visions of Social Control: Crime, Punishment, and Classification'* (eds, Polity 1985).

young as 10 to be subjected to sentencing,[324] and vulnerable to up-tariffing.[325] The latter approach is illuminated through statistics, which demonstrate custodial sentences to be more frequent, severe, and for less violent crimes.[326]

Aligning with the latter perspective, the current sentencing response is failing to acknowledge the significance of a child's age and welfare when delivering their sentence. For instance, any sentence delivered to a child will feel comparatively longer, in comparison to adults, because it is likely to disrupt their education and biological development.[327] Consequently, the latter perspective casts doubt on the former perspective, which attempts to justify sentencing discretion. The former perspective is too simplistic because it fails to recognise judicial discretions contribution towards harsher sentences, longer sentences, and recidivism. All of which are contrary to the YJS aims, objectives, and principles. It illuminates the current inadequacy of youth sentencings and how there is a need for an alternative approach, such as minimum intervention and maximum diversion which would have greater potential at preventing offending.

Accordingly, a minimum intervention and maximum diversion approach could be operationalised through diversion away from custody and wider implementation of the referral order,[328] due to its alignment with restorative justice. This alternative would provide a more effective approach at reducing recidivism as a result of its focus on rehabilitation and re-building relationships between the offender and community, through collaboration on their wrongs and potential reconciliations. There is a stronger focus placed on re-integration and support networks which would tackle criminal labels and discriminatory behaviours following offending.[329] Restorative justice focuses on the presence of a support network, employment, and accommodation to ensure successful re-integration which deters re-offending.[330] Henceforth, a minimal intervention and maximum diversionary approach would be useful to prevent recidivism and align with the YJS aims, objectives, and principles.

---

[324] Children and Young Persons Act 1933**,** s 50
[325] *R v K (A N)* [2014] EWCA Civ 45.
[326] Tim Bateman (n 70).
[327] Sentencing Council, 'Sentencing Children and Young People' [2017], s 2.
[328] Youth Justice and Criminal Evidence Act 1999.
[329] Crown Prosecution Service, Restorative Justice [2023] < https://www.cps.gov.uk/legal-guidance/restorative-justice > accessed 14th April 2025.
[330] Howard Zehr, *The Little Book of Restorative Justice* (Good Books, 2002); Crown Prosecution Service, Restorative Justice [2023] < https://www.cps.gov.uk/legal-guidance/restorative-justice > accessed 14th April 2025.

4. Individual Positivism

Despite the favourability towards a minimal intervention and maximum diversion approach to youth offending previously outlined, an alternative approach to tackling youth crime is established through individual positivism. Individual positivism challenges diversionary strategies for failing to consider the inherent nature of offending where criminals are biologically different,[331] such as in weight and height.[332] It offers an alternative perspective to offending which questions whether any approach is adequate when offenders cannot be changed as offending is a natural, innate, and genetic behaviour. Such perspective emphasises the importance of custodial sentences in ensuring public protection as a result of re-offending being rendered impossible whereas, diversion outside of custody allows for potential re-offending. From this perspective, custodial approaches would be most beneficial. However, whilst individual positivism offers an alternative approach to youth offending it appears to be over-ambitious. This is owed to individual positivism being an incomplete explanation which fails to address environmental influences on youth offending. There is a primary focus on the biological influence to offending, and not environmental influences. Consequently, individual positivism' failure to consider social and environmental factors on youth offending means that such a perspective cannot be used to fully discredit diversion.

5. Barriers to Implementation

As previously outlined, a minimum intervention and maximum diversion approach to youth offending could be successful in achieving the YJS aims and objectives. However, whilst there is potential for success, there must also be recognition that such approach is constrained by barriers to implementation. These restrictions reduce the success minimum intervention and maximum diversion could have. They include previously explored barriers such as the political climate, public expenditure, and due process concerns but another consideration would be time for intervention. This is because, unlike custodial sentences which group offenders together in institutions, divisionary approaches outside of custody focus on the offender as an individual. This, for example, is exemplified through restorative justice which concerns collaboration between the offender and their community.[333] Whilst this is idealistic,

[331] Earnest Hooton, *'The American criminal: an anthropological study. Vol. 1. The native white criminal of native parentage' [1939]* Harvard University Press.
[332] Charles Goring, *"The English Convict: A Statistical Study"* (HMS 1913), 281.
[333] Crown Prosecution Service, Restorative Justice [2023] < https://www.cps.gov.uk/legal-guidance/restorative-justice > accessed 14th April 2025.

a key problem with these diversionary strategies is the limited time the youth justice system has. It is evident that the limited time cannot facilitate each individual case to such depth and detail thus, it would be an unrealistic and impractical ideal of the youth justice system. Whilst minimum intervention and maximum diversion would be most useful where each individual is closely supported on a one-to-one basis this is unachievable in reality thus, introducing another barrier to implementation which would hinder its success.

Conclusion

This paper critically evaluated the potential implementation of a minimum intervention and maximum diversion policy as an approach to youth offending. This was achieved by questioning the ethical dilemma of whether it is appropriate to impose custodial sentences on children and highlighting the damaging effects the current punitive and custodial approach has. It acknowledged how academics have casted doubts on the custodial approach and research which has suggested minimal intervention to be a more beneficial approach, which is achieved through diversion from both custody and the justice system as a whole. The paper explored both forms of diversion, however, found diversion from custody to be more achievable and result in significant improvements, such as mitigating the impact of labelling.

The paper then went on to consider the underlying tension between the youth justice aims and questioned whether public protection and rehabilitation were successfully being achieved. The recognition of custody's detrimental impacts on children presented minimal intervention and maximum diversion to be a more balanced approach in achieving these principles. Such response provides greater facilitation for re-integration and individualistic approaches to offending, which contribute to reduced recidivism. The recognition minimum intervention and maximum diversion has for the differences between children and adults enable it to be more successful at mitigating issues associated with public misconceptions, punitive responses, and the age of criminal responsibility. The paper outlined how such an approach to youth offending is not only supported by academics, but also through empirical evidence relating to the age crime curve and offending statics. Thus, despite potential barriers to implementation and contrary perspectives, the paper concludes that there should be implementation of minimum intervention and maximum diversion in the YJS. It is an approach which has the potential to be successful and efficient approach to tackling youth offending, providing there is recognition of the outlined barriers to implementation.

International Law and Its Role in Conflicts: Through The Lens of Sudan

Gregory Cooke

Abstract

This article looks at the ongoing civil war in Sudan and asks how international law and international criminal law can help in resolving the conflict. The focus is on the role of the International Criminal Court (ICC). It is argued in this article that the ICC will likely have a negative impact on conflict resolution. The ICC arrest warrants can cause the parties subject to them to abandon attempting peace in fear of being arrested as soon as peace occurs. Additionally, the arrest warrants can be manipulated for political purposes further entrenching the conflict. The article also dismisses the International Court of Justice (ICJ) as a venue for conflict resolution due to this conflict not involving multiple states. The article finally discusses how human rights law is also unlikely to have any impact in resolving the conflict due to a lack of willingness from both parties involved in the conflict.

Introduction

In this article, the role that both human rights law and international criminal law play in resolving conflicts and bringing about justice will be analysed. In particular, this paper will examine the role these areas of law can play in the current conflict in Sudan. The Sudanese civil war has been rumbling on since fighting started on the 15th of April 2023, between two military groups who were both partially in charge of governing Sudan.[334] Both sides in this conflict have been accused of serious human rights violations and humanitarian law violations since the fighting broke out.[335] Sudan and its surrounding region have been dogged by civil war and instability for most of their existence since it gained independence in 1956, because of this, an abundance of international legal work has already been carried out in the area and the neighbouring states.[336] In looking at the previous attempts made by

[334] '100 days of conflict in Sudan: A timeline' *Al Jazeera* (24 July 2023) <https://www.aljazeera.com/news/2023/7/24/100-days-of-conflict-in-sudan-a-timeline> accessed 23 December 2024; Sam Hancock and Dulcie Lee 'I haven't slept, I'm terrified': Sudan capital rocked by fighting' *BBC News* (17 April 2023) <https://www.bbc.co.uk/news/live/world-africa-65285254> accessed 23 December 2024.
[335] 'Sudan 2023' *Amnesty international* <https://www.amnesty.org/en/location/africa/east-africa-the-horn-and-great-lakes/sudan/report-sudan/> accessed 22 December 2024.
[336] 'Sudan country profile' *BBC news* (London, 13 September 2023) <https://www.bbc.co.uk/news/world-africa-14094995> accessed 6 May 2025.

the international legal community, we can establish what might be effective this time in bringing about peace and justice.

This article will discuss international law and international criminal law and argue how it is likely to have a negative impact if it is to intervene in the peace process. It will be argued in the article that in pursuing justice, international criminal law often prevents peace from occurring. Secondly, this work will state that human rights law suffers from being ineffectual in relation to bringing about any real change within the context of Sudan. The fundamental nature of Sudan is one of a war-torn country that is concurrently one of the poorest in the world.[337] These factors mean that any real chance of human rights being respected and then being enforced is minimal, as a result, international human rights law can have little impact in Sudan. Overall, this article will demonstrate that, unfortunately, there is little that international criminal or human rights law can do in Sudan, it is a matter beyond the legal realm. That is not to suggest that the international community should do nothing in Sudan, but any discussion beyond the legal realm is beyond the scope of this article.

1. International Criminal Court: Justice and Peace

The International Criminal Court (ICC) was established by the Rome Statute on the 17th of July 1998 and came into force on the 1st of July 2002.[338] The ICC was not the first criminal court to be international in scope; however, it is currently the sole international criminal court and the first to be permanent. The previous ad hoc courts did not have jurisdiction over Sudan, so they will not be discussed in great detail in this article. Due to the prominence of the ICC and the intrastate nature of the Sudanese conflict, this article will largely be discussing the ICC when discussing international criminal law. The International Court of Justice is largely irrelevant in the case of Sudan, as this is an internal conflict and the ICJ deals with disputes between states. Despite this, this article will still briefly discuss the role the ICJ, and other states could play in bringing about peace.

It is worth noting that the ICC and international criminal law primarily promote peace and justice through their role as a deterrent. For the ICC to be engaging with a current conflict, it has, in many ways, already failed in its main contribution to peace and justice. That being

[337] 'GDP per capita (current US$)' (*World Bank Group*) <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD> accessed 6 May 2025

[338] Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3 (RSICC).

said, international criminal justice can still, to a lesser extent, contribute to peace and justice once a conflict has broken out. The main way in which the ICC aims to achieve this is via pursuing prosecutions and issuing arrest warrants for those in breach of the four crimes of the Rome Statute, those being genocide, crimes against humanity, war crimes and the crime of aggression.[339] In this section of the article, the effectiveness of this method by the ICC and how it can be implemented in Sudan will be discussed.

In 2005, the United Nations (UN) Security Council granted the ICC jurisdiction over Sudan and the acts occurring in the Dafur region from 2002 onwards.[340] With no members of the Security Council having large political ties to Sudan, the ICC's jurisdiction looks likely to continue.[341] So, how the ICC proceeds in attempts to ensure justice and peace occur, and what impact their arrest warrants have, remains an important question. Looking at the ICC's involvement in neighbouring countries, until South Sudan's creation in 2011, Uganda and their previous involvement in Sudan reveals the ultimately negative impact the ICC arrest warrants have had on peace. They often prolong conflicts and entrench the conflicts all while attempting to achieve justice, as will be discussed.

The main issue the ICC faces in its attempts to resolve current conflicts is that it does the opposite, its actions often prolong them. The well-intentioned ICC does this by issuing arrest warrants, these come with a plethora of issues and consequences. The main problem with the arrest warrants, as critics put it, is that they undermine the peace process.[342] Combatants in a civil war, such as in Sudan, are more likely to keep fighting if they know that if peace were to ensue, then they will face prosecution from the ICC.[343] So they maintain the status quo that does not bring about peace, but does mean that the ICC cannot get access to the combatants.

However, this widespread criticism is questioned by academic Allard Duursma. He states that 'In short, ICC arrest warrants and summonses to appear provide conflict parties with incentives to engage in mediated peace negotiations'.[344] His logical argument is that those

[339] RSICC Article 5.
[340] United Nations Security Council (UNSC) Res 1593 (2005) (31 March 2005) S/RES/1593.
[341] Aidan Lewis 'Sudan's conflict: Who is backing the rival commanders?' *Reuters* (12 April 2024) < https://www.reuters.com/world/africa/sudans-conflict-whos-backing-rival-commanders-2023-05-03/> accessed 23 December 2024.
[342] J Michael Greig and James D Meernik 'To Prosecute or Not to Prosecute: Civil War Mediation and International Criminal Justice' (2014) 19 International Negotiation 257.
[343] Allard Duursma, 'Pursuing justice, obstructing peace: the impact of ICC arrest warrants on resolving civil wars' (2020) 20 Conflict, security & development 335, 335.
[344] ibid 340.

being prosecuted by the ICC may well seek mediation and a peace process rather than face international prosecution. This is especially true if the negotiations involve an amnesty for any events that occurred during the conflict, as is often the case.[345] The ICC would still have jurisdiction over any eventual Sudanese government, and any amnesties granted would be, in principle, irrelevant to the ICC. Although, in practice, the ICC has found it almost impossible to function without the relevant states' cooperation, ultimately, the ICC is at the behest of state sovereignty.[346] Thus, an ICC arrest warrant in the current conflict could well help speed up the mediation process. Another purpose that the ICC arrest warrant could have is its impact on the wider international community. ICC arrest warrants can put international pressure on a party, especially on the governmental side of the conflict. If they receive arrest warrants and a narrative that they are the "bad guys" starts to form, then even greater pressure can build. The need to mitigate this pressure in order to avoid sanctions, maintain governmental diplomacy or maintain military aid can often be achieved by attending mediation or starting talks and the peace process.[347] The effectiveness of this approach in the Sudanese approach may be limited, as there is not a clear government and rebel distinction between the two groups, as both groups splintered from the power sharing leadership of Sudan, and both sides have equally been accused of atrocities.

While Allard Duursma states the positives that can be found from the ICC attempts at prosecution, he, along with other scholars, is keen to highlight the downsides that also come with this approach. Critics to this view state that the act of beginning mediation and may well be a bluff to alleviate the international pressure for a moment, rather than any actual attempt at resolving the conflict.[348] Additionally, when ICC arrest warrants are only aimed at one side of the conflict, they can have a harmful effect. This was shown in the previous Sudanese conflict when rebel groups were fighting the previous Leader of Sudan, Omar Al-Bashir. The rebel leaders used the ICC arrest warrants as a way of attacking Bashir, they attacked his legitimacy as the president of Sudan. More importantly, they used it as a lightning rod, they intensified their military activity and refused peace talks with the Sudanese Government, all with their own added sense of legitimacy, in part from their

[345] ibid 338.
[346] Kate Gibson, 'Reliance Upon and Complications with State Cooperation', United Nations International Criminal Tribunal for Rwanda, A Compendium on the Legacy of the ICTR and the Development of International Law (2014).
[347] Duursma (n 10) 339.
[348] Greig and Meernik (n 9) 269.

leaching of ICCs legitimacy.[349] The ICC became such a part of warfare that the rebels in Darfur actively wanted to be seen engaging with the ICC and made it a point of difference for them from the Sudanese government.[350] Whether the rebel groups were justified in their stances or not is beside the point, the ICC arrest warrants acted to further entrench the conflict rather than end it. It is hard to see how if the ICC is one sided again in its arrest warrants how this current conflict will be any different.

On the governmental side of the conflict, the ICC arrest warrants also had an entrenching effect. The government was more focused on achieving a military victory than ever. A peace deal and power sharing arrangement could easily lead to the incumbent government being ultimately deposed and therefore at the mercy of the ICC, which would be a great political outcome for the hypothetical government replacing Bashir.[351] Out of the seven arrest warrants issued, only one has resulted in a trial yet.[352] The trade-off for this was both sides digging in and fighting more bitterly. The ICC arrest warrants did nothing but embitter and drag on the last conflict in Sudan. It is hard to see how ICC arrest warrants this time around will serve in resolving the conflict currently engulfing Sudan or in bringing about justice.

The ICC, as well as Sudan, was also heavily involved in the Civil War in Uganda, their actions in this nearby state can also be a guide on how this conflict can be resolved by the ICC. The ICC's involvement in Uganda was dogged by issues throughout, issues of bias, naivety and a general lack of political nous. It is everything that the ICC should avoid in Sudan. The Ugandan conflict was similar in nature to the current conflict in Sudan, both being civil wars in bordering lands of Sudan and Uganda, with the Sudanese government even supporting the rebel group in Uganda, the Lords Resistance Army (LRA).[353] Lessons from one can be applied to the other.

The first issue that soon became apparent in the peace process in Uganda was that of a lack of political nous and inflexibility in relation to state-based amnesty programs. Often in civil wars, a resounding victory or loss is never achieved, and instead, an uneasy peace occurs.

---

[349] Sarah MH Nouwen and Wouter G Werner 'Doing Justice to the Political: The International Criminal Court in Uganda and Sudan' (2010) 21(4) EJIL 941, 957.
[350] ibid 956.
[351] Duursma (n 10) 340.
[352] 'Darfur, Sudan' (International Criminal Court) <https://www.icc-cpi.int/darfur> accessed 10 June 2025.
[353] 'Crises in Sudan and Northern Uganda' (*Human Rights Watch*, 29 July 1998) <https://www.hrw.org/legacy/campaigns/sudan98/testim/house-07.htm> accessed 6 June 2025; 'Ugandan army says Sudan is backing Joseph Kony's LRA' (*BBC news*, 30 April 2012) <https://www.bbc.co.uk/news/world-africa-17890432> accessed 23 December 2024.

Key to maintaining that fragile peace most often is power sharing agreements and an agreement that neither side will be arrested for crimes. A line will be drawn in the sand, and there is an agreement that in order for things to progress, the past must be moved on from, this occurs in the form of an amnesty. They are incredibly important to the peace process, although they are often not just in nature.

This presents a particular issue for the ICC as they take a legalist rather than a pragmatist approach. Legalists are more concerned with justice than peace. There are many arguments that support this legalist view. Scholars Vinjamuri and Snyder argue the many advantages to the ICC view.[354] The first of which is that war crimes trials and ICC trials have an educational role. They teach the people of fledgling democracies and dictatorial states how society should be run and how they deserve justice and human rights.[355] This, in turn, forges peace and stability as more accountability is formed. This is the model that ideally the ICC works through. This model is ripe for criticism for attempting to westernise places that are unwilling or ill-suited to that world viewpoint. Beyond the criticism of this viewpoint, a pragmatist's retort is based in real world truths. They argue that the establishment of peace should be the first priority as all future possibilities of justice are hinged on their first being "social peace and unshakeable democratic institutions".[356] So while the ICC's approach would and does oppose amnesties being granted for international crimes, the pragmatist approach does not. This approach is the more effective one when dealing with situations like Uganda and Sudan, the failings of the legalist approach are shown in Uganda.

Adam Branch discusses the failing of the ICC's inflexible legalist approach. The first point he makes is that these amnesty programmes are popular, in Uganda especially. He states that 'arrest warrants fly in the face of the popular demand for general amnesty, rendering the act inapplicable to the very people to whom it most needs to be applied for peace to arrive'.[357] Without divulging too much into arguments of state sovereignty and state consent, the ICC arguably morally should not ignore the delicate peace process in Uganda that is happening at the behest of the people. Uganda is a member of the ICC, though, so any attempt at this in Sudan is an attempt based on even less legitimacy, as Sudan is not a signatory of the Rome Statute. It also has a damaging effect on citizens of these states and their faith and trust in

[354] Leslie Vinjamuri and Jack Snyder, 'Advocacy and scholarship in the study of the international war crime tribunals and transitional justice' (2004) 7 Annual Review of Political Science 345.
[355] ibid 348.
[356] ibid 353.
[357] Adam Branch, 'Uganda's Civil War and the Politics of ICC Intervention' (2007) 21 Ethics & International Affairs 179, 184.

international criminal law. The ICC's actions show a lack of faith and trust that these states can make their decisions, alienating the population and ultimately reducing the ICC's effectiveness. On top of this, it significantly reduces the state's ability to negotiate peace and to resolve the conflicts. Branch states that the process of getting conflicting parties to the negotiating table is incredibly tricky, but to do that and at the same time stating they will have to appear in the Hague is impossible.[358] He states that in Uganda, 'The unsealing of the arrest warrants prompted head mediator Betty Bigombe to abandon the peace process'[359]. The inflexible nature of the arrest warrants can mean that both sides of the conflict can use them for political and military aims, often prolonging conflicts. The sum of all that has been discussed above leads Branch to claim that the logical conclusion was that the 'Ugandan government called in the ICC against the LRA not to help bring the war to an end but to entrench it'.[360] The Ugandan government used the ICC to further their war efforts and make sure the LRA kept fighting. The war was for the Ugandan government politically useful, and they used the ICC to prolong it by entrenching the LRA and bolstering their perception of legitimacy and, with it, securing foreign aid.

A counter to this narrative is that in Uganda, the ICC did play a part to end the conflict, but not in the way they would have liked. By using the ICC, the Ugandan government were able to make their side of the conflict appear more legitimate, consolidate international support and apply pressure on the Sudanese support of the LRA. The Sudanese government had been funding the rebel movement in Uganda for some time. Supporters of the ICC state that in criminalising the rebel group, this makes it even more difficult for the Sudanese government to support the LRA and continue funding them.[361] Another argument would be that the prosecutions isolate the LRA leadership, making them easy targets or less effective.[362] This could be argued to have been the case, with the LRA now a highly splintered and no longer an effective organisation. However, these claims should be taken with a grain of salt as the fact as, although the Sudanese government officially stopped its support, claims still continued of Sudanese backing the LRA.[363] On top of this, the ICC's contribution to the

[358] ibid 183.
[359] ibid 184.
[360] ibid 185.
[361] ibid 183.
[362] ibid 183.
[363] 'Ugandan army says Sudan is backing Joseph Kony's LRA' *BBC News* (30 April 2012) < https://www.bbc.co.uk/news/world-africa-17890432> accessed 23 December 2024.

LRA's splintering is impossible to calculate due to the lack of information from within the LRA.

The most significant drawback is that it sets a dangerous precedent. The ICC needs to be independent for any proper justice or conflict resolution to occur. The ICC actions in Uganda focused too heavily on the LRA, while governmental forces had no prosecutions, many viewed the ICC's actions in Uganda as biased.[364] If the ICC were to use their arrest warrants as a means of ending conflicts, that would massively impact their impartiality. It would be well beyond their scope, not just theoretically but practically. As Branch states, the ICC lacks the resources or legitimacy to make decisions on the 'democratic credentials of those calling for intervention and to make predictions as to the possible repercussions of its intervention'.[365] Any such action would be a scandal and likely ruin the reputation and legitimacy of the ICC. Therefore, the ICC must stay out of actively trying to resolve conflicts through anything other than prosecuting those guilty of crimes under the Rome Statute.

Uganda shows us that the ICC, through its power and influence, can be used to prolong or alter conflicts. In modern day Sudan, the ICC must learn from Uganda, and if it wishes to provide a conflict resolution role, then it perhaps should avoid issuing arrest warrants until after an amnesty and peace deal have been brokered. If it is to issue arrest warrants, then it should be even handed in how it dishes out these warrants and not target one side.

To summarise, the ICC must be very careful about how it applies its jurisdiction in Sudan. In general, the ICC arrest warrants can be useful for bringing sides to the negotiating table, but they remain quite ineffective at helping bring an actual resolution to the conflict. Uganda has also shown the many weaknesses that the ICC approach can have when dealing with amnesties, the inflexible nature of the ICC can be harmful to peace and conflict resolution. If the ICC is to issue arrest warrants, then they should follow the United States (US)' lead in targeting both sides, to not destabilise the peace process.[366] Ultimately, the ICC is too resolute on achieving justice that it foregoes peace.

---

[364] Mark Kersten, *Justice in Conflict: The Effects of the International Criminal Court's Interventions on Ending Wars and Building Peace* (1st edn, OUP 2016).
[365] Branch (n 24) 195.
[366] 'U.S. says Sudan's rebels have committed genocide and sanctions their leaders' *NBC News* (8 January 2025) <https://www.nbcnews.com/news/world/us-says-sudans-rebels-committed-genocide-sanctions-leaders-rcna18674> accessed 8 January 2025.

2. The ICJ and Its Role in Peace

The ICJ is worth briefly mentioning, as although the ICC is the court most likely to be utilised, in the context of the Sudanese conflict, the ICJ could hypothetically be used. The ICJ deals with disputes between states and does not target individuals like the ICC. Due to this, the ICJ would only be used in an interstate context where one state has approached the ICJ about the conduct of another state. Within the context of Sudan, the ICJ would most likely be approached by a state in relation to genocide and the Convention on the Prevention and the Punishment of the Crime of Genocide.[367] The Genocide Convention does give the ICJ jurisdiction over state committed genocide, this was seen when Bosnia and Herzegovina sued Yugoslavia and more recently with South Africa and Israel.[368] In both these situations, the ICJ gave interim orders to stop acts that were genocidal in nature.[369] A second state taking action to sue Sudan in the ICJ could well be a method of ending the genocide and ensuring a more peaceful Sudan is secured.

However, it is concerning that if any such action did occur, it would be largely ineffectual. The ICJ's main problem is in the enforcement of the orders. In Bosnia and Herzegovina, the interim order from the court was followed up by a second interim order. The second order suggested that Yugoslavia had not adhered to the first order and that it had failed to prevent genocide from occurring.[370] This was commented on by Judge Tonka, who pointed out that the two orders did little to help and could prevent further killings such as the 1995 Srebrenica killings.[371]

The ineffective nature of the ICJ in preventing genocide was again seen recently in Israel and Gaza. The ICJ found that Israel had breached its obligations under the Genocide Convention and should discontinue its actions in the Gaza region.[372] Despite this, the Israeli Government have made no real change in its treatment of the Palestinian people. One reason for the ICJ's failure to stop the genocide in these cases is that of the enforcement method. The UN Charter for the UN Security Council gives the Security Council the power to force

[367] The Genocide Convention (adopted 9 December 1948, entered into force 12 January 1951) UNTS 276.
[368] *Case concerning application of the convention on the prevention and punishment of the crime of genocide (Bosnia and Herzegovina v. Yugoslavia)* (preliminary objections) [1996] ICJ Rep 595; *Application of the convention on the prevention and punishment of the crime of genocide on the Gaza strip (South Africa v Israel)* (provisional measures) [2024] ICJ Rep 3.
[369] *Bosnia and Herzegovina v. Yugoslavia* (n 35).
[370] John Quigley, "International Court of Justice as a Forum for Genocide Cases" (2007) 40 CASE W. RES. J. INT'L L 243, 246.
[371] ibid 251.
[372] *South Africa v Israel* (n 35).

compliance with any ICJ order, this has yet to be used.[373] However, Sudan may well have better luck with the Security Council enforcing compliance than in Bosnia and Gaza. The Sudanese conflict has no connection to or involvement from any of the Security Council states. In Bosnia, Russia was invested in the conflict, and in Gaza, the US is heavily involved, making Security Council intervention virtually impossible.[374] But Sudan has no such links, and so where Gaza and Bosnia were failed, Sudan may not be. However, the lack of links or incentives in the outcome of Sudan means that the Security Council and the wider international community will likely be indifferent and inactive in preventing the genocide.

There are further issues with using the ICJ as a method to secure peace and justice. I would argue that the ICJ is an inappropriate setting for dealing with the Sudan conflict. One of the core issues is that this is a civil war and as such any interstate element is frustrated by lack of clarity on what the Sudanese state is. The ICJ cannot prosecute individuals, so the ICJ must only deal with the state. But what the state is during a civil war is unclear. The nature of the civil war in Sudan is that both sides have splintered off from the government, there is no clear rebel group and no clear governmental or "legitimate" group. The ICJ can only deal with the state, but it is not clear which side represents the state. It would be very damaging and possibly destabilising for the ICJ to weigh in on the conflict and deem one side the quote unquote legitimate and rightful group. The ICJ also lacks the expertise to make that judgement. The requirement for the ICJ to operate at a state level means it is hampered and ineffectual at dealing with genocide issues when a civil war is being waged. With no clear governmental side, an individual approach is what is needed. That is why, along with the ineffectual nature of the interim judgments, the ICJ is the improper court for resolving the Sudanese conflict.

3. Human Rights and Resolving Conflict

Human rights and the international systems built around them are key and instrumental to maintaining a healthy society. However, the key issue that human rights face in attempting to achieve justice is that it can be quite meaningless and ineffectual when encountered by

[373] Quigley (n 37) 245.
[374] Dimitar Bechev, 'Between the EU and Moscow: How Russia Exploits Divisions in Bosnia' (2024) The Carnegie Endowment for International Peace <https://carnegieendowment.org/research/2024/06/bosnia-between-russia-eu?lang=en> accessed 10 June 2025; Tinshui Yeung, 'Trump says Israel will hand over Gaza to US after fighting ends' *BBC news* (6 February 2025) <https://www.bbc.co.uk/news/articles/c4g9xgj2429o> accessed 10 June 2025.

realpolitik. Realpolitik is an approach that is based on real world practicality rather than beliefs based on morality and ideals. The basic stumbling blocks this article will demonstrate is that what human rights suffer from is that in worn-torn states, human rights and the justice and conflict resolution they can bring are unachievable, as the people in power don't value or want human rights.

Dr Manikkalingam highlights many of the issues surrounding human rights. |The core of the issues this article will discuss can be shown in his statement that 'human rights are often promoted as non-negotiable and absolute; conflict resolution is usually about finding the common ground between political actors who hold different positions, including about what is right and wrong'.[375] This statement summarises the lack of real-world applicability and the inflexibility that is by its nature how human rights operate. The universality that is core to the functioning of human rights is perhaps what makes human rights ill-equipped to deal with conflict resolution and achieving justice in situations such as Sudan.

The first practical issue with using human rights mechanisms to achieve justice and peace in Sudan is the real-world application of UN funding and Non-Governmental Organisation (NGO) outreach programs. The UN, through its various agencies, funded many projects and programmes during the previous Sudanese conflict. These programmes aimed at educating the citizens of Sudan of their human rights and, in general, increasing their awareness and therefore protection of their human rights. Whilst these programmes had good intentions and probably did have a positive impact on some, they were riddled with issues, as academic Mark Massoud points out. The first issue that Massoud highlights is that of improper use of the UN funding.[376] Truly independent NGOs in Sudan are often unable to register as charities, as a result, they form a wealthy tax base for the Sudanese government, funded by international money.[377] UN money that was being spent to try and stop the government's human rights violations instead ended up indirectly funding it. Furthermore, due to the relatively high wages the UN paid the programme employees, the 'human rights education programs may have helped fund an elite civil society'.[378] This has made the poorest people's access to human rights resources guarded by a new group of wealthy elites who were reaping the rewards of international funding. Instead of uniting those in need of human rights, the

[375] Ram Manikkalingam, 'Promoting Peace and Protecting Rights: How are Human Rights Good and Bad for Resolving Conflict?' (2008) 5 Essex Human Rights Review 1, 9.
[376] Mark Fathi Massoud, 'Do Victims of War Need International Law? Human Rights Education Programs in Authoritarian Sudan' (2011) 45 Law & Society Review 1, 26.
[377] ibid.
[378] ibid 19.

funding caused a segregated society to emerge. Massoud finally states that NGOs that had close ties to the government abused UN funding. Massoud spoke to activists who stated that 'organizations that they claimed had cozy connections to the regime would also adopt the language of human rights in order to obtain funding from the UN'[379]. It is clear that an authoritarian government could pressure NGOs into attaining UN funding under conditions that the Sudanese government approved of. This issue of funding and corruption is one that would be equally relevant in the current Sudanese conflict. With a lack of proper government now in Sudan, the wealth and resources that the UN provides to the NGOs could very easily become targets to both sides of the conflict. The good will of the UN could easily be co-opted into funding the military campaigns from both sides.

A further and more prominent issue faced by human rights organisations is simply the ineffective nature of human rights in a legal landscape that does not respect them. Massoud suggests that human rights became 'myth like' in Sudan.[380] The main issue is simply the political and military situation in countries like Sudan. For example, UN agencies that trained the Sudanese police on human rights protections have limited impact.[381] A lawyer in Sudan stated that 'They don't violate human rights because of their ignorance … Every totalitarian regime that wants to stay in power will find itself forced to violate human rights. So, you cannot change people through these workshops'[382]. The simple fact of the matter is that those with power in Sudan do not care for human rights, especially while a conflict is ongoing. Massoud even goes as far as to suggest that human rights education can be dangerous, as it creates an expectation from the citizens that the government may adhere to the citizens' demands and therefore advance society in a more peaceful and just direction.[383] Normally, in a properly functioning society, this would be lauded as great success, but in the violent and highly volatile environment of modern day Sudan, this is probably a dangerous thing to instil in the population. Particularly when groups trying to assert their rights is often a cause of conflict in politically volatile situations where the aggrieved parties are unlikely to get any form of justice.[384]

Human rights, on the contrary, it has been argued, can have a positive effect and hasten peaceful resolutions in conflicts. Violence and human rights violations can be reduced by

[379] ibid 18.
[380] ibid 26.
[381] ibid 18.
[382] ibid 19.
[383] ibid.
[384] Manikkalingam (n 42) 1.

the presence of human rights monitors.[385] Dr Manikkalingam states this has been shown to work in Kosovo through the work of the Organisation for Security and Co-operation in Europe and their part in reducing violence in Kosovo.[386] He further states that due to the universal nature of human rights, they can provide an effective and neutral framework for resolving claims and helping transition the peace process away from one of political and military might into one of reasoning and understanding.[387] It should be stated that using human rights as a framework also adds a greater sense of legitimacy to the proceedings. Whilst these principles on paper seem like a compelling success of human rights law, in reality they are not. Manikkalingam goes on to correctly point out that the monitoring work in Kosovo was quickly stopped due to the outbreak of much fiercer fighting.[388] The only real resolution in Kosovo and the surrounding region was due to the NATO bombing and military intervention, not the human rights and international criminal attempts. Attempts at peace via the International Criminal Tribunal for the Former Yugoslavia were largely ineffectual in resolving the conflict. And whilst using human rights as a framework may be useful, any real mediation or negotiation relies on the political context in which peace is being made. A dealing needs to be based on confidence from both sides that peace is the most desirable outcome and one that is likely. Raising human rights violations at a delicate time like that can significantly derail any chance of peace.[389] Naturally, once fighting stops, a significant number of human rights violations will stop, however, social, economic and political rights in states like Sudan would still need to be guaranteed. This could certainly have a destabilising effect and would require far more negotiation that could at any moment cause a break down in the peace process.

In general, human rights and their lofty goals struggle to co-exist with the harsh reality that is present in current day Sudan. There is simply not enough money for, nor the political capital in the idea of human rights, for it to be effective in bringing justice or a peaceful resolution to the Sudanese civil war.

Conclusion

The current conflict in Sudan is nearing its second year and does not look like it will end any time soon. This article has stated that international criminal law and human rights law are

[385] ibid 6.
[386] ibid.
[387] ibid 5.
[388] ibid 6.
[389] ibid 7.

ill-equipped to deal with the conflict in Sudan. The ICJ is not the correct venue for a civil war dispute, leaving other avenues of international law to play their part. International criminal law and the ICC risk jeopardising any meaningful progress towards peace if they intervene in the civil war and the peace process therein. The previous war in Sudan and in the neighbouring state, at the time, of Uganda show that the ICC's arrest warrants can entrench conflicts and stop the peaceful resolution of war. Additionally, human rights law has shown that the grand ideals of human rights are ineffectual when confronted with the harsh reality of conflict, where neither side values human rights all that highly. That is not to say that the international community should or can do nothing, it is to say that the best avenues for affecting real change in Sudan lie outside the legal world.

BIBLIOGRAPHY

Treaties

Rome Statute of the International Criminal Court (adopted 17 July 1998, entered into force 1 July 2002) 2187 UNTS 3

Convention on the Prevention and Punishment of the Crime of Genocide (The Genocide Convention) (adopted 9 December 1948, entered into force 12 January 1951) UNTS 276.

Table of cases

Application of the convention on the prevention and punishment of the crime of genocide on the Gaza strip (South Africa v Israel) (provisional measures) [2024] ICJ Rep 3

Case concerning application of the convention on the prevention and punishment of the crime of genocide (Bosnia and Herzegovina v. Yugoslavia) (preliminary objections) [1996] ICJ Rep 595

Journal articles

Branch A 'Uganda's Civil War and the Politics of ICC Intervention' (2007) 21 Ethics & International Affairs 179

Duursma A, 'Pursuing justice, obstructing peace: the impact of ICC arrest warrants on resolving civil wars' (2020) 20 Conflict, security & development 335

Greig JM and Meernik JD, 'To Prosecute or Not to Prosecute: Civil War Mediation and International Criminal Justice' (2014) 19 International Negotiation 257

Manikkalingam R, 'Promoting Peace and Protecting Rights: How are Human Rights Good and Bad for Resolving Conflict?' (2008) 5 Essex Human Rights Review 1

Massoud MF, 'Do Victims of War Need International Law? Human Rights Education Programs in Authoritarian Sudan' (2011) 45 Law & Society Review 1

Nouwen SMH and Werner WG, 'Doing Justice to the Political: The International Criminal Court in Uganda and Sudan' (2010) 21 EJIL 941

Quigley J, 'International Court of Justice as a Forum for Genocide Cases' (2007) 40 CASE W. RES. J. INT'L L 243

Vinjamuri L and Snyder J, 'Advocacy and scholarship in the study of the international war crime tribunals and transitional justice' (2004) 7 Annual Review of Political Science 345

Books

Kersten M, Justice in Conflict: The Effects of the International Criminal Court's Interventions on Ending Wars and Building Peace (1st edn, OUP 2016)

Online News

Hancock S and Lee D 'I haven't slept, I'm terrified': Sudan capital rocked by fighting' *BBC News* (17 April 2023) <https://www.bbc.co.uk/news/live/world-africa-65285254> accessed 23 December 2024

Lewis A 'Sudan's conflict: Who is backing the rival commanders?' *Reuters* (12 April 2024) < https://www.reuters.com/world/africa/sudans-conflict-whos-backing-rival-commanders-2023-05-03/> accessed 23 December 2024

'Sudan country profile' *BBC news* (London, 13 September 2023) <https://www.bbc.co.uk/news/world-africa-14094995> accessed 6 May 2025.

Yeung T, 'Trump says Israel will hand over Gaza to US after fighting ends' *BBC news* (6 February 2025) <https://www.bbc.co.uk/news/articles/c4g9xgj2429o> accessed 10 June 2025.

'Ugandan army says Sudan is backing Joseph Kony's LRA' *BBC News* (30 April 2012) < https://www.bbc.co.uk/news/world-africa-17890432> accessed 23 December 2024

'U.S. says Sudan's rebels have committed genocide and sanctions their leaders' *NBC News* (8 January 2025) <https://www.nbcnews.com/news/world/us-says-sudans-rebels-committed-genocide-sanctions-leaders-rcna18674> accessed 8 January 2025

'100 days of conflict in Sudan: A timeline' *Al Jazeera* (24 July 2023) <https://www.aljazeera.com/news/2023/7/24/100-days-of-conflict-in-sudan-a-timeline> accessed 23 December 2024

Miscellaneous

‘Crises in Sudan and Northern Uganda’ (*Human Rights Watch*, 29 July 1998) <https://www.hrw.org/legacy/campaigns/sudan98/testim/house-07.htm> accessed on 6 June 2025

‘Darfur, Sudan’ (international Criminal Court) <https://www.icc-cpi.int/darfur> accessed on 10 June 2025

Bechev D, ‘Between the EU and Moscow: How Russia Exploits Divisions in Bosnia’ (2024) The Carnegie Endowment for International Peace

<https://carnegieendowment.org/research/2024/06/bosnia-between-russia-eu?lang=en> accssed 10 June 2025

‘GDP per capita (current US$)’ (*World Bank Group*)

<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD> accessed 6 June 2025

‘Sudan 2023’ *Amnesty international* <https://www.amnesty.org/en/location/africa/east-africa-the-horn-and-great-lakes/sudan/report-sudan/> accessed 22 December 2024

United Nations Security Council (UNSC) Res 1593 (2005) (31 March 2005) S/RES/1593

The Dogmatic Nature of Section 17: Reassessing Utmost Good Faith in Marine Insurance

Lambros Andreades

Abstract

This essay undertakes a critical examination of the historical trajectory and contemporary salience of the doctrine of uberrima fides (utmost good faith) within English insurance contract law, with particular emphasis on Section 17 of the Marine Insurance Act 1906. It posits that the traditional, singular remedy of avoidance for pre-contractual breaches proved inherently impractical and anachronistic, thereby precipitating substantial judicial and legislative reform efforts.

The paper meticulously analyses how the Insurance Act 2015 and the Consumer Insurance (Disclosure and Representations) Act 2012 have fundamentally reconfigured disclosure obligations, effectively circumscribing much of Section 17's erstwhile purview. Subsequently, it delves into the proposed future modalities for the doctrine of good faith, including its potential as an interpretive principle for contractual stipulations, a foundational basis for the implication of terms under the business efficacy test, and a source of judicial flexibility. While acknowledging the inherent limitations attending these emergent roles, the essay concludes that, notwithstanding the abrogation of the draconian avoidance remedy and the advent of more explicit statutory regimes, the foundational principle of good faith retains a pervasive, albeit nuanced, influence. This enduring relevance is manifest in its capacity to inform the content of implied terms, furnish a basis for damages in the event of such breaches, and facilitate equitable intervention against an insurer's unreasonable reliance on contractual prerogatives.

Introduction

In insurance law, since the eighteenth century, it has been well embedded that the doctrine of utmost good faith is reciprocal.[390] The common law rules on pre-contractual information duties found their way to the formation of the Marine Insurance Act 1906 with section 17

[390] *Carter v Boehm* (1766) 3 Burr 1905

specifically emphasising that a marine contract is based upon the utmost good faith,[391] and it identifies the remedy of avoidance if the duty is not observed.[392]

A myriad of case law connotes that the sole remedy offered by this section, namely the "avoidance" of the contract, has created enormous difficulties, as numerous scholars have argued that at the pre-contractual stage, the remedy of avoidance is an impractical remedy for the assured, especially if a breach is detected after the insured has suffered a loss of their property,[393] which consequently lead some to question whether *uberrima fides* is anachronistic.[394] The nature of the duty at the post-contractual phase is often considered to be rigorous,[395] which forced the judiciary to recourse to pioneering interpretive techniques to restrict the scope of the duty of good faith.

It is submitted that the practical scope of the modification of the provision might accelerate the natural elaboration of the doctrine, as an "interpretative principle"[396], implying contractual terms under the business efficacy test and giving some room for flexibility to the judiciary.[397] This paper will explore the evolution of the doctrine of utmost good faith in section I. and hence examine whether or not the abolishment of the remedy of avoidance has the potential to strengthen the duty of utmost good faith.

1. Historical Foundations and the Doctrinal Genesis of 'Uberrima Fides'

Conducive to determining the range and nature of the remedies accessible in the judicial toolbox for the failure to observe the duty of good faith, a discernment of the juristic basis of the duty and the remedies linked to the duty is vital.[398] Throughout utmost good faith's evolutionary journey, numerous scholars have advocated that there is a motion afoot to safeguard that the panoply of remedies within the realm of the law might be utilized for any

[391] Marine Insurance Act 1906, Section 17
[392] ibid, section 20
[393] B. Soyer "*The insurer's duty of good faith: is the path now clear for the introduction of new remedies?*" in M Clarke and B Soyer (eds), *The Insurance Act 2015, A New Regime For Commercial and Marine Insurance Law* (Informa Law from Routledge 2016), 38
[394] Francis Achampong, "*Uberrima Fides in English and American Insurance Law: A Comparative Analysis*" (1987) 36 (2) I.C.L.Q. 329, <https://www.jstor.org/stable/759999> accessed 20 May 2023
[395] *Manifest Shipping Co Ltd v Uni-Polaris Shipping Co Ltd (The Star Sea)* [2001] UKHL 1; [2003] 1 AC 469, [6]; *Pan Atlantic Insurance plc v Pine Top Insurance* [1994] 3 All E.R. 581; Norma J. Hird, "*Utmost good faith – forward to the past*" (2005) Journal of Business Law 257, 258
[396] Law Commission and Scottish Law Commission, *Insurance Contract Law: Business Disclosure; Warranties; Insurers' Remedies for Fraudulent Claims; and Late Payment* (Law Com No 353, July 2014)
[397] ibid, [30.22]
[398] *Banque Financière de la Cite v. Westgate Insurance Co. Ltd* [1990] 1 Q.B. 665; *aff'd* [1991] 2 A.C. 249

transgression, most notably a breach of the duty of utmost good faith.[399] Devoid of principle, the urge for justice can often descend into implacable justice or merciful, unjust beneficence.[400] The doctrine of utmost good faith never took hold in general contract law as there is no duty to disclose information that would be likely to act on the other party's decision to conclude the contract,[401] as insurance contracts belong to the exceptional category of contracts.[402]

1.2. Juristic Foundations and Early Challenges

There have been some suggestions that the duty of utmost good faith had its roots in shared moral values[403] and equity, as Lord Steyn designated it to give effect to reasonable expectations[404] of honest people.[405] The notion that parties should not vigorously deceive one another was highlighted by Lord Bingham in *Interfoto,*[406] as in his view the doctrine of good faith is perceived by numerous civil law jurisdictions as being an overriding principle of contract law which ultimately entails fair and open dealing in the making and performance of contracts. These duties might have the capability to encompass a spirit of supporting the contract's objectives which can be extended to the duties of not misrepresenting material facts to encouraging harmonic cooperation and openness in dealing, which divergences undermining these goals of the duty to not act capriciously or unreasonably.[407]

Albeit the fact that the impulse of good faith has a customarily clear and persistent meaning through case law, fair dealing on the contrary, may have an amorphous connotation as to the exact fidelity normally expected of implied terms by custom and practice. Lord Steyn

---

[399] Sir A.Mason, "*The Place of Equity and Equitable Remedies in the Contemporary Common Law World*" (1994) 110 LQR 238; Sir P.Millett, "*Equity's Place in the Law of Commerce*" (1998) 114 LQR 214.
[400] Peter MacDonald Eggers, "*Remedies for the failure to observe the utmost good faith*" (2003) LMCLQ 249, 251; Johanna Hjalmarsson, "*Insurers and the law of fraud - A success story and the case for regulatory intervention*" in D. Rhidian Thomas "*The Modern Law of Marine Insurance*" (1st ed, Informa Law from Routledge 2023), ch.6, 123; Marc Galanter, "*Why the'Haves' Come Out Ahead: Speculations on the Limits of Legal Change*", (1974) 9:1 Law and Society Review; *Feise v Parkinson* (1812) 4 Taunton 640, 128 ER 482; Third Parties (Rights against Against Insurers) Act 2010
[401] B. Soyer, "*Reforming the assured's pre-contractual duty of utmost good faith in insurance contracts for consumers: are the Law Commissions on the right track?"* (2008) 5 J.B.L. 386, 386
[402] *HIH Casualty & General Insurance Ltd v Chase Manhattan Bank* [2003] UKHL 6; [2003] 2 Lloyd's Rep 61; [2003] 2 Lloyd's Rep IR 230 (Woodborough LJ) [85]; *Bell v Lever Bros Ltd* [1932] AC 161, 227.
[403] *Versloot Dredging v HDI Gerling (The DC Merwestone)* [2016] UKSC 45; [2016] 2 Lloyd's Rep 198; [2016] Lloyd's Rep IR 468; [2017] AC 1, [10] (Sumption LJ)
[404] *First Energy (UK) Ltd. v. Hungarian International Bank Ltd. [*1993] 2 Lloyd's Law Reports 194, 196; Ewan McKendrick "*Good Faith in the Performance of a Contract in English Law*" in Larry DiMatteo and Martin Hogg (eds.), Comparative Contract Law: British and American Perspectives (OUP 2016), 196
[405] Johan Steyn *'Contract Law: Fulfilling the Reasonable Expectations of Honest Men'* (1997) 113 LQR 433
[406] *Interfoto Picture Library Ltd v Stiletto Visual Programmes Ltd* [1987] EWCA Civ 6; [1989] QB 433
[407] Howard N. Bennet, "*Utmost good faith in the House of Lords*" (1995) 111 L.Q.R. 181, 182; Margaret Hemsworth, '*The fate of "good faith" in insurance contracts*' (2018) L.M.C.L.Q. 143, 145

laconically stated that 'fair dealing' is a behaviour which observes reasonable commercial standards,[408] which in the framework of insurance contracts might be best appropriately captured by the abundance of expectations of the financial services regulators, precisely by ICOBS and COBS,[409] and can even give a rise to a claim in damages.[410] However, there are limits to the effect of good faith and fair dealing in insurance contracts, as they do not give rise to rights of disclosure or loyalty, the true intentions are revealed by the implied and express terms, albeit that s.17 allows a range of meaning of good faith depending on the context.[411]

1.2. The Jurisprudential Strain: From Pre-Contractual Ideal to Post-Loss Impasse

Consequently, Lord's Manfield's approach to insurance contract cases had become a 'panacea' for truthful behaviour between the parties,[412] where the law requires parties to act in *uberrimae fidei* in their mutual dealings at the pre-contractual stage,[413] restating a passive duty to refrain from misrepresentation but equally important imposing a duty to both parties to volunteer certain information.[414] Conversely, this proposition was controversial among scholars, as the act appeared to imply that it was designed for the pre-contractual presentation of the risk to the insurer but nevertheless applied after the presentation of the risk, upon the exposition of a claim.[415] The rhetoric of good faith intends to be applied between parties once entered into a contract, but no definition of their joint contractual duties beyond those in the contract was ever clear. Hjalmarsson stressed that by presuming that this doctrine was held to denote a mutual duty, every opportunity to allow it to outline the insurer's duties was in some way lost.[416] Hence, case law signifies that in a failure to disclose fraud by the insured's agent, the insurer did not owe its insured damages,[417] nor where their payment of the indemnity was unreasonably late.[418]

---

[408] Johann Steyn, "*The role of good faith and fair dealing in contract law: a hair-shirt philosophy*?" (1991) Denning LJ 131; George Leggatt, "*Contractual duties of good faith*" (Lecture to the Commercial Bar Association on 18 October 2016) < https://www.judiciary.uk/wp-content/uploads/2016/10/mr-justice-leggatt-lecture-contractual-duties-of-faith.pdf> accessed 09 May 2023
[409] Hemsworth (n.18), 148
[410] Financial Services and Markets Act 2000, s.150
[411] *Yam Seng Pte Ltd v International Trade Corp Ltd* [2013] EWHC 111 (QB); [2013] 1 Lloyd's Rep 526; *MSC Mediterranean Shipping Co SA v Cottonex Anstalt* [2016] EWCA Civ 789; [2016] 2 Lloyd's Rep 49
[412] Hjalmarsson (n.11), 125
[413] Soyer "*Reforming the assured's pre-contractual duty of utmost good faith in insurance contracts for consumers: are the Law Commissions on the right track?"* (n.12), 386
[414] Hemsworth (n.11), 146
[415] ibid, 125
[416] ibid
[417] *La Banque Financière de la Cité SA* (n.1).
[418] *Sprung v Royal Insurance (UK) Ltd* [1999] Lloyd's Rep IR 111

Therefore, the ambiguity of the application of the pre-contractual paradigm to the post-contractual state ultimately led to its collapse, and it was left to the judicial branch to cultivate an alternative remedy.[419] Albeit critics,[420] numerous scholars advocated the need for reform,[421] and recommendations have been made by the Law Commission since 1957.[422] Hjalmarsson stressed that complementary sanctions such as damages for breach ought to have been deemed available, given that the "idiosyncratic" remedy of avoidance *ab initio* was undoubtedly inappropriate to the mutual contractual context,[423] as this would have required utmost good faith to be perceived as an implied term, which has already been rejected by the Court of Appeal.[424]

3. Initiatives for The Future Role of Utmost Good Faith

3.1. Reforming Pre-Contractual Disclosure: The Fair Presentation Duty

The Law Commission had several aspirations to orchestrate a greater role of Section 17, it has been argued that the duty of utmost good faith that it may serve *inter alia* as a powerful interpretative principle,[425] albeit that none of the suggestions features within the reformed legislation of the 2015 Act,[426] which got rid of all the abysmal shortcomings of the remedy of avoidance and changed the doctrine into a renewed system of interpreting the duty of "fair presentation".[427] The Insurance Act 2015 further clarifies the knowledge of an insured, as with respect to individuals they are deemed to what is known to them,[428] but if the insured

---

[419] Hjalmarsson (n.11), 125; Philip Rawlings and John Lowry, "*Insurance Fraud: The 'Convoluted and Confused' State of the Law*" (2016) LQR (132) 119, 119; Deborah Tompkinson, "*Duration of the duty of utmost good faith*" 5(2) Int. I.L.R. 1997, 40

[420] R. A. Hasson, "*The Doctrine of Uberrima Fides in Insurance Law – A Critical Evaluation*" (1969) 32 (6) TMLR < https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2230.1969.tb01239.x> accessed 29 April 2023; R. A. Hasson*, "The Basis of the Contract Clause" in Insurance Law"* (1971) 34(1) TMLR 29 <https://www.jstor.org/stable/1094022> accessed 28 April 2023; Richard Aikens, "*The post-contract duty of good faith in insurance contracts: is there a problem that needs a solution?"* (2010) 5 J.B.L. 379, 393

[421] A. Tarr and J. Tarr, "*Utmost Good Faith in Insurance: Reform Overdue?*" (2003) 10 Asia Pacific Law Review 171, 183; *Soyer* "*Reforming the assured's pre-contractual duty of utmost good faith in insurance contracts for consumers: are the Law Commissions on the right track?"* (n.12), 387; *Rose* (n.18), 224

[422] Law Reform Committee Fifth Report, *Conditions and Exceptions in Insurance Policies* (Cmnd. 62, 1957); Law Commission of England and Wales, *Insurance Law: Non-Disclosure and Breach of Warranty* (Law Com 104, 1980); *Law Commission and Scottish Law Commission* (n.12)

[423] *Banque Financière* (n.14), [563] (Slade LJ)

[424] ibid, [548] (Slade LJ); Hjalmarsson (n.11), 126; *Joel v Law Union Insurance Co* [1908] 2 K.B. 863 [886]; March *Cabaret Club Ltd v London Insurance* [1975] 1 Lloyd's Rep. 169 [175] (May LJ); Banque *Keyser Ullmann SA v Skandia (UK) Insurance Co Ltd [*1990] 1 Q.B. 665 [701] – [702] (Steyn LJ)

[425] B. Soyer & A. Tettenborn, "*Mapping (Utmost) Good Faith in Insurance Law- Future Conditional?"* (n.69) 620

[426] ibid; Hemsworth (n.18), 154

[427] Insurance Act 2015, Section 3(1); Mingting Zhu, "*The Utmost Good Faith in Marine Insurance: The Nature*" (2020) 11 Beijing Law Review 99, 104 < https://www.scirp.org/iourna I/birResearch> accessed 10 May 2023

[428] ibid, s 4(2)

is not an individual, they are deemed to know what is known by the insured's senior management.[429] The notion of fair presentation of risk is rooted on the basis that both parties are anticipated to act in good faith in exchanging information, but if the court discovers that the insured had deliberately refrained from disclosing only the bare minimum of information, in anticipation that the insurer would fail to make auxiliary enquiries to reveal the full picture, therefore there would be a breach of the duty of fair presentation, as the insured did not act in good faith in a deliberate or reckless manner.[430]

3.2. Section 17's Role as An Interpretative Principle: Limitations and Challenges

However, Hemsworth has submitted her scepticism, as this role might open the floodgates for the likelihood of a different test from that expounded by CIDRA and the Insurance Act 2015, or to insinuate the prescribed remedied by the statute may be subject to the application of "discretionary judicial restraint" instituted upon the absence of good faith.[431] If the insurer is notified and advances to ask further questions, in this scenario and the insured responds as the law entails, then there is no infringement as to the express provisions are concerned and it is challenging to perceive what capacity there is for some "overriding judicial assessment of his *bona fides,"* but nevertheless, the equivocal wording of the statutory provisions do not guarantee clarity on this issue. [432] Conversely, in the case that the insured endures in his concealment in response to the insurer's auxiliary requests, then there is modest exertion in employing the suitable statutory provision to determine that there is a breach of the duty of fair presentation of the risk under section 3 of the Insurance Act 2015, [433] or in breach of the duty to take reasonable care in avoiding misrepresenting matters material to the risk under section 2 of CIDRA.[434] Nevertheless, a failure to respond is probable to designate an affirmation[435] or a positive statement by interference[436] gives rise to the insurer's remedies under the schedule of each Act.[437]

---

[429] ibid, s 4(8)(c).
[430] Law Com No 353, [30.23]; Christopher Chen Chao-hung, "*Insurance Act 2015 in the UK and Lessons for Singapore*" (2016) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2928418> accessed 23 April 2023
[431] *MSC Mediterranean Shipping Co SA* (n.51), [40] (Rix LJ)
[432]Hemsworth (n.18), 154
[433] ibid, 155
[434] Wang Xin and Shan H. Jun, "*Review of the current hull insurance law of China: in comparison with the UK rules*" (2017) 2 J.B.L. 121, 125
[435] *Grupo Nacional Provincial SA v WISE* [2004] 2 Lloyd's Rep 483
[436] R. M. Merkin, "*Colinvaux's Law of Insurance*" (13th edn, Sweet & Maxwell, 2022), ch.7, 7-238
[437] Hemsworth (n.18), 155

Scholars have commented that with regard to section 17, does not enhance the explicit statutory provisions, but conversely, it complicates and undermines them, as the scope is less clear in the case where the insurer neglects to respond to further enquiries. Arguably, there is some tension between the effect of the insured's privation of good faith and the insurer's failure to make an enquiry, as the latter in certain scenarios itself unveils some absence of good faith by the insurer, due to the fact that the statutory provisions do not offer for such a possibility and the absence of that might encourage an extrapolation that the statutes do not propose the insurer to have a remedy as both CIDRA and the Insurance Act 2015 prescribe "remedies" as opposed to "rights", which perhaps leaves a confined space for judicial discretion.[438] Soyer and Tettenbron share a similar view with Hemsworth, that there is no room for a separate concept of "utmost good faith interpretation" interpretation envisaged by the Law Commission, thus section 17 does not have the effect of altering the outcomes by the mode of remedy specified by the statutory provisions, [439]and the more explicit regime set out by CIDRA and the Insurance Act 2015 has effectively displaced the role of section 17 in this setting.[440]

The Law Commission has envisioned a second role for Section 17 to "inform the need to imply contractual terms into the policy under the traditional 'business efficacy' test, as good faith provides a background when considering whether it is necessary to imply a particular term."[441] In the context of the implication of terms, good faith should be deemed as a component of business efficacy where the contractual relationship is one of good faith, rather than s.17 specifically.[442] This implies that following the abolishment of the avoidance remedy, the courts may exercise more freely to conclude that good faith, notwithstanding the semblance of a more explicit form of the implied term, forms a part of a particular insurance contract, as this approach aims to grant greater judicial flexibility to temper contractual in a manner consistent with general contract law, therefore the courts can imply terms which would not be considered essential outside the field of insurance, by allowing both the implication of positive rights as well as restrictions.[443]

[438] ibid
[439] B Soyer & A Tettenborn, "*Mapping (Utmost) Good Faith in Insurance Law- Future Conditional* (n. 36), 622
[440] *(The Star Sea)* (n.6), [7] (Clyde LJ); Hemsworth (n.18), 155
[441] Law Com No 353, [30.23]
[442] Ibid, [30.58]
[443] Ibid, [30.60]

It is worth noting, that the intention is not that good faith is itself the implied term,[444] but rather there is the outlook of the doctrine as a facet of the law and as a pervasive presence in the contractual relationship in insurance, hence it may make it easier for courts than has previously been before possible to imply some less amorphous term.[445] But the vision of this fades somewhat, in light of the existing restrictions, specifically that a term will not be implied when it would be in conflict with an express term,[446] and the business efficacy will often present an impediment to transcend, notwithstanding that the term sought to be implied necessitates not being apparent.[447] Furthermore, as seen in the recent case of *Bristol Groundschool*,[448] by applying the rationale in *Yam Seng*[449] the implications prerequisites to be beyond merely reasonable, equitable or fair, as business efficacy in its practical definition implies that deprived of the term, the contract would lack commercial sense rather than its presence is unequivocally essential,[450] as the implied term must be identified with a high level of precision, and hence an austere version of good faith is implausible to be sufficient. This position is differs in Australian jurisprudence,[451] where the 1984 Insurance Contract Act made utmost good faith an implied obligation in every insurance contract.[452] Tar highlighted the observations made by the Supreme Court of Canada,[453] as the common law has resisted acknowledging any generalized and independent doctrine of good faith in the performance of contracts, and by applying two incremental steps, namely acknowledging that good faith contractual performance is a general organizing principle which underpins various rules and obligations[454] and secondly to recognizing, as a further manifestation of the doctrine, indicates that there is a common law duty which applies to all contracts to act honestly in the performance of contractual obligations. The court stated that a duty would be

[444] *Banque Keyser Ullmann SA* (n.65)
[445] *K/S Merc-Scandia XXXXII v Certain Lloyd's Underwriters (The Mercandian Continent)* [2001] EWCA Civ 1275; [2001] 2 Lloyd's Rep 563 ; [2001] Lloyd's Rep IR 802
[446] *Marks and Spencer Plc v BNP Paribas Securities Services Trust Co (Jersey) Ltd* [2015] UKSC 72; [2016] AC 742; *Federal Mogul Asbestos Personal Injury Trust v Federal-Mogul Ltd (formerly T & N Plc)* [2014] EWHC 2002 (Comm); [2014] Lloyd's Rep IR 671
[447] *Société Générale v Geys* [2012] UKSC 63; [2013] 1 AC 523, [55] (Hale LJ)
[448] Bristol Groundschool Ltd v Intelligent Data Capture Ltd [2014] EWHC 2145 (Ch.)
[449] *Yam Seng Pte Ltd v International Trade Corp Ltd [2013] EWHC 111 (QB)*; [2013] 1 Lloyd's Rep 52
[450] ibid; *cf Mediterranean Salvage & Towage Ltd v Seamar Trading & Commerce Inc (The Reborn)* [2009] EWCA Civ 531; [2009] 2 Lloyd's Rep 639
[451] Kate Lewins, "*Going walkabout with Australian insurance law: the Australian experience of reforming utmost good faith*" (2013) 1 J.B.L. 22; 4
[452] Insurance Contracts Act 1984, ss 13–14;
[453] Julie-Anne Tarr, "*A growing good faith in contracts*" (2015) 5 J.B.L. 410, 413
[454] *Bhasin v Hrynew* 2014 SCC 71, [33]

established that is just, accords with the "reasonable expectations of commercial parties and is sufficiently precise that it will enrich rather than detract from commercial certainty.[455]

Equally, there is a potential force that good faith could be an influential tool in which an express or implied term is construed, hence good faith could be capable of seizing control over the way in which the term is to be contented.[456] The Law Commission validated the rationale in *Gan Insurance*[457] as a "qualification to an absolute right to withhold endorsement to a settlement to the effect that any such withholding had to be exercised in good faith, not capriciously or arbitrarily[458] and had to be constructed on precise facts rather than by reference to exclusive superfluous matters on the form of conduct which no reasonable person would have envisaged".[459] The decision in *Groom v Crocker*[460] supports the notion of the law exercising restraint on the decision-making capacity of an insurer, specifically the liability of the insurer's invariable right to take over and control legal proceedings against the insured, in light of the obvious potential conflict of interest involved, required the insurer to give due considerations to the circumstances, albeit no direct reference to utmost good faith has been made.[461]

Furthermore, there have been some suggestions that the doctrine of good faith may have a greater role to inform the interpretation of express rights in the context of liability insurance, although there is no clear scope as having any independent force, as it requires both the insured and the insurer to have a common interest and a united stance in regard for the commercial interests of one another. Moreover, there has been some cautious judicial reference to the doctrine of utmost good faith in the context of the construction of the contract terms, particularly in relation to those terms which allow the insurer rights of control over the conduct of the insured's defence, in the scenario where there is some doubt as to the extent of cover based on commercial considerations.[462] This inherent potential for conflict of interest, was raised in *Goshawk*[463] where the Court of Appeal was prepared to imply a

[455] ibid, [34]
[456] Hemsworth (n.18), 154; Jonathan Gilman, Robert Merkin, Claire Blanchard, Mark Templeman, "*Arnould's Law of Marine Insurance and Average*", 18th edn (2013), [18.19]
[457] *Gan Insurance Co Ltd v Tai Ping Insurance Co Ltd (No 2)* [2001] EWCA Civ 1047; [2001] Lloyd's Rep IR 667 .
[458] *Paragon Finance plc v Nash* [2001] EWCA Civ 1466; [2002] 1 WLR 685, [36] (Dyson LJ); *Socimer International Bank Ltd v Standard Bank London Ltd* [2008] EWCA Civ 116; [2008] 1 Lloyd's Rep 558
[459] Ibid, [68] (Mance LJ)
[460] [1939] 1 KB 194
[461] Ibid, [203] – [204]; *Cox v Bankside Members Agency* [2985] CLC 671, 680, 684
[462] Hemsworth (n.18), 158
[463] *Goshawk Dedicated Ltd v Tyser & Co Ltd* [2006] EWCA Civ 54; [2007] Lloyd's Rep IR 224 .

term into the contract to a right to inspect and take copies of placing and claiming documents held by the brokers, but Rix LJ indicated that albeit the implication was to be made on the traditional basis that is essential for business efficacy, the doctrine of good faith that applies in the insurance context supported that conclusion.[464] Notwithstanding the reference to the "mutual duty of good faith" in numerous cases, the source of this implied term in the insurance contract is not section 17, thus the amended statutory provision is unlikely to lead to any change.[465]

The final role envisaged for the doctrine of good faith is to allow some space for judicial flexibility, as it can provide a solution to stiff cases or emergent difficulties, albeit the Law Commission has acknowledged the infrequency of such cases and the possibility of the concept's being utilized to avert an insurer from relying on a right to deny a claim, which could be deemed as unfair, as it conflicts with the duty of honesty and fair dealing.[466] The Law Commission in its review has concluded that the role of good faith might play a role beyond the pre-contractual stage, namely variations to risk/held-covered clauses,[467] renewals, cancellation, information; fraudulent claims, occasions where good faith may be implied, and litigation.

Following the decision of *Star Sea,*[468] it is evident that the Law Commission has acknowledged that as to litigation, the duty of good faith cannot endure beyond the commencement of proceedings,[469] from which point in time the parties' conduct is monitored by procedural rules and by the integral jurisdiction of the court. In regards to the provision of information in the post-contractual stage, the insured is not obliged to make disclosure and this is not altered even if the insurer may exercise his cancellation rights.[470] The Law Commission deems it rational for the courts to have the right to have imperilled such terms to close scrutiny and to interpret their realm restrictively, considering that the insurer assesses the risk and the appropriate premium as at inception, in any event, such terms are probably to take effect as exclusions or as warranties, in light of ss 9 and 10 of the Insurance Act 2015 and Part 2 of the Consumer Rights Act 2015, which pursues to control unfair terms.

---

[464] ibid, [53]; *Eagle Star Insurance Co Ltd v Cresswell* [2004] EWCA Civ 602; [2004] Lloyd's Rep IR 537 , [54]
[465] Soyer "*The insurer's duty of good faith: is the path now clear for the introduction of new remedies?*" (n.4), 48; Hemsworth (n.18), 159
[466] ibid
[467] Insurance Act 2015 section 2; CIDRA 2012 Section 2
[468] *The Star Sea* (n.6)
[469] *Emirates Trading Agency LLC v Prime Mineral Exports Pte Ltd* [2014] EWHC 2104 (Comm); *Cable & Wireless v IBM* [2002] EWHC 2059 (Comm).
[470] *The DC Merwestone* (n.14), [67] (Hughes LJ)

Moreover, consumers and small businesses may seek a remedy using the services of FOS[471] if it appears to be fair and reasonable.[472] As seen in *Alfred*[473] the court accepted that intentional concealment of misinformation might be a breach of the duty of utmost good faith and similarly in *Phoenix*,[474] Hobhouse J agreed that there would be an obligation of good faith in cases of liability insurance and reinsurance to keep adequate records and make them available to the insurer to allow him to keep track of his liabilities.[475] However, there is no clear scope of what remedy would be in such a case, albeit it is safe to conclude that there is no reason why remedies other than avoidance should not be available.[476]

The relationship between the duty of utmost good faith and the fraudulent claims rule has become more perspicuous, following the amendments made by Part 4 of the Insurance Act 2015 as well as the rationale in *The DC Merwestone.*[477] The consequences of a fraudulent claim are laid out in ss12 and 13 of the Insurance Act 2015, which builds on prior case law and confirms the "guillotine effect" to the contract, where the insurer is not liable to pay the claim, might be able to recover any sum paid on it and may terminate the contract while retaining any premium already paid.[478] Soyer has argued that no well-defined attempt was made through the statutory reforms to elucidate or modify the scope of fraudulent claims, nor to clarify the relationship with the fraudulent claim provisions and the duty of utmost good faith, albeit the removal of the avoidance remedy in this context has expediated the reform through the common law.[479] It deserves to be mentioned that the rationale in *The DC Merwestone,*[480] indicated that the duty of utmost good faith and the fraudulent claims rule is closely related as both are seen as rules of law,[481] albeit distinct concepts, as a dishonest lie in the post-contractual stage, will not result in a loss of claim, nor the loss of insurance cover. The epithet "collateral lie" connotes a dishonest claim and is accessed through the materiality test with the application of retrospection to the existing facts.[482] Yet, according to Hemsworth's analysis a dishonest but collateral lie, albeit not sufficient to place the insured

[471] Financial Ombudsman's Service
[472] FCA Handbook, DISP; The Financial Services and Markets Act 2000, s.228.
[473] *Alfred McApline plc v BAI (Run-Off Ltd)* [2000] 1 Lloyd's Rep. 437
[474] *Phoenix General Insurance Co v Halvanon Insurance Co Ltd* [1985] 2 Lloyd's Rep 599
[475] ibid, [614]
[476] *The Mercandian Continent* (n.93), [22] (Longmore LJ)
[477] *The DC Merwestone* (n.14)
[478] Hemsworth (n.18), 160
[479] Soyer "*The insurer's duty of good faith: is the path now clear for the introduction of new remedies?*" (n.4), 51
[480] *The DC Merwestone* (n.14)
[481] ibid, [8] (Sumption LJ)
[482] ibid, [111]; [127–128

in breach of the fraudulent claims rule, may still leave him in breach of the duty of utmost good faith. However, notwithstanding the amendment to section 17 and absent clear express contractual terms to another effect, such a breach seems to have no legal costs in the insurance contract realm.[483]

3.3. Redefining Post-Loss Obligations: The Implied Duty to Pay Claims and the Contraction of Avoidance

On the other hand, good faith may be implied, as it has been addressed by the reform made by the Insurance Act 2015, Part 4A, specifically section 13A, which imposes an implied term on an insurer to pay a valid claim within a reasonable time period, where the nature of the conduct of the insurer will be the determining factor in an issue of a breach.[484] The Law Commission has stressed whether the duty of good faith is capable of confining avoidance under s.17, notwithstanding the addition of the proportionate remedies under the statutory reforms, however, it is an open-ended question among commentators.[485] The notion that the duty of utmost good faith is mutual and there are deliberate attempts to uphold it in instances where the insurer asserts a right to avoid for non-disclosure.[486] In obiter remarks, the matter was more distinctly addressed in *Drake*[487] as Rix LJ[488] and Clark LJ[489] noted that the doctrine of good faith should be capable of limiting an insurer's right to avoid situations where that "draconian" remedy would operate deceitfully, and Pill LJ respectively suggested before taking the exorbitant step of avoidance, the insurer had a duty to make further enquiry of the facts.[490] Arguably, the magnitude of these dicta has been distinctly reduced by the duty of disclosure by the reformed Acts, and there is a significant shift in the balance between insured and insurer, as in light of the historical context of the dicta noted in *Drake*, the remedy of avoidance was capable of operating harshly and it was the sole available remedy. The dynamic for the justification for the use of the doctrine of good faith has been arguably substantially reduced, as seen in *Drake* that an insurer seeking to rely on the insured's non-disclosure of material information was obliged to demonstrate that it would not have entered

[483] Hemsworth (n.18), 161
[484] ibid, 162
[485] Law Com No 353, [30.55]
[486] *Pan Atlantic Insurance plc* (n.6), 555
[487] *Drake Insurance Plc v Provident Insurance Plc* [2003] EWCA Civ 1834; [2004] 1 Lloyd's Rep 268; [2004] Lloyd's Rep IR 277; [2004] QB 601.
[488] ibid, [87]
[489] ibid, [144]
[490] ibid, [177]

into the contract or would have charged a different premium,[491] which in the instance of consumer insurance, CIDRA abolished the duty of disclosure, replacing it with a duty to not misrepresent material circumstances in the insurer's enquiries.[492] Recently in *Axa Versicherung*[493] the court stressed it will need to consider what needs to have been said for a fair presentation from the standpoint of a reasonable insurer[494] and the court may have to determine what considerations played in the occurrence of a "hypothetical fair presentation" would have played in the mind of the insurer.[495] Thus, there is little scope for the doctrine of good faith, as arguably the insurer in *Drake* would fail under the current law to show inducement as it would have necessitated a further discussion of previous accidents on the hypothetical fair presentation of the risk.[496]

4. Elucidating the scope for a future role of Section 17

It is noteworthy, that Soyer has suggested that the removal of the avoidance remedy, coupled with the reaffirmation of the underlying nature of insurance contracts requiring utmost good faith could tempt the court to extend the doctrine interstitially to other situations where the remedy can be more nuanced, such as the loss of the right to rely on a limit of liability.[497] Notwithstanding the award of damages raised particular difficulties in *Banque*,[498] the rationale in *Conlon*[499] supports that there is the prospect of such liability in damages if there is a lack of any positive misrepresentation, even if there was an unsatisfied duty to speak.[500]

The case of *Interfoto*[501] demonstrates that the issue of unfair, unreasonable or unexpected terms is largely achieved by the requirements as to incorporations of terms, hence good faith is operating in the sense of process rather than substance, as it is the common law which requires the insurer to take reasonable steps in the pre-contractual stage to specify terms to the attention of the insured, as the statutory reforms and the Consumer Rights Act 2015

[491] ibid; [62]; Kehinde Anifalaje, "*Statutory reform of the doctrine of uberrimae fidei in insurance law: a comparative review*" (2019) 63(2) J.A.L. 251, 265
[492] CIDRA 2012, ss. 2 - 3
[493] *Axa Versicherung AG v Arab Insurance Group* [2017] EWCA Civ 96; [2017] Lloyd's Rep IR 216
[494] ibid [58].
[495] ibid, [60]
[496] Hemsworth (n.18), 161
[497] Soyer & Tettenborn, "*Mapping (Utmost) Good Faith in Insurance Law- Future Conditional?"* (n.36) 636
[498] *Banque Keyser Ullmann SA (*n.35)
[499] *Conlon v Simms* [2006] EWHC 401 (Ch); [2006] 2 All E.R. 1024; [2006] EWCA Civ 1749; [2008] 1 W.L.R. 484
[500] *HIH Casualty & General Insurance Ltd* (n.13), [48] – [49]
[501] *Interfoto Picture Library Ltd* (n.17)

requires transparency and clarity of terms, and as to striking down those terms might create a notable imbalance between the parties to the consumer's detriment.[502]

4.1. Estoppel

It is worth noting, that the potential application of estoppel could limit the relevance of the doctrine of good faith as seen in *Shardley*[503] Rix Lj highlighted in *obiter* that it was not consistent with the duty of utmost good faith as well as the insurer's requirement to present an alternation on previously agreed agreements but failing to draw the terms to the consumer's attention.[504] The recent case of *Ted Baker*[505] illustrates that albeit an insurer is not generally under any duty to warn an insured of the necessity to comply with policy terms,[506] estoppel by acquiescence could arise if a reasonable person alleging the estoppel would expect the other party to act honestly and responsibly, even if a contract is not dependent of being one of "utmost good faith" as it was noted that it might increase the prospect of party's having a duty to speak.[507]

4.2. Waiver by election

On the other hand, unlike the suspensory nature of estoppel,[508] waiver by election requires that the insurer is possessed of the full fact sufficient to make an informed decision about the release of its rights and concerns the active choice evidenced to have been made between conflicting options, and once such decision is communicated, it becomes final.[509] Contrary to estoppel, which is not dependent on the insurer's state of knowledge but upon some reasonable reliance by the insured, hence in the context of a breach waiver by election could be engaged rather than estoppel.[510] Therefore, Hemsworth argued that if an insurer wants to enact his statutory remedies, set out by the reformed acts of CIDRA and the Insurance Act 2015, it appears that the future scope for good faith is limited, as the common law appears to be timid to authorize the implication of a term established on good faith in the setting of

[502] Hemsworth (n.18), 164
[503] *Sherdley v Nordea Life and Pensions SA* [2012] EWCA Civ 88; [2012] 2 All ER (Comm) 725.
[504] ibid, [51]
[505] *Ted Baker Plc v Axa Insurance UK Plc* [2017] EWCA Civ 4097.
[506] *Privy Council in Diab v Regent Insurance Co Ltd (Belize)* [2006] UKPC 29; [2006] 1 Lloyd's Rep IR 779; [2007] 1 WLR 797.
[507] *Starbev GP Ltd v Interbrew Central European Holdings BV* [2014] EWHC 1311 (Comm)
[508] Kammins Ballrooms Co Ltd v Zenith Investments (Torquay) Ltd [1971] AC 850, (Diplock LJ) [882] – [883]
[509] *Hemsworth* (n.18), 167
[510] *Motor Oil Hellas (Corinth) Refineries SA v Shipping Corp of India (The Kanchenjunga)* [1990] 1 Lloyd's Rep. 391, [398] – [399].

a termination of a contract as distinctive from matters arising in its formation and performance.[511]

Conclusion

The development of the insurer's duty of good faith has been hampered due to the grandiloquent nature of section 17, due to being viewed as the source of it, bringing the unfitting remedy of avoidance into action in case of its breach.[512] The reformed Acts have replaced the duty of disclosure with a duty to not misrepresent material circumstances and to make a fair presentation of the risk. Hence the potential future scope of good faith in its pervasive relevance as an interpretative tool has practical limitations as the doctrine of good as section 17 has no independent power to create, modify or restrict expressed contractual terms.[513] The pervasive influence of good faith may have a pragmatic application to inform the content of a term to be implied with the prospect for a remedy in the form of damages in a case of a breach of a term. However, an insurer who decides to rely on a contractual right in an unreasonable fashion[514] might be prevented from doing so by the intervention of common law and equity principles.

[511] *Ilkerler Otomotiv Sanayai ve Ticaret Anonim v Perkins Engines Co Ltd* [2017] EWCA Civ 183; [2017] 4 WLR 144, [29] (Longmore LJ)
[512] Soyer "*The insurer's duty of good faith: is the path now clear for the introduction of new remedies?*" (n.4), 52
[513] *Hemsworth* (n.18), 170
[514] *Massoud v NRMA Insurance* (1995) 62 NSWLR 657.

BIBLIOGRAPHY

Cases

Abu Dhabi National Tanker Co v Product Star Shipping Ltd (the Product Star) No 2 [1993] 1 Lloyd's Rep 397

Alfred McApline plc v BAI (Run-Off Ltd) [2000] 1 Lloyd's Rep. 437

Axa Versicherung AG v Arab Insurance Group [2017] EWCA Civ 96; [2017] Lloyd's Rep IR 216

*Banque Keyser Ullmann SA v Skandia (UK) Insurance Co Ltd [*1990] 1 Q.B. 665 [701] – [702] (Steyn LJ)

Banque Financière de la Cite v. Westgate Insurance Co. Ltd [1990] 1 Q.B. 665; *aff'd* [1991] 2 A.C. 249

Bell v Lever Bros Ltd [1932] AC 161

Bhasin v Hrynew 2014 SCC 71

Braganza v BP Shipping Ltd [2015] UKSC 17; [2015] 1 WLR 1661; [2015] 2 Lloyd's Rep 240

Bristol Groundschool Ltd v Intelligent Data Capture Ltd [2014] EWHC 2145 (Ch.)

Britton v Royal Insurance Co (1866) 4 F & F 905

Brownlie v Campbell (1880) 5 App Cas 925

Carter v Boehm (1766) 3 Burr 1905

Conlon v Simms [2006] EWHC 401 (Ch); [2006] 2 All E.R. 1024; [2006] EWCA Civ 1749; [2008] 1 W.L.R. 484

Cox v Bankside Members Agency [2985] CLC 671, 680, 684

Deutsche Ruckversicherung AG v. Walbrook Insurance Co. Ltd [1995] 1 Lloyd's Re. 153, 164

Drake Insurance Plc v Provident Insurance Plc [2003] EWCA Civ 1834; [2004] 1 Lloyd's Rep 268; [2004] Lloyd's Rep IR 277; [2004] QB 601.

*Emirates Trading Agency LLC v Prime Mineral Exports Pte Ltd [*2014] EWHC 2104 (Comm); Cable *& Wireless v IBM* [2002] EWHC 2059 (Comm).

Federal Mogul Asbestos Personal Injury Trust v Federal-Mogul Ltd (formerly T & N Plc) [2014] EWHC 2002 (Comm); [2014] Lloyd's Rep IR 671

Feise v Parkinson (1812) 4 Taunton 640, 128 ER 482

*First Energy (UK) Ltd. v. Hungarian International Bank Ltd. [*1993] 2 Lloyd's Law Reports 194

Gan Insurance Co Ltd v Tai Ping Insurance Co Ltd (No 2) [2001] EWCA Civ 1047; [2001] Lloyd's Rep IR 667.

Goshawk Dedicated Ltd v Tyser & Co Ltd [2006] EWCA Civ 54; [2007] Lloyd's Rep IR 224.

Hayes v Willoughby [2013] UKSC 17; [2013] 1 WLR 935

HIH Casualty & General Insurance Ltd v Chase Manhattan Bank [2003] UKHL 6; [2003] 2 Lloyd's Rep 61; [2003] 2 Lloyd's Rep IR 230

Ilkerler Otomotiv Sanayai ve Ticaret Anonim v Perkins Engines Co Ltd [2017] EWCA Civ 183; [2017] 4 WLR 144

Joel v Law Union Insurance Co [1908] 2 K.B. 863

K/S Merc-Scandia XXXXII v Certain Lloyd's Underwriters (The Mercandian Continent) [2001] EWCA Civ 1275; [2001] 2 Lloyd's Rep 563; [2001] Lloyd's Rep IR 802

Kammins Ballrooms Co Ltd v Zenith Investments (Torquay) Ltd [1971] AC 850, (Diplock LJ) [882] – [883]

Manifest Shipping Co Ltd v Uni-Polaris Shipping Co Ltd (The Star Sea) [2001] UKHL 1; [2003] 1 AC 469

March Cabaret Club Ltd v London Insurance [1975] 1 Lloyd's Rep. 169

Marks and Spencer Plc v BNP Paribas Securities Services Trust Co (Jersey) Ltd [2015] UKSC 72; [2016] AC 742

Mediterranean Salvage & Towage Ltd v Seamar Trading & Commerce Inc (The Reborn) [2009] EWCA Civ 531; [2009] 2 Lloyd's Rep 639

Motor Oil Hellas (Corinth) Refineries SA v Shipping Corp of India (The Kanchenjunga) [1990] 1 Lloyd's Rep. 391

MSC Mediterranean Shipping Co SA v Cottonex Anstalt [2016] EWCA Civ 789; [2016] 2 Lloyd's Rep 49

Nottingham Patent Brick & Tile Co v Butler (1886) 16 QBD 778 (CA).

Pan Atlantic Insurance plc v Pine Top Insurance [1994] 3 All E.R. 581

Paragon Finance plc v Nash [2001] EWCA Civ 1466; [2002] 1 WLR 685

Phoenix General Insurance Co v Halvanon Insurance Co Ltd [1985] 2 Lloyd's Rep 599

Polhill v Walter (1832) 3 B & Ad 114; 110 ER 43

Privy Council in Diab v Regent Insurance Co Ltd (Belize) [2006] UKPC 29; [2006] 1 Lloyd's Rep IR 779; [2007] 1 WLR 797.

Redgrave v Hurd (1881) 20 Ch D 1

Sherdley v Nordea Life and Pensions SA [2012] EWCA Civ 88; [2012] 2 All ER (Comm) 725.

Société Générale v Geys [2012] UKSC 63; [2013] 1 AC 523

Socimer International Bank Ltd v Standard Bank London Ltd [2008] EWCA Civ 116; [2008] 1 Lloyd's Rep 558

Sprung v Royal Insurance (UK) Ltd [1999] Lloyd's Rep IR 111

Starbev GP Ltd v Interbrew Central European Holdings BV [2014] EWHC 1311 (Comm)

Starbev GP Ltd v Interbrew Central European Holdings BV [2014] EWHC 1311 (Comm)

Ted Baker Plc v Axa Insurance UK Plc [2017] EWCA Civ 4097.

Traill v. Baring (1864) 4 De G.J & S. 318

Versloot Dredging v HDI Gerling (The DC Merwestone) [2016] UKSC 45; [2016] 2 Lloyd's Rep 198; [2016] Lloyd's Rep IR 468; [2017] AC 1

Yam Seng Pte Ltd v International Trade Corp Ltd [2013] EWHC 111 (QB); [2013] 1 Lloyd's Rep 52

Yam Seng Pte Ltd v International Trade Corp Ltd [2013] EWHC 111 (QB); [2013] 1 Lloyd's Rep 526

Interfoto Picture Library Ltd v Stiletto Visual Programmes Ltd [1987] EWCA Civ 6; [1989]

*Midland Bank v. Hett, Stubbs & KempQB 433 [*1939] 1 KB 194

Sharp v Sphere Drake Insurance Co Ltd (The Moonacre) [1992] 2 Lloyd's Rep. 207

Statutes

Consumer Insurance (Disclosure and Representations) Act 2012

Financial Services and Markets Act 2000

Marine Insurance Act 1906

Insurance Act 2015

Insurance Contracts Act 1984

Third Parties (Rights Against Insurers) Act 2010

Books

Soyer B, "*The insurer's duty of good faith: is the path now clear for the introduction of new remedies?*" in Clarke M and Soyer B (eds), *The Insurance Act 2015, A New Regime for Commercial and Marine Insurance Law* (Informa Law from Routledge 2016)

Hjalmarsson J, "*Insurers and the law of fraud - A success story and the case for regulatory intervention*" in D. Rhidian Thomas "*The Modern Law of Marine Insurance*" (1st ed, Informa Law from Routledge 2023)

Gilman J, Merkin R, Blanchard C, Templeman M, "*Arnould's Law of Marine Insurance and Average*", 18th edn (2013)

Merkin R. M., "*Colinvaux's Law of Insurance*" (13th edn, Sweet & Maxwell, 2022)

McKendrick E, "*Good Faith in the Performance of a Contract in English Law*" in Larry DiMatteo and Hogg M, (eds.) "*Comparative Contract Law: British and American Perspectives*" (OUP 2016), 196

Journal Articles

Anifalaje K, "*Statutory reform of the doctrine of uberrimae fidei in insurance law: a comparative review*" (2019) 63(2) J.A.L. 251

Bennet H.N., "*Utmost good faith in the House of Lords*" (1995) 111 L.Q.R. 181

Galanter M, "*Why the 'Haves' Come Out Ahead: Speculations on the Limits of Legal Change*", (1974) 9:1 Law and Society Review

Lewins K, "*Going walkabout with Australian insurance law: the Australian experience of reforming utmost good faith*" (2013) 1 J.B.L. 22

Margaret Hemsworth, '*The fate of "good faith" in insurance contracts* (2018) L.M.C.L.Q. 143

Mason A, "*The Place of Equity and Equitable Remedies in the Contemporary Common Law World*" (1994) 110 LQR 238

Millett P, "*Equity's Place in the Law of Commerce*" (1998) 114 LQR 214

Norma J. Hird, "*Utmost good faith – forward to the past*" (2005) Journal of Business Law 257

Peter MacDonald Eggers, "*Remedies for the failure to observe the utmost good faith*" (2003) LMCLQ 249

Philip Rawlings and John Lowry, "*Insurance Fraud: The 'Convoluted and Confused' State of the Law*" (2016) LQR (132) 119

Rose FD, "*Informational asymmetry and the myth of good faith: back to basics*" (2007) LMCLQ 181, 184

Soyer B & Tettenborn A, "*Mapping (Utmost) Good Faith in Insurance Law- Future Conditional?"* (2016) 132 Law Quarterly Review 618

— — "*Reforming the assured's pre-contractual duty of utmost good faith in insurance contracts for consumers: are the Law Commissions on the right track?"* (2008) 5 J.B.L. 386

Steyn J *'Contract Law: Fulfilling the Reasonable Expectations of Honest Men'* (1997) 113 LQR 433

— — "*The role of good faith and fair dealing in contract law: a hair-shirt philosophy*?" (1991) Denning LJ 131

Tarr J and Tarr, "*Utmost Good Faith in Insurance: Reform Overdue?*" (2003) 10 Asia Pacific Law Review 171

— — "*A growing good faith in contracts*" (2015) 5 J.B.L. 410

Tompkinson D, "*Duration of the duty of utmost good faith*" 5(2) Int. I.L.R. 1997

Xin W and Jun S.H., "*Review of the current hull insurance law of China: in comparison with the UK rules*" (2017) 2 J.B.L. 121

Online Journal Articles

Achampong F, "*Uberrima Fides in English and American Insurance Law: A Comparative Analysis*" (1987) 36 (2) I.C.L.Q. 329, <https://www.jstor.org/stable/759999> accessed 20 May 2023

Chao-Hung C.C., "*Insurance Act 2015 in the UK and Lessons for Singapore*" (2016) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2928418> accessed 23 April 2023

Leggatt G, "*Contractual duties of good faith*" (Lecture to the Commercial Bar Association on 18 October 2016) < https://www.judiciary.uk/wp-content/uploads/2016/10/mr-justice-leggatt-lecture-contractual-duties-of-faith.pdf> accessed 09 May 2023

Hasson R.A.*, "The Basis of the Contract Clause" in Insurance Law"* (1971) 34(1) TMLR 29 <https://www.jstor.org/stable/1094022> accessed 28 April 2023; Richard Aikens, "*The post-contract duty of good faith in insurance contracts: is there a problem that needs a solution?"* (2010) 5 J.B.L. 379

— — "Constitutional Hasson, "*The Doctrine of Uberrima Fides in Insurance Law – A Critical Evaluation*" (1969) 32 (6) TMLR < https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2230.1969.tb01239.x> accessed 29 April 2023

Zhu M, "*The Utmost Good Faith in Marine Insurance: The Nature*" (2020) 11 Beijing Law Review 99, 104 < https://www.scirp.org/iourna I/birResearch> accessed 10 May 2023

Command papers and Law Commission reports

Law Commission and Scottish Law Commission, *Insurance Contract Law: Business Disclosure; Warranties; Insurers' Remedies for Fraudulent Claims; and Late Payment* (Law Com No 353, July 2014)

Law Reform Committee Fifth Report, *Conditions and Exceptions in Insurance Policies* (Cmnd. 62, 1957)

Law Commission of England and Wales, *Insurance Law: Non-Disclosure and Breach of Warranty* (Law Com 104, 1980)

Miscarriages of Justice: Are Pre-Trial Panels The Solution?

Nora Belkhiter

Abstract

Miscarriages of justice have long been a cause for concern in the UK's criminal justice system. Lucy Letby's case, though not a miscarriage of justice, has brought forward some of the debates over the safety of the process, especially where complex scientific evidence is involved. Among the various measures proposed for the prevention of miscarriages of justice, in these types of cases, is that of 'a panel of experts' proposed by Dr Phil Hammond. This article discusses the extent to which such a proposal would successfully stop all miscarriages of justice.

In relation to expert evidence, the introduction of a panel of experts can significantly reduce the risk of miscarriages of justice. Firstly, it can ensure that admissibility criteria is applied appropriately, thereby creating a consistent standard of scrutiny. Secondly, it will equip judges with a comprehensible expert report, enabling them to better direct the jury on the evidence. Thirdly, the report will allow lawyers to correctly examine and cross-examine experts. Finally, the experts can support the jury in its understanding of the evidence and as such improve the likelihood of arriving at the correct decision.

However, a panel of experts for expert scientific evidence would not prevent miscarriages of justice caused by other significant factors such as the police investigative process and eyewitness evidence. As such, this article discusses that a modified panel for the monitoring of police conduct, and eyewitness evidence reliability is likely to have similar positive results in the reduction of miscarriages of justice.

Introduction

The conviction of Lucy Letby[515], a highly publicised case in the media, reignited long standing debates on the appropriate procedures for the admissibility of expert evidence.[516] Though not a proven miscarriage of justice, Letby's case has provoked commentary on

[515] 'Lucy Letby found guilty of baby murders' (CPS, 18 August 2023) <https://www.cps.gov.uk/mersey-cheshire/news/lucy-letby-found-guilty-baby-murders> accessed 4 January 2025.
[516] 'Lucy Letby' (The Guardian, December 2024) <https://www.theguardian.com/uk-news/lucy-letby> accessed 4 January 2025; 'Lucy Letby' (BBC) <https://www.bbc.co.uk/news/topics/cvp59396q42t > accessed 4 January 2025; 'Lucy Letby' (The Telegraph) , <https://www.telegraph.co.uk/lucy-letby/> accessed 4 January 2025.

reforming the criminal justice system to reduce procedural and substantive errors. Dr Phil Hammond proposed the introduction of 'a panel of experts'[517] which would assess complex scientific and statistical evidence, and produce a 'written report'[518] to the Crown Prosecution Service (CPS). This would assist the CPS in their decision to prosecute and would also provide evidence in trial, should the case proceed. This article critically discusses whether the materialisation of this proposal would prevent all miscarriages of justice. It begins by defining the term and introducing the main causes that are discussed throughout (I). The second section advances the argument that a panel of experts would prevent most miscarriages of justice by remedying the issues of over reliance in the trial procedures and actors. It examines other proposed measures such as the Daubert criteria, primers, pre-hearing discussions, accreditation and written summaries, which fail to prevent most miscarriages of justice because they do not address the procedural issues that cause them (II). The third and fourth sections adapt Hammond's idea of a panel of experts and discuss its effect in preventing miscarriages of justice where biggest contributing factors are the police investigative process of suspects and eyewitness testimony (III-IV). Before concluding, it addresses the financial feasibility of these panels, submitting that this should not be the deciding factor of whether the proposal would be implemented (V).

1. Definition of Miscarriages of Justice and The Main Causes

Given that there is no universal definition of miscarriages of justice, it is important to qualify the term. According to Naughton, a case can only be characterised as a miscarriage of justice if and when there is a successful appeal against conviction.[519] Contrarily, a miscarriage of justice under the Criminal Justice Act 1988 occurs only when new facts prove 'beyond reasonable doubt that the person did not commit the offence'.[520] This definition is overly restrictive as it not only imposes the highest burden of proof operated in law, but also only focuses on substantive miscarriages of justice, where the defendant is factually innocent. Furthermore, it excludes successful conviction appeals, which Naughton defines as miscarriages of justice. Easton helpfully elaborates on Naughton's definition, asserting that a miscarriage of justice can refer not only to the 'wrongful conviction of the innocent,…

[517] Phil Hammond, 'My own view is that…' (*X*, 26 July 2024) <https://x.com/drphilhammond/status/1816887965144019452> accessed 20 December 2024.
[518] ibid.
[519] Micheal Naughton, *Rethinking Miscarriages of Justice: Beyond the Tip of the Iceberg* (Palgrave 2007) 17.
[520] Criminal Justice Act 1988, s133(1ZA).

wrongful acquittals of the guilty'[521] but also where the defendant is factually guilty 'but the integrity of the process has been undermined'.[522] Naughton's and Easton's definition complement each other because a successful appeal against conviction can occur where there are procedural improprieties as in *Mullen* where the defendant, although guilty, had been unlawfully removed from Zimbabwe.[523] Thus this article employs both definitions. A miscarriage of justice is a successfully appealed case, regardless of whether the cause was of a substantive or procedural nature.

Beyond defining miscarriages of justice, it is important to note their causes. Expert evidence and testimony features as one of the main reasons in Nurse's list of systemic reasons for modern miscarriages of justice. [524] Thus its discussion is fundamental to answering the question of whether a panel of experts would stop all miscarriages of justice. However, it is not the only cause. Miscarriages of justice are also caused by the police investigation and interview phases which Poyser and Milne viewed as 'significant and enduring'. [525] Failures of police procedure are not only significant but likely appear as a contributing factor into all miscarriages of justice.[526] Finally, eye witness testimony that misidentifies the defendant as the culprit, 'one of the principal causes of miscarriages of justice'.[527] As early as 1976, Lord Devlin took the view that such testimony bears a 'special risk of wrong conviction'[528] and thus the progression of cases solely relying on it should not be allowed to progress. It is obvious that alone Hammond's proposed panel of experts would not stop all miscarriages because it would fail to address the above two fundamental causes of it. As such, in its critical examination of the question this article adapts the proposed panel to consider whether this idea would prevent miscarriages of justice caused by police investigations and witness testimony.

[521] Susan Easton, *Silence and Confessions: The Suspect as the Source of Evidence* (Palgrave Macmillan 2014) 129.
[522] ibid 129.
[523] *R v Mullen* [2000] Q.B. 520, 535.
[524] Sam Poyser, Angus Nurse and Rebecca Milne, *Miscarriages of Justice: Causes, Consequences and Remedies* (Policy Press 2018) as referenced in Jon Robins, 'Miscarriages of Justice in England and Wales' in Jon Robins (ed), Murder, Wrongful Conviction and the Law (Routledge, 2023).
[525] Sam Poyser and Rebecca Milne, 'The time in between a case of 'wrongful' and 'rightful' conviction in the UK: Miscarriages of justice and the contribution of psychology to reforming the police investigative process' (2021) 23 IJPSM 5, 5.
[526] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 5.
[527] Graham Davies and Laurence Griffiths, 'Eyewitness Identification and the English Courts: A Century of Trial and Error' (2008) 15 Psychiatry Psychol. Law 435.
[528] Lord Devlin, *Report to the Secretary of State for the Home Departments of the Departmental Committee on Evidence of Identification in Criminal Cases* (London Her Majesty's Stationery Office, 26 April 1976) 149.

2. Expert Opinion Evidence

2.1. Admissibility And Reliability

In England and Wales, opinion evidence is not admissible, with the exception of expert opinion evidence. However, it will only be admitted if it is 'relevant'[529] to issues raised in the process, if it falls 'outside the court's own knowledge and experience'[530], the expert is 'competent'[531] and the opinion is 'sufficiently reliable'.[532] Experts must give opinion that is 'objective and unbiased'[533] and is on a topic within their 'expertise'.[534] In aid of establishing reliability, 'the court may'[535] consider a range of factors, including 'the quality of data'[536] relied upon, the 'validity of the methodology'[537], and whether the evidence falls within 'the expert's own field of expertise'[538], among others.

Scientific evidence often 'is either unreliable or of unknown reliability'[539], something which the criteria does not adequately scrutinise because unreliable evidence has been admitted into trial and presented with utmost confidence. For example, in *Dallagher*[540] an expert supported that he was 'absolutely convinced'[541] the ear print was the defendant's and in *Cannings*[542] the statistical evidence was 'confidently presented to the jury as virtually overwhelming'[543] both of which led to the wrongful conviction of the defendants. A panel of experts, though requiring a more radical change in the criminal justice system, would have the sole purpose of assessing the evidence and reporting on the likelihood of conviction based on it. Therefore, even cases which do not involve complex scientific evidence such as ear print evidence, would benefit from an individual assessment of the evidence in the context of that case, reducing the risk for such miscarriages.

[529] Criminal Practice Directions 2023, part 7.1.1(a).
[530] ibid part 7.1.1(b).
[531] ibid part 7.1.1(c).
[532] ibid part 7.1.1(d).
[533] Criminal Procedure Rules 2020, SI 2020/759 part 19.2.(1)(a)(i).
[534] ibid part 19.2.(1)(a)(ii).
[535] Criminal Practice Directions 2023, part 7.1.2.
[536] ibid part 7.1.2(a).
[537] ibid part 7.1.2(b).
[538] ibid part 7.1.2(f).
[539] Gary Edmond and Mehera San Roque, 'The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial' (2012) 24 CICJ 51, 52.
[540] *R v Dallagher* [2002] EWCA Crim 1903, [2005] 1 Cr App R 12.
[541] ibid [32].
[542] *R v Cannings* [2004] EWCA Crim.
[543] ibid [156].

2.2. Accreditation

However, a less radical measure of compulsory accreditation for raising the standard of reliability has been supported by the House of Commons Science and Technology Committee and Lord Auld.[544] Hartshorne and Miola considered this, finding that the existing accreditation councils impose stringent criteria and scrutiny on experts, with 'a rigorous selection process'.[545] They thus supported a move towards making accreditation compulsory because in their view, experts who are regulated by codes of practice, breach of which leads to expulsion, would 'be approaching their task with greater, objectivity, integrity and professionalism'.[546] Therefore, it also has the potential to raise the standard in all scientific fields. However, it is questionable whether accreditation, alone, can raise the standard of evidence.[547] For example, Roy Meadows, who gave evidence in the murder trials of *Clark*[548] and *Cannings*[549], would have no issue gaining accreditation due to his 'experience and profile'.[550] Consequently, any sanctions would 'be a case of closing the stable door after the horse had bolted'[551] because they would come after his evidence, contributing to the conviction of innocent people. Instead, accreditation could be used in conjunction with the panel of experts which would prevent such miscarriages by assessing the reliability of the evidentiary techniques used. This is not to say that experience is not irrelevant but as convincingly argued by Edmond and Roque it is only valuable when the techniques applied are themselves 'demonstrably reliable'[552], something which a panel of experts focusing on assessing the scientific evidence, and not the expert, would ensure.

2.3. Daubert Criteria

Given that a panel of experts would assess evidence before trial, but would not change the current reliability criteria, the Daubert Criteria applied can be considered a better alternative.[553] The five main criteria include assessing whether the evidence 'can be (and has

[544] Lord Auld, *Review of the Criminal Courts of England and Wales* (2001) 573 para [131]; House of Commons Science and Technology Committee, 'Forensic Science on Trial' (HC 96-1 2005).
[545] John Hartshorne and Jose Miola, 'Expert Evidence: Difficulties and Solutions in Prosecutions for Infant Harm' (2010) 30(2) Legal Studies 279, 282.
[546] ibid 282.
[547] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011) para 5.104-5.105.
[548] R v Clark [2003] EWCA Crim 1020.
[549] Cannings (n28).
[550] John Hartshorne and Jose Miola, 'Expert Evidence: Difficulties and Solutions in Prosecutions for Infant Harm' (2010) 30(2) Legal Studies 279, 282
[551] ibid 282.
[552] Edmond and Roque (n25) 55.
[553] Daubert v Merrell Dow Pharmaceuticals [1993] 509 US 579.

been) tested'[554], whether it 'has been subjected to peer review and publication'[555], whether there is a 'known or potential rate of error'[556], whether there is 'the existence and maintenance of standards controlling its operation'[557] and whether there is 'widespread acceptance'[558] for it. This provides clear expectations for the evidence the experts submit and is a pre-trial process, so the screening of evidence happens before it reaches the jury. Similar criteria now exist in the Criminal Practice Directions.[559] However, they are also exposed to judicial discretion for their application which has 'routinely misinterpreted and broadened the reach of Daubert'.[560] Their focus on testability and peer review has also excluded 'novel'[561] but sound cases, thereby failing to prevent miscarriages of justice. Furthermore, similar to what the panel of experts would offer, Edmond and Roque support the idea of independent pre-trial scrutiny of the reliability of evidence.[562] In consulting the Law Commission on evidence, the General Medical Council also opined that 'a robust assessment of its admissibility prior to the trial is critical'.[563] These statements suggest that there is a common view that the best scrutiny of expert evidence takes place before the trial. This in turn implies support for a proposal such as Hammond's and supports the contention that it would be effective in significantly preventing miscarriages of justice caused by expert evidence.

2.4. Reliance on Judges and The Idea of Primers

Even so, the admissibility of expert evidence falls on judges, as principal gatekeepers of expert opinion evidence. They are also responsible for 'summing-up'[564] the case and evidence at the end of the trial. There is therefore, an inherent expectation that the judge will identify the limitations of such evidence and exclude the unreliable and irrelevant evidence.[565] Far from this, the norm is that judges will opt to admit expert evidence and leave

[554] ibid 593.
[555] ibid 593.
[556] ibid 594.
[557] ibid 580.
[558] ibid 594.
[559] Criminal Practice Directions 2023, part 7.1.2.
[560] Tellus Institute, 'Daubert: the most influential supreme court ruling you've never heard of' (Project on Scientific Knowledge and Public Policy, 2003) 3.
[561] Jane Ireland and John Beaumont, 'Admitting scientific expert evidence in the UK: reliability challenges and the need for revised criteria – proposing an Abridged Daubert' (2015) 17 Journal of Forensic Practice 3, 7.
[562] Edmond and Roque (n25) 54.
[563] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011) para 1.16.
[564] Lord Auld, *Review of the Criminal Courts of England and Wales* (2001) 573, 519 [17].
[565] Edmond and Roque (n25) 54.

the assessment of how much 'weight'[566] should be placed on it, to the jury. Limitations in the exercise of their roles, have also been identified in the guidance they provide to jurors.[567] This proposition was supported by Edmond et al in their extensive analysis of 300 cases involving 'latent fingerprint evidence'.[568] They found that the majority of the judges misrepresented the value of the evidence, taking it as 'definitive'[569], whilst the minority that summarised the evidence more modestly did so under the impression that the words they were using were synonymous to definitive.[570] These issues appear in Barry George's case. [571] The Court of Appeal was critical of the judge's summary of the firearm discharge residue evidence because he failed to present it as 'neutral'[572] but instead 'as supporting the proposition'[573] that George was the murderer.

A panel of experts would aid the judge in the admissibility and explanation of expert evidence. First, the report would be provided to judges when deciding on the admissibility of evidence, highlighting its problematic areas. Second, upon admission the report would aid the judge in properly summarising the evidence's impact. For example, had there been a report in George's case it would have highlighted the neutrality of the firearm discharge residue evidence, allowing the trial judge to either exclude the evidence or accurately summarise its effect. However, it could be argued that such a report would have the same limited effect as the use of a 'series of 'primer' documents'[574], like those published by the Royal Society to assist in understanding the 'limitations' of evidence presented.[575] However, a panel of experts would not be, as primers are, limited to analysing areas of scientific 'consensus'[576]. Quite the opposite, it would seek to fill the gap of misunderstanding complex evidence as it appears in each case, reducing the significant risks, as Edmond et al demonstrated, of judges misinterpreting the evidence.

---

[566] ibid 54.
[567] Eva Ribbers and Marika Linnéa Henneberg, 'Evidentiary Instructions Improve Mock Juror Assessment of Feature-Comparison Evidence' (2018) 22 Int'l J. Evidence & Proof 262, 263-264.
[568] Gary Edmond, Emma Cunliffe, and David Hamer, 'Fingerprint Comparison and Adversarialism: The Scientific and Historical Evidence' (2020) 83 MLR 1287, 1288.
[569] ibid 1290.
[570] ibid 1292.
[571] *R v George (Barry)* [2007] EWCA Crim 2722.
[572] ibid [51].
[573] ibid [51].
[574] Sir Brian Leveson, *Review of Efficiency in Criminal Proceedings* (Judiciary of England and Wales January 2015) 63 [237].
[575] 'Courtroom science primers launched today' (The Royal Society, 22 November 2017) <https://royalsociety.org/news/2017/11/royal-society-launches-courtroom-science-primers/> accessed 2 January 2025.
[576] Leveson (n 60)

2.5. Reliance on Lawyers and Cross Examination, and The Idea of Expert Discussions

Adversarial systems rely on trial procedures such as cross examination to scrutinise expert evidence upon admission. Scholarly discussion supports that these procedures overly rely on lawyers to possess the knowledge and ability to effectively challenge expert evidence.[577] The Law Commission adopted the same 'doubtful'[578] view in its examination of the ability of cross-examination to effectively question expert evidence. This results from lawyers placing too much emphasis on the 'experts' credibility'[579] and potential conflict of interest[580], instead of challenging the 'fundamental methodological and interpretive issues'[581] as well as the 'validity and reliability'[582] of the evidence. Substandard scrutiny is also a product of the defence not calling expert witnesses, often attributable to the inequality of resources between the prosecution and the defence.[583] Even where experts are consulted, the defence may choose not to call them for evidence in court, as in Lucy Letby's case.[584] Though this case is not characterised as a miscarriage of justice, there is an obvious risk of it where the jury only hears evidence from the prosecution.

A panel of experts poses a viable solution to the discussed issues. It would supply the defence lawyers with a report that analyses the techniques and methods equipped by the expert, giving them enough detail to enable them to direct their cross-examination to the limitations of the techniques. Furthermore, it would balance the resource inequalities by independently scrutinising evidence produced by the prosecution. However, the extent to which this would stop all miscarriages of justice caused by expert evidence can be challenged by reference to current measures for pre-trial hearings between the experts.[585] There, experts discuss what they intend to raise in their evidence, identifying the extent of their 'agreement'[586] and 'disagreement'.[587] The aim is 'to agree and narrow issues'[588] as much as possible, which can improve the reliability of the evidence presented at trial. This is similar enough to what the

[577] Edmond and Roque (n25) 55-56; Paul Roberts, 'Expert Evidence' in Paul Roberts and Adrian Zuckerman (eds), *Roberts & Zuckerman's Criminal Evidence* (3rd edn, OUP 2022) 532.
[578] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011) para 1.20.
[579] Ibid para 1.21.
[580] *R v Letby (Lucy)* [2024] EWCA Crim 748 [102].
[581] Edmond and Roque (n25) 56.
[582] Edmond et al (n54) 1321.
[583] Ribbers (n53) 263. See also R v Smith [2011] EWCA Crim 1296 [11].
[584] *R v Letby (Lucy)* [2024] EWCA Crim 748 [5].
[585] Criminal Practice Directions 2023, part 7.3; Criminal Procedure Rules 2020, part 19.6.
[586] ibid part 7.3.2(a).
[587] ibid part 7.3.2(b).
[588] ibid part 7.3.2.

panel of experts would be doing and yet has not proven to prevent miscarriages. Furthermore, these discussions are not designed with the aim of complete agreement and where the evidence is complex there is no guarantee that the panel of experts would come to an agreement either. If that is the case, one could argue the report would be of no use to lawyers in the cross-examination phase as it would not provide an unambiguous interpretation of the evidence and thus would not help to rebut expert testimony where appropriate.

However, the panel of experts can be differentiated in two ways. First, it would not rely on the judge to direct a hearing, which is what is arguably currently causing their use to be 'the exception rather than the rule'.[589] Leveson and Lord Auld have recommended a more frequent use of these discussions due to their helpfulness.[590] Second, given that the discussions happen between prosecution and defence witnesses, it is reasonable to assume that they are incumbent on the defence calling witnesses. This leaves cases where the defence has not done so, unprotected from the level of screening these hearings provide. A panel of experts would assess the evidence regardless of whether the defence is planning on bringing experts. As a result, the consistent evaluation of the evidence strengthens the impact on preventing miscarriages of justice. Even where experts may disagree as in Letby's case, their report can reflect on the areas of disagreement, providing an aid to lawyers to better scrutinise the evidence during trial and successfully stop miscarriages of justice by aiding the jury in their final decision.

2.6. Reliance on Jury and The Aid of Written Evidential Summaries

Partly as a result of poor trial scrutiny of expert evidence, jurors are not equipped to critically analyse it and often accept unreliable evidence.[591] Academic discourse describes jurors as overly deferential due to experts representing 'epistemic authority'[592] which makes jurors take their evidence 'at face value'[593], especially considering the existence of criteria for expert evidence.[594] Primary mock jury research has also demonstrated this, particularly in relation to feature comparison evidence which has appeared in various miscarriages of

[589] Lord Auld, *Review of the Criminal Courts of England and Wales* (2001) 514, 579 [145].
[590] Leveson (n60) 63 [236].
[591] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011) para 1.20.
[592] Paul Roberts, 'Expert Evidence' in Paul Roberts and Adrian Zuckerman (eds), *Roberts & Zuckerman's Criminal Evidence* (3rd edn, OUP 2022) 535.
[593] Ribbers (n53) 266. See also Edmond and Roque (n25) 58.
[594] Ribbers (n53) 266.

justice like *Barry George*[595] and *Dallagher*.[596] A panel of experts would prevent these miscarriages of justice by focusing its scrutiny on the evidence and not the expert, providing a helpful report for the jury's benefit. In their study which simulated the aforementioned cases, Ribbers and Henneberg found that once an evidentiary instruction was provided to jurors, they became less deferential to the expert evidence, demonstrating confidence to reject expert evidence.[597]

Jury deference can also be explained by reference to their limited ability to comprehend and challenge the reliability of complicated expert evidence.[598] This is supported by Thomas' study of jurors' ability to comprehend legal instructions which revealed that 'only 31%'[599] correctly identified the judge's legal directions. However, it also revealed that where written instructions are provided to the jury, comprehension increases and the accuracy with which jurors are able to state their understanding also improves.[600] This leads to a reduced risk of the jury convicting the innocent. It can be asserted that a panel of experts is not required for this as the evidentiary instruction can come from the judge. However, given the limitations of judges' themselves in understanding expert evidence it is submitted that better results would be achieved by a panel of experts. Ribbers and Henneberg also recognised that these instructions should be 'controlled by an independent party, in cooperation with the scientific community'.[601] This is significant support for the conclusion that a panel of experts would produce the most effective evidentiary instructions which would help jury comprehension and subsequently reduce miscarriages of justice.

3. Police Investigation of Suspects

Having established that a panel of experts would reduce miscarriages of justice where scientific evidence is the main cause, the discussion necessarily turns to whether a modified panel would have the same effect where the main cause is police procedure. A main issue associated with investigations is that often the police believe that the person they arrested is

[595] *R v George (Barry)* [2007] EWCA Crim 2722.
[596] Ribbers (n53).
[597] Ribbers (n53) 280.
[598] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011) para 1.15.
[599] Cheryl Thomas, *Are Juries Fair?* (Ministry Of Justice Research Series 1/10, February 2010) 36.
[600] ibid 38-39.
[601] Ribbers (n53) 281.

the perpetrator.[602] This belief creates a 'confirmation bias'[603] whereby the rest of the investigation process is driven by a 'tunnel vision'[604] focus on finding evidence that confirms their initial judgement. Consequently, they may rely heavily on information that is misleading and end up pursuing only one suspect, disregarding other potential ones.[605] As a result any other information that appears to be inconsistent with the suspect's guilt is 'overlooked or dismissed as irrelevant, incredible or unreliable'.[606] This is best demonstrated by the case of Colin Stagg[607] who was wrongfully accused for the murder of Rachel Nickell. The real murderer, Napper, had come to police attention 'at least seven times'[608] and there were hints about his involvement that should have been pursued further.[609] The police relied upon the description of the killer not being tall when in fact Napper was.[610] A modified panel made up of experienced investigators would significantly reduce miscarriages of justice caused by this failure of procedure. It would examine the process that the police followed in identifying the suspect, reporting on missed clues and lines of inquiry before the CPS decides on whether to pursue the case.

Issues in the investigative process appear in interviews as well. Responding to calls for improvement, 'the PEACE model'[611] of interviewing was introduced, with 'over 120,000 officers'[612] being trained in this information gathering approach. This was determined to have 'improved the quality of suspect interviews'[613] by reducing 'leading questions'[614],

[602] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 9
[603] ibid 10.
[604] ibid 10.
[605] Alec Samuels, 'The Rachel Nickell case: reflections on the significance' (2012) 52(3) Medicine, Science and the Law 181, 18-182.
[606] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) Findley and Scott, 2006: 293.
[607] Alec Samuels, 'The Rachel Nickell case: reflections on the significance' (2012) 52(3) Medicine, Science and the Law 181.
[608] Sandra Laville, 'Nickell Case: Missed Clues that Allowed Napper to Kill Again' The Guardian (London, 18 December 2008) <https://www.theguardian.com/uk/2008/dec/18/robert-napper-clues> accessed 23 December 2024.
[609] Alec Samuels, 'The Rachel Nickell case: reflections on the significance' (2012) 52(3) Medicine, Science and the Law 181, 18.
[610] ibid 181.
[611] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 14.
[612] ibid 14.
[613] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 14 Clarke and Milne 2001.
[614] Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 14.

encouraging police interviewers to 'assume nothing'[615], 'believe nothing'[616] and 'challenge everything'[617]. Unfortunately, Sanders found that the 'listening skills…remained poor', with 10 per cent of interviews studied containing possible breaches of the Police and Criminal Evidence Act[618] (PACE).[619] Furthermore, despite the efforts for improvement, interviews still often have 'a confession focus'[620] leading to a heightened risk of miscarriages of justice. Consequently, there appears to be a need for further reform, to enhance the effectiveness of the interview stage of the investigative process. The modified panel would not prevent any misconduct as it would review the process after the fact. However, the need for its introduction can still be justified if considered in its procedural benefits sense. Firstly, it would act as a check on whether suspect interviews have been carried out in accordance with PACE.[621] Its role would not be to strike out cases, but it would indicate to the CPS whether there are any breaches that could risk a miscarriage of justice. Secondly, its reports could be used a tool for raising standards by reflecting on the panel's comments. It is however conceded that this approach would not be the most effective alone and it should be accompanied with other reforms in raising the standard of interviews.[622]

4. Eyewitnesses

Finally, it is worth considering how a modified panel would affect the miscarriages of justice caused by eyewitness testimony. Eyewitness testimony is governed by the *Turnbull*[623] guidelines which distinguish between 'good'[624] and 'poor'[625], cautioning the admittance of the latter. The judge can exclude eyewitness evidence where it would adversely affect the fairness of proceedings.[626] This area has also benefitted from PACE[627] which governs the process of identification of suspects. The importance of adhering to the Code was

[615] Forensic Interview Solutions, 'The Science of Interviewing: P.E.A.C.E. A Different Approach to Investigative Interviewing' (FIS) < https://www.fis-international.com/assets/Uploads/resources/PEACE-A-Different-Approach.pdf> accessed 29 December 2024, 6.
[616] ibid, 6.
[617] ibid, 6.
[618] 1984.
[619]As referenced in Sam Poyser and John Grieve, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), *Investigation: Psychology into Practice* (Routledge 2018) 14.
[620] ibid 15-16.
[621] PACE (n104).
[622] A panel of experts would be most effective if used in conjunction with better and consistent police training in the PEACE model of interviews and in recording details of their investigative process.
[623] *R v Turnbull* [1977] QB 224.
[624] ibid [229].
[625] ibid [229]-[230].
[626] PACE (n104) s 78.
[627] PACE (n104); Code of Practice D (2023).

highlighted in *Forbes*.[628] Similar criticisms to expert evidence apply here, as Davies and Griffiths found that judges have not given them 'due weight' and failed to use them until requested by the defence.[629] Though not with a modified panel in mind, Roberts asserted that the most effective way to challenge these risks is for 'pre-trial procedures'.[630] It is submitted that the most effective pre-trial procedure is a modified panel. It would prevent miscarriages of justice by assessing the witness evidence, proposing pursuance or not of the case and drafting a report on the circumstances of the eyewitness testimony and the process with which this was elicited.

However, there are two points to consider. First, the introduction of a modified panel could inappropriately intervene with the jury's role in assessing the evidence. This is because the assessment of such evidence is within the jury's common knowledge and therefore outside the boundaries of opinion evidence. Second, even if the report aids the judge in properly applying the admissibility criteria, eyewitnesses can be completely convinced that what they saw is what actually happened, thus implicating the innocent. For example, the first of the witnesses positively identified George on her own[631] and so did one of the witnesses in Victor Nealon's case.[632] One consideration therefore, is that a modified panel would not be able to prevent miscarriages of justice where witnesses are mistakenly convinced they have identified the perpetrator. However, it would remedy the inconsistent application of the guidelines.[633] Its report would also clearly set out the circumstances under which identification was made, making the limitations of eyewitness testimony transparent. For example, in George's case[634], it would have highlighted that only one of sixteen witnesses identified him, seventeen months after the incident.[635] It would also highlight that she estimated to have seen him for 'six seconds'[636]. Most importantly, it would state that the two other witnesses who identified him did so after sharing a car with the first witness who discussed who she selected and her statement to the police.[637] It would have perhaps

[628] *R v Forbes* [2001] 1 AC 473.
[629] Graham Davies and Laurence Griffiths, 'Eyewitness Identification and the English Courts: A Century of Trial and Error' (2008) 15 Psychiatry Psychol. Law 435, 440.
[630] Andrew Roberts, 'The Problem of Mistaken Identification: Some Observations on Process' (2004) 8 E & P 100, 100.
[631] *R v George (Barry)* [2007] EWCA Crim 2722; (n115) p 441.
[632] *R v Nealon (Victor)* [2014] EWCA Crim 574 [33].
[633] Graham Davies and Laurence Griffiths, 'Eyewitness Identification and the English Courts: A Century of Trial and Error' (2008) 15 Psychiatry Psychol. Law 435.
[634] *R v George (Barry)* [2007] EWCA Crim 2722.
[635] Graham Davies and Laurence Griffiths, 'Eyewitness Identification and the English Courts: A Century of Trial and Error' (2008) 15 Psychiatry Psychol. Law 435, 441.
[636] ibid 441.
[637] ibid 441.

prevented the testimony of the two later witnesses and if not it would have clearly stated the process by which they identified George.[638] It would also have stated that the two witnesses who had allegedly seen the suspect leave Dando's house, failed to identify Barry George.[639] The panel would act in similar fashion in relation to Sam Hallam's case, by emphasising that neither of the two witnesses mentioned him in their initial interviews and only did so when they were 'prompted[640] by others. By highlighting the process by which eyewitness testimony arose, the modified panel would not be assuming the jury's role. Instead, it would aid the jury in making the appropriate inferences about the testimony, thus reducing the number of miscarriages of justice.

5. Practical Feasibility

Before concluding, it is important to address the practical feasibility of these reforms. It is beyond the scope of this article to calculate their financial costs and benefits as there are no official estimations of how much it would cost and save. However, the Law Commission's report of expert evidence and the government's subsequent response is of use. The Law Commission's proposals were more moderate in that they referred to the codification of the reliability test in statute and called for more frequent pre-trial discussions.[641] The government's response referred to the uncertainty 'as to the offsetting of savings which might be achieved'[642] and due to 'resource constraints'[643] declined 'to implement the proposals in full'.[644] It is therefore reasonable to assume that a proposal for a much more fundamental change in the criminal justice process would not be well received due to its potential cost. However, having examined examples of miscarriages of justice, the non-financial benefits of preventing the conviction of innocent people or the conviction of the guilty by violating due process right, should take precedent.

Conclusion

The proposal for a panel of experts is a novel idea which this article has submitted can significantly reduce miscarriages of justice where the main cause is expert opinion evidence.

---

[638] ibid 442.
[639] ibid 441.
[640] *R v Hallam (Sam)* [2012] EWCA Crim 1158 [77].
[641] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales* (Law Com No 325, 2011).
[642] Ministry of Justice, 'The Government's response to the Law Commission report: "Expert evidence in criminal proceedings in England and Wales" (Law Com No 325)' (MOJ 21 November 2013) para 3
[643] ibid para 3.
[644] ibid para 3.

It addresses the procedural issues linked with relying on individual actors to successfully and consistently implement vague standards, scrutinise expert evidence in trial and have enough knowledge and confidence to reject it. Other solutions have been considered such as compulsory accreditation, the Daubert criteria, the use of primers, pre-trial discussions and written summaries. However, these fail to address the need for robust pre-trial assessments of the expert evidence. It has been conceded that a panel of experts, in the form it was proposed by Hammond, would not serve a solution for all miscarriages of justice. However, when applying a modified version for reviewing police investigative processes and eyewitness testimony it has been submitted that similar results would be achieved. Consequently, this article contends that a panel of experts alone will not solve all miscarriages of justice but together with the modified panel, has promising potential in enhancing proper scrutiny of evidence and procedure, consequently preventing miscarriages of justice.

BIBLIOGRAPHY

--‘Courtroom science primers launched today’ (The Royal Society, 22 November 2017) <https://royalsociety.org/news/2017/11/royal-society-launches-courtroom-science-primers/> accessed 2 January 2025

--‘Lucy Letby’ (BBC) <https://www.bbc.co.uk/news/topics/cvp59396q42t > accessed 4 January 2025

-- ‘Lucy Letby found guilty of baby murders’ (CPS, 18 August 2023) <https://www.cps.gov.uk/mersey-cheshire/news/lucy-letby-found-guilty-baby-murders> accessed 4 January 2025

--‘Lucy Letby’ (The Guardian, December 2024) <https://www.theguardian.com/uk-news/lucy-letby> accessed 4 January 2025

--‘Lucy Letby’ (The Telegraph), <https://www.telegraph.co.uk/lucy-letby/> accessed 4 January 2025.

Auld, Review of the Criminal Courts of England and Wales (2001) 573 para [131]; House of Commons Science and Technology Committee Forensic Science on Trial Seventh Report of the Session 2004-05, HC 96-1 2005.

Davies G and Griffiths L, ‘Eyewitness Identification and the English Courts: A Century of Trial and Error’ (2008) 15 Psychiatry Psychol. Law 435

Devlin, Report to the Secretary of State for the Home Departments of the Departmental Committee on Evidence of Identification in Criminal Cases (London Her Majesty’s Stationery Office, 26 April 1976)

Easton S, Silence and Confessions: The Suspect as the Source of Evidence (Palgrave Macmillan 2014) 129

Edmond G, Cunliffe E, and Hamer D, ‘Fingerprint Comparison and Adversarialism: The Scientific and Historical Evidence’ (2020) 83 MLR 1287

Edmond G and Roque M.S, ‘The Cool Crucible: Forensic Science and the Frailty of the Criminal Trial’ (2012) 24(1) CICJ 51

Forensic Interview Solutions, 'The Science of Interviewing: P.E.A.C.E. A Different Approach to Investigative Interviewing' (FIS) < https://www.fis-international.com/assets/Uploads/resources/PEACE-A-Different-Approach.pdf> accessed 29 December 2024

Hammond P, 'My own view is that…' (X, 26 July 2024) <https://x.com/drphilhammond/status/1816887965144019452> accessed 20 December 2024.

Naughton M, Rethinking Miscarriages of Justice: Beyond the Tip of the Iceberg (Palgrave 2007) 17

Hartshorne J and Miola J, 'Expert Evidence: Difficulties and Solutions in Prosecutions for Infant Harm' (2010) 30(2) Legal Studies 279, 282

House of Commons Science and Technology Committee, 'Forensic Science on Trial' (HC 96-1 2005).

Ireland J and Beaumont J, 'Admitting scientific expert evidence in the UK: reliability challenges and the need for revised criteria – proposing an Abridged Daubert' (2015) 17 Journal of Forensic Practice 3, 7

Laville S, 'Nickell Case: Missed Clues that Allowed Napper to Kill Again' The Guardian (London, 18 December 2008) https://www.theguardian.com/uk/2008/dec/18/robert-napper-clues accessed 23 December 2024

Law Commission, Expert Evidence in Criminal Proceedings in England and Wales (Law Com No 325, 2011)

Leveson B, Review of Efficiency in Criminal Proceedings (Judiciary of England and Wales January 2015)

Ministry of Justice, 'The Government's response to the Law Commission report: "Expert evidence in criminal proceedings in England and Wales" (Law Com No 325)' (MOJ 21 November 2013) para 3

Poyser S and Grieve J, 'Miscarriages of Justice: What Can We Learn?' in Rebecca Milne and Andy Griffiths (eds), Investigation: Psychology into Practice (Routledge 2018)

Poyser S and Milne R, 'The time in between a case of 'wrongful' and 'rightful' conviction in the UK: Miscarriages of justice and the contribution of psychology to reforming the police investigative process' (2021) 23 IJPSM 5

Poyser S, Nurse A and Milne R, Miscarriages of Justice: Causes, Consequences and Remedies (Policy Press 2018)

Ribbers E and Henneberg M.L, 'Evidentiary Instructions Improve Mock Juror Assessment of Feature-Comparison Evidence' (2018) 22 Int'l J. Evidence & Proof 262

Roberts A, 'The Problem of Mistaken Identification: Some Observations on Process' (2004) 8 E & P 100

Roberts P, 'Expert Evidence' in Paul Roberts and Adrian Zuckerman (eds), Roberts & Zuckerman's Criminal Evidence (3rd edn, OUP 2022)

Robins J, 'Miscarriages of Justice in England and Wales' in Jon Robins (ed), Murder, Wrongful Conviction and the Law (Routledge, 2023)

Samuels A, 'The Rachel Nickell case: reflections on the significance' (2012) 52(3) Medicine, Science and the Law 181

Tellus Institute, 'Daubert: the most influential supreme court ruling you've never heard of' (Project on Scientific Knowledge and Public Policy, 2003)

Thomas C, Are Juries Fair? (Ministry Of Justice Research Series 1/10, February 2010)

Table Of Cases

<u>UK Cases</u>

R v Cannings [2004] EWCA Crim 1

R v Clark [2003] EWCA Crim 1020

R v Forbes [2001] 1 AC 473

R v George (Barry) [2007] EWCA Crim 2722

R v Hallam (Sam) [2012] EWCA Crim 1158

R v Harris and others [2005] EWCA crim 1980

R v Letby [2024] EWCA Crim 748 [5]

R v Mullen [2000] Q.B. 520, 535

R v Nealon (Victor) [2014] EWCA Crim 574 [33]

R v Smith [2011] EWCA Crim 1296

R v Turnbull [1977] QB 224

US Cases

Daubert v Merrell Dow Pharmaceuticals [1993] 509 US 579, 594

Table Of Legislation

Criminal Justice Act 1988

Police and Criminal Evidence Act 1984,

Criminal Practice Directions 2023

Criminal Procedure Rules 2020, SI 2020/759

How Has The Commission Adapted Article 102 Of The TFEU To Abuse of Dominance in The Digital Markets?

Petros Rouvas

Abstract

The Commission is responsible for the protection of the internal market, ensuring fair competition amongst undertakings. The driving idea behind the Commissions objective of protecting competition within the internal market is the improvement of consumer welfare, which can only be attained through fair competition between undertakings. Although dominance within the market is not by itself considered harmful to consumer welfare, the Commission utilizes the powers given to it under Article 102 the Treaty for the Functioning of the European Union to detect and punish such undertakings from abusing there dominant position. With the growth of the internet and the rise of digital markets the Commission is called upon to ensure fair competition and prevent abusive conduct in a new and unique market. This article reviews the methods used by the Commission in defining the relevant market, establishes the dominance of the undertaking and determines where there has been abusive conduct by the dominant undertaking in the digital market. Due to the differences between traditional markets and digital economy, the Commission has been called upon to adapt the methods it uses to better suit the task at hand. As seen in the article the Commission has responded to this call with varying degrees of success. The analysis within the article is based on a review of recent decisions by the Commission and the judgement by the European Court of Justice relating to conduct of undertakings within the digital market. The Commission has shown to be unwilling to include the unique aspects of the digital market in defining the market, preferring more traditional methods, which led to a smaller market definition and an easier task of establishing dominance. In establishing the dominance of the undertaking in the market and finding abusive conduct, the Commission has proved to be more adaptable, adapting traditional methods to the digital market and developing new tools were necessary.

Introduction

The digital market is defined by its dynamic nature which can create or change markets through the creation of digital platforms.[645] Despite its dynamic nature the digital market still falls under the EU anti-competitive laws in order to ensure that the market remains competitive and prevent any abuse of power.[646] This is done through Article 102 which prohibits undertakings that are dominant or hold a substantial part of the internal market from abusing their position.[647] Dominance is defined in the case law as a position within the market that allows the undertaking to act independently from other competitors without the normal competitive restraints.[648] This independence is normally expressed in the case law as the ability to maintain high prices without losing customers.[649] Dominance does not mean that there is no competition within the market but that the dominant undertaking can influence the competition to the extent that it does not go against its interest within the market.[650] This definition of dominance is reflected within the Guidance given by the Commission on the application of the Article 102.[651] The Commission takes a two stage approach to establishing the undertakings in question dominance, where it first defines the market and then assesses the undertakings position within the market in order to analyse the competitive restraints on the undertaking.[652] However as will be explained below the characteristics of the digital market and the platform that make part of it the identification of these competitive restraints becomes very difficult.

Furthermore, the multisided nature of products within the digital market make the assessment of abusive conduct through established theories of harm very difficult.[653] This is because anti-competitive behaviour on one side of the market can result in the benefit of the

[645] Daniel Mandrescu, 'Applying (EU) competition law to online platforms: Reflections on the definition of the relevant market' (2018) 41 World Competition: Law and Economic Review 1, 1.
[646] Ibid (Mandrescue) pg.2.
[647] Consolidated Version of the Treaty on the Functioning of the European Union [2016] OJ C202/1 (TFEU) Art.102.
[648] Case 27/76 United Brands Company and United Brands Continental BV v Commission of the European Communities [1978] ECR 207 par. 65-66.
[649] Case C-457/10 P AstraZeneca AB and AstraZeneca plc v European Commission [2012] ECLI:EU:C:2012:770. par. 180.
[650] Case 85/76 Hoffmann-La Roche & Co AG v Commission of the European Communities [1979] ECR 461 par. 39.
[651] Communication From the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings [2009] [2009] OJ C 45 (Article 102 Guidance) par. 10.
[652] Alison Jone, Niamh Dunne and Brenda Sufrin, *Jones & Sufrin's EU Competition Law: Text, Cases & Materials*, (8th edn, Oxford University Press 2023) pg. 331.
[653] David S. Evans and Richard Schmalensee, The Antitrust Analysis of Multi Sided Platform Businesses, Working paper 18783, 28.

consumer.[654]To ensure that normal competition is maintained the Commission has developed novel theories of harm however the nature of the digital market makes it hard to distinguish conduct which is in line with competition on merits and abusive behaviour which leads to anti-competitive foreclosure.[655]

This article will look at how the Commission applies Art. 102 to the digital market and how the methods and approaches it uses have changed due to the unique nature of the digital market. In section 2 looks at how the Commission has traditionally defined the relevant market in which the undertaking in question is operating and how it has adapted for the digital market. In section 3 examines the factors the Commission relies on in order to establish whether the undertaking in dominant within the relevant market and how it has developed its approach for the digital market. Section 4 focuses on which actions by the undertaking can be considered as an abuse of dominance and how the Commission has approached this question in relation to digital markets.

1. Defining The Market

Within this section examines the need to define a relevant market in order to establish the dominance of an undertaking within that market and how these methods have adapted to the digital market. However, as will be pointed out in the analysis below the traditional methods used by the Commission in defining the market do not adequately consider the unique aspects of the digital market.

Analysis of dominance under Article 102 has the starting point of identifying the relevant market that the undertaking is concerned with in order to then assess the position of the undertaking within the market.[656] Identifying the relevant market allows the Commission to identify competitors that can restrain the undertaking from acting independently within the market.[657] The relevant market comprises of the relevant product and geographical market.[658] The relevant product market is defined as all products that are regarded as substitutable or interchangeable by the consumer.[659] This market should include all products that are competitors in order to properly assess the undertakings possible dominance within

[654] Ibid 30.
[655] Pinar Akman, The Reform of the Application of Article 102 TFEU, (2016) 81 Antitrust Law Journal 145, 152.
[656] Mandrescu (n.1) pg. 5 and 2.
[657] Commission Notice on the definition of relevant market for the purpose of Community competition law (1997) OJ C 372 (1997 Relevant Market Notice) par. 2.
[658] Ibid par.4.
[659] Ibid par. 7.

that market and the extent that they can operate independently of normal competitive restraints.[660] The identification of the relevant market product and the competitors within the market is done through an assessment of demand and supply substitution, which have been identified as the most important ways of assessing the competitive restraints on the undertaking.[661] The assessment of demand substitution used to include the use of the SSNIP test.[662] Supply side substitution is not given as much attention by the Commission.[663] Supply substitutability of the product is considered the ease of which a supplier changes his production in order to produce a similar product as in the relevant market.[664] Whether there is supply substitution is determined by the costs, risks and incentives for the supplier to enter the market.[665] If such a change can be done with relative ease then the relevant market can be expanded to include those undertakings as competitors of the undertaking in question.[666] Defining the relevant market within the digital economy has become difficult and prone to errors due to the nature of the product and services present in the market.[667] These different characteristics make the application of the standard approach to market analysis difficult for the digital market.[668] The reason for these difficulties is that traditionally the relevant market was identified against a static market where the price of the product is a central factor.[669] The digital market is defined by the innovation and the rapid development of new products.[670] This dynamic nature of the digital market makes it very difficult to apply traditional static tools for defining the market.[671] Despite the difficulties of defining the relevant market in digital this task is still considered a necessary element of the anti-competitive assessment.[672] This is because defining the relevant market assists competition authorities in the assessment of the relevant position of the undertaking, which can help determine the competitive restraints on the undertaking.[673] Due to the scale of

[660] Commission Notice on the Definition of the Relevant Market for the Purposes of Union Competition Law [2024] OJ C, C/2024/1645 (2024 Relevant Market Notice) para. 25.
[661] Ibid para. 23-24.
[662] 1997 Relevant Market Notice (n.13) par. 15.
[663] Ibid par. 13.
[664] 2024 Relevant Market Notice (n.16) par. 32.
[665] Ibid par. 33.
[666] Ibid par. 34.
[667] Jens-Uwe Franck and Martin Peitz, 'Market Definition in the Platform Economy' (2021) 23 Cambridge Yearbook of European Legal Studies 91, 95.
[668] Jones (n.8) pg 1219.
[669] Viktoria H.S.E. Robertson, 'ANTITRUST LAW AND DIGITAL MARKETS' in Heinz D. Kurz and others (eds), *The Routledge Handbook of Smart Technologies: An Economic and Social Perspective* (1st edn, Routledge 2021) 435; Tone Knapstad, 'Digital dominance: assessing market definition and market power for online platforms under Article 102 TFEU' (2024) 20 European Competition Journal 412, 417.
[670] Case T-79/12 Cisco Systems v Commission, EU: T:2013:635, para 69.
[671] Robertson (n.25) 435; Knapstad (n.25) 417.
[672] Franck (n.23) 98.
[673] 1997 Relevant Market Notice (n.13) Par. 2.

platforms on the digital market the Commission has a difficult job in defining the market in which the platform operates. On the one hand if the Commission defines the market too broadly then it will arguably be difficult to establish the monopoly of online platforms such as Google, allowing them to expand un-impeded, creating a monopoly and limiting the options available for users. However, if the Commission defines the markets very narrowly then arguably it might penalize the legitimate competitive practices, disincentivizing competition in the EU digital market. The approach taken by the Commission fails to adequately consider the unique aspects of the digital market, preferring to rely on the traditional approach in defining the market. This approach has been effective up till this point in curtailing the monopolisation of the market by platforms such as google, however as will be seen below the Commission fails to take into account various commercial realities of the digital market. An argument can be made that the Commission is reluctant to change its approach towards defining the relevant market in order to ensure the effectiveness of Art. 102.

The key difference between the traditional approach of defining the relevant market in digital markets is the existence of multi-sided platforms.[674] The digital market is defined by the creation of online platforms within it which change or create new markets and can be very substitutable with each other.[675] These platforms are described as multi-sided due to the fact that they can attract users from a number of different markets by offering a number of services from the same platform.[676] By bringing together users in different markets, platforms can create network effects where the more users on each side of the market interacts with the platform, the more the opposite side wants to do so as well.[677] In advertisement supported platforms such as the general search services in the Google Shopping decision the network effect is generated by the amount of users that are present on one side of the market which can indicate the demand for the platform on the advertiser side.[678] Strong network effects make it hard to assess the relevant market as it can be affected by the activity of the users.[679] An issue created by the network effects is whether the relevant market should be defined through the inclusion of all sides or whether each side constitutes

[674] Robertson (n.25) 435.
[675] Mandrescue (n.1) 5.
[676] Evans (n. 9) pg.1.
[677] Ibid pg. 9.
[678] Franck (n.23) 92; Google Search (Shopping) (Case AT.39740) Commission Decision [2017] (Google Shopping decision) para. 293.
[679] Knapstad (n.25) 417.

a separate market.[680] This creates difficulties for traditional market assessments as they primarily focus on one-sided markets.[681] If we were to focus on the user's side of the multi-sided market then we risk losing valuable insight on how the platform interacts with the advertisers that is necessary for establishing dominance.[682] It has been argued that due to the network effects the sides of the market should be considered as one market and dividing the market would not acknowledge the interdependence of the market sides.[683] Not considering the interdependence between the sides of the platform makes the market definition unreliable.[684] However, if we consider all sides of the multi-sided market then the relevant market might be defined too broadly, including markets that are not substitutable and making it harder to show that the undertaking in question is dominant within the market.[685] Due to this issue defining the relevant market in the digital economy needs to be more flexible.[686] Further than a flexible market analysis, some experts have advocated that the nature of the digital market challenges the traditional notion of establishing dominance through the identification of the relevant market.[687] Instead they argue that dominance of an undertaking within the market should be established through an assessment of the effects of its conduct within the relevant market.[688] Such a change in approach needs to come from the EU courts, which up till now have maintained the traditional application of Article 102 as seen in the Google Shopping case.[689] The Commission has accepted the issues brought by the digital market but has not moved away from the traditional application of Article 102, opting for a case by case analysis of the market.[690]

In the Google Shopping decision the Commission identified 2 relevant markets, those being that of the general search services market and the comparison shopping market.[691] When defining the relevant market the Commission did not consider the multi-sidedness of the general search service platform, instead merely referring to it when defining such a service

[680] Ibid 418.
[681] Evans (n.9) pg. 23.
[682] Robertson (n.25) 435-436.
[683] Bundeskartellamt, Working Paper - The Market Power of Platforms and Networks, Executive Summary (2016) B6-113/15, 5-6.
[684] Evans (n.9) pg.9.
[685] Robertson (n.25) 436.
[686] Franck (n.23) 96.
[687] Heike Schweitzer, Justus Haucap, Wolfgan Kerber and Robert Welker, Modernizing the Law on abuse of market power, Report for the Federal Ministry for Economc Affairs and Energy (Germany) (2018) (September 17, 2018). <https://ssrn.com/abstract=3250742>pg. 7.
[688] Akman (n.11) 148.
[689] Schweitzer (n.43) 7; Case T-612/17 Google and Alphabet v Commission, ECLI:EU: T:2021:763 (Google Case)
[690] Jones (n.8) pg. 330.
[691] Google Shopping Decision (n.34) para. 154.

was considered an economic activity.[692] As seen in the Microsoft/ LinkedIn decision the approach taken with multi-sided platforms is that the assessment of the relevant market for online platforms depends on its functionality.[693] However this approach focuses only on the user's side functionality splitting the markets of the general search services platform.[694]

For the general search services market the Commission deemed that from the users perspective the service had insufficient interchangeability with other web search services, which was based on the different characteristics and functionality of the different services.[695] A key argument made by the Commission in the decision was that the relevant market did not include specialized search services because they did not aim to offer all possible results for a search, unlike the general search service.[696] However as argued by Bergkamp this shows the inadequacy in the Commission's reasoning as providing an incomplete answer to a request is not a suitable business model that undertakings like Yelp would pursue.[697] Furthermore the Commission saw the use of specialized search services as complementary to that of general search services, ignoring evidence that showed the propensity of users to rely on other general or specialized search services when Google failed to give them a satisfactory answer.[698]

Through this it can be seen that the Commission fails to understand the habits of users in the digital market, opting to look only look the certain service that the certain platform provides to the user rather than how users utilize the services provided by the platforms. In doing so an argument can be made that the Commission narrowly defined the market, not taking into account the fact that Google as a general search service was competing with other search services. I do not believe that this is due to ignorance by the Commission in the habits of users, but rather an effort to ensure that the market is narrowly defined in order to more easily establish an undertaking dominance.

Similarly the Commission defined the comparison shopping market as a distinct relevant market due to lack of interchangeability between the service and other online search or

[692] Robertson (n.25) 436; Google Shopping Decision (n.34) para. 159.
[693] Case M.8124 - Microsoft/LinkedIn Commission Decision C(2016) 8404 final, para. 36.
[694] Knapstad (n.25) 418.
[695] Google Case (n.45) Para. 44-45.
[696] Google Shopping Decision (n.34) para. 168.
[697] Penelope A Bergkamp, 'The European Commission's Google Shopping decision: Could bias have anything to do with it?' (2019) 26 Maastricht Journal of European and Comparative Law 524, 531.
[698] Ibid 531.

retailing services.[699] However when assessing the substitutability of competitive shopping services and merchant platforms the Commission did consider the different sides of the market.[700] In relation to supply side substitution the Commission looked only at the differences in monetization of the interactions and the functionality provided to the retailers.[701] Scholars such as Bergkamp have argued that the commission's assessment failed to take into account the innovation present in the digital market and the fact that these merchant platforms later introduced features that made the need of comparative search services obsolete.[702] This idea is supported by the court's decision in the Continental Cans case, the fact that the other merchant platforms innovated and created competition in the comparative shop services market should have shown that both services were in the same market.[703] Bergen argues that this shows that the Commission failed to consider the commercial realities of the digital market and the innovation present within it, leading it to define an overly narrow relevant market.[704] The view that the Commission needs to take into account the potential development of the platform in the digital market in order to adequately define the relevant market is agreeable, but there is a need for the Commission to develop new methods as to how this can be assessed. This is because assessing the potential innovation of the market is not something that the Commission can adequately assess in defining the market since through the traditional approach of defining the market the Commission needs to define the relevant market through the evidence it has at the time. Although there are arguments setting out that the Commission should have considered the fact that the merchant platform innovated and entered the comparative shopping market, therefore taking into account the dynamic nature of the digital market, an argument can equally be made there was no evidence that such a development would take place. Requiring the Commission to be extremely cautious in defining the relevant market, taking into account all potential innovation paths for the platforms could lead to a very wide definition of the relevant market as every similar service can potentially be considered as competitors in the same relevant market.

Another difficulty related to defining relevant markets in the digital market is the existence of platforms which provide the users with a service free of charge.[705] In the traditional

[699] Google Shopping Decision (n.34) para. 192.
[700] Franck (n.23) 104.
[701] Google Shopping Decision (n.34) para. 225 and 226.
[702] Bergkamp (n.53) 533.
[703] Case 6/27 Continental Can Company Inc v Commission ECLI:EU:C: 1973:22 par. 33 and 34.
[704] Bergkamp (n.53) pg .539.
[705] Robertson (n.25) 435.

approach of defining the relevant market the Commission relied on the SSNIP test in order to assess demand substitution.[706] With this approach the Commission determined whether products were within the same market based on whether customers switched over to them where there was a small but significant increase in the price of the product.[707] The application of the SSNIP test can be seen in the France Telecom case where the Commission introduced a hypothetical 10% increase in the costs of high speed internet access.[708] Based on the fact that 80% of the respondents stated that they would not switch their internet service, the Commission differentiated the market for internet service.[709] In multi-sided platforms however this increase needs to be on all sides of the market in order to properly assess in what way the users of the platform will react to it.[710] This is because an increase to one side of the platform might not have an effect on the users of the platform.[711] This can be very difficult in digital economies as many platforms provide a zero-price service despite monetizing their interactions with another part of the market.[712] This is the case in the Google Shopping decision, since although the users could access the platforms for free, the advertisers on the other side of the market needed to pay to advertise on the platform.[713] In such situations the Commission declined to apply the SSNIP test due to the zero-price market.[714] An alternative test has been suggested in later decisions for zero-price platform in the Google Android case, referred to as the SSNDQ.[715] This test defines the relevant market by considering what platform users would switch if there was a small but significant decrease in the quality of the platform.[716] However the application of this test is unclear.[717]

This section sets out that the traditional approach to defining the market is not easily applied onto the digital market. Although the Commission has accepted that there needs to be another approach in defining digital market it still relies on the traditional approach, which has led to the narrow definition of various digital markets as seen above. It has been argued that the approach does not adequately consider how the digital platforms are utilized by the users,

[706] 1997 Relevant Market Notice (n.13) para. 16.
[707] 1997 Relevant Market Notice (n.13) para. 16.
[708] Wanadoo Interactive – (COMP/38.233) Commission Decision (2003) para. 199.
[709] Ibid para. 199.
[710] Evans (n.9) 22.
[711] Dirk Auer and Nicolas Petit, Two-sided markets and the Challenges of Turing Economic Theory into Antitrust Policy, (2015) 60 Antitrust Bulletin <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2552337> pg. 27.
[712] Bundeskartellamt (n.39) Pg. 7.
[713] Google Shopping Decision (n.34) para. 159.
[714] Google Shopping Decision (n.34) para. 245.
[715] Jones (n.8) 343.
[716] Google Android (Cae AT. 40099) Commission Decision C(2018) 4761 para. 267.
[717] Franck (n.23) 121.

limiting the relevant market to a specific set out characteristics which do not adequately describe the use of the platform. Furthermore, it has been argued that the traditional approach does not consider the innovation of the market. However, the Commission cannot predict the potential development of the platforms and requiring it do so could potentially result in very wide relevant markets. Additionally, through the analysis given by the Commission in defining the relevant market, the network effects created by the use of the search engine where taken into consideration, revealing a weakness in the approach taken by the Commission as it does not take into account an important aspect of digital markets. However as seen by the development SSNDQ test, the Commission has begun to adapt its approach to defining the relevant market to the commercial realities of the digital market, even if the application of the test is still not clearly defined. Through this narrow definition of the market the Commission was able to more easily establish the dominance of the undertaking, as examined below. This narrow definition can however slow down the innovation of the digital market, limiting technological development and competition in the EU. However, it should be noted that the Commission has an interest in narrowly defining the relevant markets. If the arguments of Bergkamp made above where accepted, and the markets where defined as broadly as possible to adequately encompass the digital market, then Art. 102 would not be enforceable. This is because in such a wide market it would be difficult to establish a specific undertakings dominance.

2. Establishing Dominance

After defining the market, the second step to finding whether there has been abuse of dominance is to define the position of the undertaking within the market.[718] This section will be looking into the methods used by the Commission in establishing dominance. The methods used by the Commission in doing so have been created for traditional markets, but the Commission has worked to adapt such tools to the digital market. Furthermore, acknowledging the weaknesses of its traditional methods the Commission has increasingly relied on newer methods which can be applied more easily on digital markets.

One of the methods used by the Commission to establish the position of the undertaking in the relevant market is through the assessment of the undertakings' market shares.[719] Under the Guidance Paper 1997 the assessment of market share was based on the number of sales

[718] Jones (n.8) 331; TFEU (n.3) Art. 102.
[719] Article 102 Guidance (n.7) par. 13.

of the relevant product in the market.[720] Although not the only factor in the assessment of dominance, market shares play a significant role where the undertaking holds a high market share.[721] In *Telekomunikacja Polska* decision this was adapted to the digital market as being the number of subscribers that the undertaking has to the relevant service it provides.[722] However for zero-price products this assessment of market power is not applicable as there are no sales to measure the undertaking market shares from.[723] Due to this the Commission in the *Google Shopping* decision determined that market share can be calculated through the number of visitors coming to the platform.[724] Through this it was established that Google had over 85% market share in the general search services market in most EEA since 2008.[725] In the 2009 Guidance Paper high market share held over a significant period of time is a strong indicator that the undertaking in question is dominant within the market.[726] In the Hoffman La Roche case the courts determined that high market share could constitute a presumption of dominance as it makes the undertaking an unavoidable trading partner.[727] However within the digital economy, market shares are not a reliable indicator of market power.[728] This is because such markets are defined by the innovation and short-term development cycles which means users can quickly migrate to other services.[729] Therefore the static indicators of market power such as market shares are not reliable in the digital economy.[730] This was also acknowledged in the Microsoft/Skype decision where the Commission accepted that market shares can quickly change in dynamic sectors such as in the digital market, limiting their ability to show dominance.[731] Despite this the Commission still believe that they are a useful tool, especially in the context of the Google Shopping decision as the high market share of Google was not affected by the short innovation cycle of the digital economy.[732]

A further difficulty in the assessment of dominance of an undertaking in the digital market are the multisided platforms. The market share of such platforms is difficult to assess

[720] 1997 Relevant Market Notice (n.13) par. 53.
[721] Hoffmann La Roche (n.6) para. 39.
[722] Telekomunikacja Polska (COMP/39.525) Commission Decision (2011) *OJ C 324/7*, Section 3.1.1.
[723] Google Shopping Decision (n.34) par. 275.
[724] ibid par. 273.
[725] ibid par. 283 and 284.
[726] Article 102 Guidance (n.7) par. 15.
[727] Hoffmann-La Roche (n.6) par. 41.
[728] Robertson (n.25) 438.
[729] Cisco Systems (n.26) para 69.
[730] Select Committee on European Union, Online Platforms and the Digital Single Market (HL Paper 129) pg.10.
[731] Microsoft/Skype (COMP/M.6281) Commission Decision C(2011) 7279 par. 78.
[732] Google Shopping Decision (n.34) par.267.

because each side of the platform offers a different service to its users.[733] In the Google Shopping case these sides of the general search service market were the user side and the online search advertising side of the market.[734] In the advertisement side of the market, market shares can be a useful tool in order to assess the undertaking's dominance in the relevant market.[735] The Commission erred in its assessment of the market shares of the undertaking by focusing on only the user side of the general search services market and did not consider the market power of Google in the advertisement side of the market.[736] According to Eleni Argentisi and Lapo Filistrucchi in order to understand the full extent of the undertaking market power in the relevant market the Commission needs to consider both sides of the market because of the interrelated nature of the platforms.[737] This interrelation in the general search services can be seen in the fact that the amount of advertisers that are willing to use the platform depends on the number of users on the other free side of the platform.[738]

Due to these issues the assessment of dominance through market share provides a very limited view of the undertaking's position within the market.[739] In acknowledging the difficulties of primarily relying on an undertaking market shares for the establishment of dominance, the Commission has developed other methods for the assessment of an undertaking position. Therefore the approach has shifted from an assessment of market shares towards other factors such as barriers to entry as a better indicator of the undertaking's position in the market.[740] The constraints imposed on an undertaking through potential expansion or entry of competitors into the relevant market can be used as an indicator of dominance.[741] The existence of high barriers of entry can also lead to entrenchment of dominance as it prevents entry of new competitors in the market where the dominant undertaking has foreclosed its competition.[742] Assessing the barrier to entry is a necessary requirement when establishing dominance as the ability for other undertakings to enter the relevant market can dilute the market power of the undertaking in question.[743] When

[733] Evans (n.9) pg.20.
[734] Google Shopping Decision (n.34) par. 159.
[735] Knapstad (n.25) 430.
[736] Google Shopping Decision (n.34) par. 275.
[737] Eleni Argentisi and Lapo Filistrucchi, 'Market Power in a Two-Sided Market: The Case of Newspapers' (2007) 22 Journal of Applied Econometrics 1247, 1264.
[738] Google Shopping Decision (n.34) para. 159.
[739] Knapstad (n.25) 431.
[740] Franck (n.23) 124.
[741] Article 102 Guidance (n.7) par. 12.
[742] Ibid par. 20.
[743] ibid par. 16.

assessing the barriers to entry the Commission considers the likelihood of an undertaking entering the market based on the costs of doing so and the risks of failing to enter.[744] Additionally the Commissions considers how quickly a potential competitor can enter the market when assessing barriers to entry.[745] These barriers can take a number of forms, such as legal barriers or access to key inputs.[746] In the digital markets, due to the dynamic nature of the market, barriers to entry relate to network effects generated by the platforms and the ability to access and use data.[747]

As discussed above the interrelation between the general search service and the online search advertisement create network effects in favour of the undertaking.[748] As the Commission found in the Google Shopping decision is because the more users are present on the platform the higher the demand is for the advertiser as there is a higher chance of them matching with a user and concluding a sale.[749] These network effects limit the supply side substitutability of the platforms and create barriers for entry and expansion within the market.[750] Once an undertaking reaches a certain size it creates this positive feedback effect between the sides of the platform.[751] This attracts users from both sides of the market, leading to a concentration in the market and entrenching the position of the undertaking within the market.[752] As acknowledged by the Commission network effects therefore raise barriers to entry as potential competitors cannot enter the market and gain a sufficient number of users to overcome the undertakings in question network effect.[753] This effectively limits the number of viable competitors that the undertaking has because even if the competitor offers a better service.[754] Although network effects have been calculated through the number of users in the platform, there is no definitive way of measuring these effects.[755] Due to this there have been arguments that the effect of network effects should not be overstated because they wear out once the platform reaches a certain threshold.[756] Furthermore, it has been suggested in a working paper produced by the Bundersktartellamt that the innovation within

[744] ibid par. 16.
[745] Ibid par. 16.
[746] ibid par. 17.
[747] Knapstad (n.25) 429 and 431.
[748] Google Shopping Decision (n.34) para. 292.
[749] ibid para. 293.
[750] Evans (n.9) 19.
[751] Evans (n.9) 19.
[752] Bruno Lasserre, Andreas Mundt, Competition Law and Big Data The enforcers' view (2017) Italian Antitrust Review, DOI: 10.12870/iar-12607 87, 95.
[753] Google Shopping Decision (n.34) para. 292; Knapstad (n.25) 433.
[754] Evans (n.9) 14; Knapstad (n.25) 431.
[755] Bundersktartellamt (n.39) pg.12.
[756] Bergkamp (n.53) 532.

the market in combination with network effects can promote the entry and expansion of competitors, allowing them to gain a sizable user base in a small amount of time.[757] However following this approach would make it difficult for the Commission to establish the dominance of an undertaking within the relevant market. In relation to the *Google Shopping* decision the Commission has not considered such arguments in the general search services markets, basing its opinion on the fact that many competitors of Google have failed to enter the market.[758] These indicates that the Commission considers some of the aspects of the digital market when deciding on the dominance of the undertaking. However, its assessment is based on pre-established trends in the digital market, such as the failure of other platforms to enter the general search service. This shows that the Commission places limited weight on the potential innovation present within the digital market.

Access to data and the ability to analyse it can establish the dominance of an undertaking within the market.[759] Access to data can create a competitive advantage for the undertaking as the data can be used to improve the quality of service provided to the user.[760] Platforms require a certain amount of data in order to remain viable within the digital market.[761] Therefore the dependence of data for the effective competition within the digital market indicates that there is a barrier to entry into the market.[762] It has been argued that the data should have a limited effect due to the diminishing returns on the improvements made to the service, however the Commission stated that these concerns had little relevance when assessing barriers to entry.[763]

In assessing the position of the undertaking in order to establish a dominant position within the relevant market, the Commission has proved itself to be very flexible in assessments of platforms in digital markets. It has adapted existing methods such as the assessment of market shares to digital markets, looking into the number of users as an indicator of dominance. It has also accepted the limitations of the above method in relation to digital markets, adopting better indicators of an undertaking position in the market such as barriers to entry. In keeping up with the development of the digital market the Commission also included network effects created by platforms as a factor indicating the existence of barriers

[757] Bundersktartellamt (n.39) pg.12; Lasserre (n.108) 95.
[758] Google Shopping Decision (n.34) para. 298-305.
[759] Robertson (n.25) 438-439.
[760] Google Shopping Decision (n.34) par. 287.
[761] Ibid par. 287.
[762] Knapstad (n.25) 432.
[763] Google Shopping Decision (n.34) par. 289; Bergkamp (n.53) 532.

to entry. Furthermore, the Commission has acknowledged the importance of date in the digital market and has effectively utilized it in its assessment of dominance. However, similarly to the methods used to define the relevant market we can see difficulties in the Commissions approach in relation to the innovation present in the digital market.

3. Abusive Conduct

The next step in determining whether an undertaking has breached Article 102 is to determine if its conduct constitutes an abuse of its position within the market.[764] Holding a dominant position is not in itself a breach of Article 102.[765] However under EU law a dominant undertaking is under a responsibility not to act in a way that would affect genuine competition within the market.[766] A dominant undertaking is still allowed to compete within the market on the merits of the service it provides.[767] Abusive conduct is considered to be practices that might harm the interest of the consumer, it also covers practices that can indirectly harm the consumer by hindering effective competition within the market.[768] Article 102 sets out some examples of abusive conducts but the list provided is not an exhaustive one.[769] Instead EU competition law aims to show the connection between the harm caused to the competition by the practice and the negative effects it has on competition in order to establish a breach.[770] This approach looks at the effects the conduct has on aspects of the product that are beneficial to the consumer such as price.[771] This means that the Commission focuses on ensuring that the market functions properly so that the consumer can benefit from the result of efficient competition.[772] However traditional theories of harm such as predatory pricing are not well suited for the digital market as the economic analysis that underpins them is based on one-sided markets.[773]

[764] TFEU (n.3) Art. 102.
[765] Case C-209/10 *Post Danmark A/S v Konkurrencerådet* [2012] ECR I-0000 para.21.
[766] Case 322/81 NV Nederlandsche Banden Industrie Michelin v Commission ECLI:EU:C:1983:313 par.57.
[767] Case T-286/09 Intel Corp v European Commission ECLI:EU: T: 2014:547 par. 205.
[768] C-286/13 P Dole Food Company Inc nd Dole Fresh Fruit Europe v European Commission ECLI:EU:C: 2015:184 par. 125.
[769] TFEU (n.3) Art. 102; Continental Can (n.59) par. 26.
[770] Jacques Cremer, Yves-Alexandre de Montjoye, and Heike Schweitzer, *Competition Policy for the Digital Era* (European Commission, 2019) 40.
[771] Amendments to the Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings (2023) OJ C119, par. 19.
[772] Akman (n.11) 151.
[773] Evans (n.9) 28.

A new theory of harm developed that is predominantly used against dominant undertakings in the digital market is self-preferencing.[774] Defining a new theory of harm therefore requires the Commission to identify the conduct that can realistically result in foreclosure effects that are not based on normal competition which can mean framing the actions of a platform as distorting competition.[775] This new theory of self-preferencing refers to the idea that the operator of the platform gives undue support to other services provided by his platform in the detriment of the other platforms which provides the same service.[776] This theory of harm is based on the pre-established theory of leveraging, which has been adapted for digital markets.[777] Leveraging refers to the use of the dominant undertaking's power in one market to increase power in another.[778] Leveraging practices are not in themselves prohibited under EU law as the exclusion of competitors in a neighbouring market does not constitute a departure from competition on the merits.[779] This is in line with the Commission purpose of protecting competition rather than the competitors within the relevant market.[780] The difficulty present in assessing abusive conduct within the digital market is whether the conduct of the undertaking in leveraging its market power complies with competition on the merits.[781] Exclusionary tactics are a part of the digital market because it is a winner takes all market, meaning the first platform to reach a certain number of users becomes dominant within the market.[782] Additionally when a platform reaches its full potential within its market it expands to neighbouring markets in order to continue to grow and generate more revenue.[783] However in order to succeed in the new market it needs to reach a certain critical mass of users in order to become viable.[784] To do so dominant undertakings can leverage their pre-existing user base to become viable within the new market which can constitute competition on the merits.[785] In relation to self-preferencing it could be argued that it is the reward for managing the platform in question, giving the undertaking an edge over its

[774] Jones (n.8) 1223.
[775] Daniel Mandrescu, 'Applying Article 102 TFEU to Multisided Online Platforms: Discrimination, Leveraging and Undefined Abuses of Dominance' in Luca Calzori, Alberto Miglio, Filippo Croci, Chiara Cellerino and Jacopo Alberti (Eds) Public and Private Enforcement of EU Competition Law in the Age of Big Data, (Giappichelli 2024) pg. 110-111.
[776] Jones (n.8) 546.
[777] Ibid
[778] Ibid 423.
[779] Google Case (n.45) par. 162 and 164.
[780] Akman (n.11) 151.
[781] Mandrescu, Applying Art. 102 (n.131) pg. 97.
[782] Ibid 95; Evans (n.9) 29.
[783] Mandrescu, Applying Art. 102 (n.131) 96.
[784] Evans (n.9) 29.
[785] Mandrescu, Applying Art. 102 (n.131) 96.

competitors for the innovations it made to outperform them.[786] Furthermore, due to the interdependence in the multi-sided platform, exclusionary tactics such as self-preferencing can increase consumer welfare.[787] This can be seen in the TomTom/Tele Atlas merger where in order for the undertaking to utilize the pro-competitive gains resulting from the vertical merger between the undertakings they had to be more favourable to their own upstream product.[788]

Self-preferencing is not therefore considered per se abusive, meaning there is not a prohibition on dominant undertakings using this tactic, but it can breach Article 102 if it has anti-competitive effects.[789] In assessing the effects of Google's conduct, the Commission argued that the traffic generated for comparison shopping service by the general search services market indispensable for the operation of third party service as they could not receive such traffic from any other source.[790] Therefore the conduct was abusive because it diverted traffic from third party service to its own service, artificially profiting from a service that it had not invented.[791] In doing so the undertaking was engaged in an exclusionary tactic aimed at depriving competitors of the user base they required to stay viable within the market, therefore weakening competition.[792] The court accepting the reasoning of the Commission stated that this deprivation of users constituted a barrier to entry, preventing the swift entrance of new competitors to provide an alternative option for users.[793] This shows that the Commission was successful in developing new tools to address new forms of abusive conduct which are found in the digital markets. In doing so the Commission accepted some commercial realities of the digital market, such as the "winner takes all" aspect of the digital market and has considered the effect of users and network effects have in the expansion of a platform within a relevant market. In doing so, however, the Commission has been given a great deal of discretion when establishing when conduct is abusive. Through this decision the court endorsed an expansive reading of a dominant undertaking's responsibility not to distort normal competition.[794] Additionally this decision showed that any conduct that interfered with competition on the merits could breach Article 102 even if it didn't fall within

[786] Pablo Ibanez Colomo, Self-Preferencing: Yet another Epithet in need of limiting principles. (2020) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3654083> Pg.7; Cremer (n.126) 66.
[787] Evans (n.9) 30.
[788] Colomo (n.142) 14-15.
[789] Cremer (n.126) 66.
[790] Google Case (n.45) par. 169.
[791] Google Shopping Decision (n.34) par. 341 and 343.
[792] Evans (n.9) 29; Google Case (n.45) par. 169.
[793] Google Case (n.45) par. 178.
[794] Elias Deutscher, Google Shopping and the Quest for a Legal Test for Self-preferencing under Article 102 TFEU (2021) 6 European Papers 1345, 1349.

pre-established theories of harm.[795] This flexibility given to the Commission to determine abusive conduct allows them to better enforce Article 102, however it can create uncertainty in what conduct could constitute a breach.[796] This uncertainty could deter undertakings from making potentially pro-consumer choices.[797] This issue can be seen through the Google Shopping decision where the courts accepted the Commission's reasoning that self-preferencing was a different theory of harm than refusal to supply.[798] However in doing so the courts and the Commission failed to set out a test to limit the reach of this new theory of harm.[799] This lack of limits means that the commission can interfere with an undertaking's discretion on how to develop and design their platform.[800] This can be seen in the measure imposed on Google, requiring them to change how the interface presented comparative shopping services.[801]

This indicates the Commissions cautiousness with the digital market, focusing on ensuring that there are no potential anti-competitive effects caused by an undertaking, even if it means limiting innovation. This is because the Commission's competition laws are based on the idea that improvement of consumer welfare comes from unimpeded competition.[802] However, this approach does not take into account the winner takes all aspects of the digital market and the innovation inherently present in the digital markets which can improve consumer welfare.

Conclusion

Through this article we have seen that the Commission has begun to adapt its approach of applying Art. 102 to the digital market.

In defining the relevant digital market the Commission faces difficulties as the multi-sided aspect of platforms within the digital market makes it hard to define the relevant market and the undertaking position within it.[803] The network effects generated by the platform create difficulties in defining the market as the different sides of the market are interdependent

[795] Ibid 1349.
[796] Akman (n.11) 152.
[797] Ibid 152.
[798] Deutscher (n.150) 1351.
[799] Ibid 1352.
[800] Ibid 1349.
[801] Google Shopping Decision (n.34) para. 700.
[802] Jones (n.8) 50.
[803] Evans (n.9) pg. 23.

meaning that the exclusion of them could result in a very narrowly defined market.[804] Despite these issues the Commission has not made any significant changes in how it defines the relevant market, essentially ignoring the unique aspects of the digital market. The Commission has rejected arguments relating to the potential innovation of platforms which can make them competitors. Furthermore, little attention is paid to the network effects generated by the platform. The Commission has, however, developed the SSNDQ test which applies to the digital market, but the application of this test is still unclear. An argument can be made that this is a conscious decision by the Commission in order to limit the relevant market by not including the potential for innovation by platforms and the multi-sided nature of such platforms. Such an approach ensure that a market is narrowly defined and making it easier to apply Art. 102 by establishing that the undertaking is dominant in the relevant market.

In relation to assessing the undertaking position in the relevant market, the characteristics of the digital market, such as the zero-price nature of services within the digital market makes it difficult to apply traditional methods of assessing demand substitutability.[805] These characteristics of the digital market make it hard to assess the position of the undertaking within the relevant market because network effects can influence the number of users a platform has.[806] Additionally, the rapid innovation within the digital market makes market shares an unsuitable method of assessing market power as through the network effects a new competitor can quickly become dominant within the relevant market.[807] However, in establishing the dominance of the undertaking the Commission has proven itself more adaptable than in defining the relevant market. It has adopted the market shares method of determining the dominance of the undertaking to the digital market by focusing on the number of users of a platform to show dominance. Furthermore, the Commission has begun taking into account the network effects created by a platform. Despite this adaptability we can see that the Commission still struggles with fitting the innovation present in the digital market in its assessment under Art. 102.

Finally, pre-established theories of harm used by the Commission are not suited for the digital market, making the development of new theories necessary.[808]In adapting to the

[804] Knapstad (n.25) 418.
[805] Bundeskartellamt (n.39) Pg. 7.
[806] Knapstad (n.25) 417.
[807] Bundersktartellamt (n.39) pg.12; Lasserre (n.108) 95.
[808] Evans (n.9) 28.

digital markets the Commission has itself innovated, developing the self-preferencing theory of harm in order to determine when conduct can be considered abusive. However, these new theories seem unable to distinguish between legitimate practices and abusive conduct.[809] This is because some exclusionary practices are commonplace within the digital market, but the Commission and the courts have failed to limit the new theories reached through the establishment of a test to determine when such conduct becomes abusive.[810] This means that the Commission is now able to determine how an undertaking will develop its platform, potentially damaging innovation within the digital market. [811]

[809] Mandrescu, Applying Art. 102 (n.131) pg. 97.
[810] Akman (n.11) 152.
[811] Ibid; Deutscher (n.150) 1349.

BIBLIOGRAPHY

Akman P, The Reform of the Application of Article 102 TFEU, (2016) 81 Antitrust Law Journal 145

Auer D and Petit N, Two-sided markets and the Challenges of Turing Economic Theory into Antitrust Policy, (2015) 60 Antitrust Bulletin <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2552337>

Argentisi E and Filistrucchi L, 'Market Power in a Two-Sided Market: The Case of Newspapers' (2007) 22 Journal of Applied Econometrics 1247

Bundeskartellamt, Working Paper - The Market Power of Platforms and Networks, Executive Summary (2016) B6-113/15

Bergkamp PA, 'The European Commission's Google Shopping decision: Could bias have anything to do with it?' (2019) 26 Maastricht Journal of European and Comparative Law 524

Cremer J, Montjoye YA, and Schweitzer H, Competition Policy for the Digital Era (European Commission, 2019)

Colomo PI, Self-Preferencing: Yet another Epithet in need of limiting principles. (2020) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3654083>

Deutscher E, Google Shopping and the Quest for a Legal Test for Self-preferencing under Article 102 TFEU (2021) 6 European Papers 1345

Evans DS and Schmalensee R, The Antitrust Analysis of Multi Sided Platform Businesses, Working paper 18783

Franck J-U and Peitz M, 'Market Definition in the Platform Economy' (2021) 23 Cambridge Yearbook of European Legal Studies 91

Jone A, Dunne N and Sufrin B, *Jones & Sufrin's EU Competition Law: Text, Cases & Materials* (8th edn, Oxford University Press 2023)

Knapstad T, 'Digital dominance: assessing market definition and market power for online platforms under Article 102 TFEU' (2024) 20 European Competition Journal 412

Lasserre B, Mundt A, Competition Law and Big Data The enforcers' view (2017) Italian Antitrust Review, DOI: 10.12870/iar-12607 87

Mandrescu D, 'Applying (EU) competition law to online platforms: Reflections on the definition of the relevant market' (2018) 41 World Competition: Law and Economic Review 1

Mandrescu D, 'Applying Article 102 TFEU to Multisided Onlne Platforms: Discrimination, Leveraging and Undefined Abuses of Dominance' in Calzori L and others (Eds) Public and Private Enforcement of EU Competition Law in the Age of Big Data, (Giappichelli 2024)

Robertson VHSE, 'ANTITRUST LAW AND DIGITAL MARKETS' in Kurz HD and others (eds), *The Routledge Handbook of Smart Technologies: An Economic and Social Perspective* (1st edn, Routledge 2021)

Schweitzer H, Haucap J, Kerber W, Welker R, Modernizing the Law on abuse of market power, Report for the Federal Ministry for Economc Affairs and Energy (Germany) (2018) (September 17, 2018)

Select Committee on European Union, Online Platforms and the Digital Single Market (HL Paper 129)

Table of Cases and Decisions

C-286/13 P Dole Food Company Inc nd Dole Fresh Fruit Europe v European Commission ECLI:EU:C:2015:184

Case 27/76 United Brands Company and United Brands Continental BV v Commission of the European Communities [1978] ECR 207

Case 322/81 NV Nederlandsche Banden Industrie Michelin v Commission ECLI:EU:C:1983:313

Case 6/27 Continental Can Company Inc v Commission ECLI:EU:C:1973:22

Case 85/76 Hoffmann-La Roche & Co AG v Commission of the European Communities [1979] ECR 461

Case C-209/10 Post Danmark A/S v Konkurrencerådet [2012] ECR I-0000

Case C-457/10 P AstraZeneca AB and AstraZeneca plc v European Commission [2012] ECLI:EU:C:2012:770

Case M.8124 - Microsoft/Linkedin Commission Decision C(2016) 8404 final

Case T-286/09 Intel Corp v European Commission ECLI:EU:T:2014:547

Case T-612/17 Google and Alphabet v Commission, ECLI:EU:T:2021:763

Case T-79/12 Cisco Systems v Commission, EU:T:2013:635

Google Android (Cae AT. 40099) Commission Decision C(2018) 4761

Google Search (Shopping) (Case AT.39740) Commission Decision [2017] (Google Shopping decision)

Microsoft/Skype (COMP/M.6281) Commission Decision C(2011) 7279

Telekomunikacja Polska (COMP/39.525) Commission Decision (2011)

Wanadoo Interactive – (COMP/38.233) Commission Decision (2003)

Statutes

Amendments to the Communication from the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings (2023) OJ C119,

Commission Notice on the definition of relevant market for the purpose of Community competition law (1997) OJ C 372

Commission Notice on the Definition of the Relevant Market for the Purposes of Union Competition Law [2024] OJ C, C/2024/1645

Communication From the Commission – Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings [2009] OJ C 45

Consolidated Version of the Treaty on the Functioning of the European Union [2016] OJ C202/1

When One Cause isn't Enough: Concurrency and The Limitations of The But-For Test in Insurance

Vernon Smith

Abstract

This article examines the reliance on the but-for test to determine causation in English insurance, arguing while the test remains effective for single-cause scenarios, it proves inadequate when multiple concurrent causes operate. Focusing on key cases, *Wayne Tank & Pump Co Ltd v Employers' Liability Assurance Corp*, *JJ Lloyd Instruments Ltd v Northern Star Insurance Co Ltd, Orient-Express Hotels v Assicurazioni Generali, and FCA v Arch Insurance*, it explores the evolution from strict but-for logic toward a broader proximate-cause approach. Drawing on scholarship by Stapleton, Song, Clarke, and others, the article shows how courts increasingly interpret policies and commercial context to address concurrency, rather than mechanically applying but-for analysis. The piece concludes that the but-for test, though still a foundational tool in simpler cases, has been judicially recognized as too narrow for multi-cause disputes, marking a pivotal shift in the way English courts reconcile coverage under overdetermined insurance claims.

Introduction

Causation is a cornerstone of English insurance law, requiring the insured to prove that a covered peril actually caused the loss.[812]

1. The But-For Test and Its Alleged Inadequacy

Traditionally, courts have relied on the but-for test, imported from tort law, to determine whether the insured peril was a factual cause. Put simply, the test asks whether the loss would have occurred but for the operation of the peril in question. If the answer is yes, then that peril is not considered to have caused the loss. The leading English formulation appears in *Cork v Kirby Maclean Ltd*, where Denning LJ explained that a cause must make a difference to the outcome this approach is examined in more detail in Part II below. However, critics have long argued that this approach can be inadequate, especially where multiple concurrent or overdetermined causes operate. In *Wayne Tank & Pump Co Ltd v Employers' Liability*

[812] Marine Insurance Act 1906, s 55(1).

Assurance Corp Ltd (Wayne Tank), the Court of Appeal grappled with the effect of one excluded and one insured peril acting together. In contrast, in *JJ Lloyd Instruments Ltd v Northern Star Insurance Co Ltd (The Miss Jay Jay)*, coverage was maintained if the second cause was merely uninsured rather than excluded.[813] The tension became more pronounced in *Orient-Express Hotels Ltd v Assicurazioni Generali SpA (Orient-Express)*, which rigidly applied a but-for analysis and drew criticism for failing to account properly for wide-area damage.[814] Ultimately, in *Financial Conduct Authority v Arch Insurance (UK) Ltd (FCA v Arch)* the Supreme Court overruled *Orient-Express* in certain contexts, expressly recognising that the but-for test can be under- or over-inclusive for multi-cause insurance disputes.[815]

Against this backdrop, this article critically examines whether it is judicially recognised that the but-for test is indeed inadequate for determining causation in insurance law. The first section outlines the historical but-for approach and its conceptual merit in simpler, single-cause cases. The second explores key concurrency principles, *Wayne Tank* and *The Miss Jay Jay* and their implications for policy construction. The third considers the rigid application of but-for in *Orient-Express* and the Supreme Court's subsequent rejection of that approach in *FCA v Arch*. The article ultimately examines the wider doctrinal and commercial consequences of this judicial shift, positing that although the but-for test remains a foundational concept, it is insufficient on its own to address disputes involving multiple perils. In closing, it is proposed that the increased focus on proximate cause and policy interpretation indicates a significant shift away from the rigid application of the but-for test, reinforcing the assertion that this method is inadequate for handling overdetermined insurance claims in English law.

2. The But-For Test in English Insurance Law

2.1 Origins and Rationale

At root, the but-for test aims to answer a factual question: But-for the act or event in question, would the damage have occurred?[816] In *Cork v Kirby Maclean Ltd,* Denning LJ (as he then was) held that if the damage would have occurred irrespective of a defendant's breach, that

[813] Wayne Tank & Pump Co Ltd v Employers' Liability Assurance Corp Ltd [1974] QB 57 (CA); JJ Lloyd Instruments Ltd v Northern Star Insurance Co Ltd (The Miss Jay Jay) [1987] 1 Lloyd's Rep 32 (CA).
[814] Orient-Express Hotels Ltd v Assicurazioni Generali SpA [2010] EWHC 1186 (Comm).
[815] Financial Conduct Authority v Arch Insurance (UK) Ltd [2021] UKSC 1.
[816] *Cork v Kirby Maclean Ltd* [1952] 2 All ER 402 (CA)

breach was not its cause.[817] Insurance law historically borrowed this threshold approach from tort to ascertain whether an insured peril played a factual role in producing the loss. Some commentators consider it a 'minimum threshold test'.[818] As Merkin and Colinvaux note in *Colinvaux's Law of Insurance*: 'It is necessarily the position that an insurer is to be liable only for those losses which occur as the result of the fortuitous operation of insured perils,' this principle is codified in the Marine Insurance Act 1906,[819] and if there is only one cause, the assured may freely classify it as falling within the policy.[820] This logic reflects the but-for test's simplicity in single-cause contexts: a peril either prevents or triggers the loss, and if absent that peril, the loss would not have happened; the peril is the factual cause.

Applying the but-for test has certain merits. It avoids imposing liability for losses arising purely from the insured's own failure to mitigate or from external events not contemplated by the policy. As Song points out, in tort or criminal cases, the condition typically is a human act or omission, so the but-for test can readily filter out baseless claims.[821] By extension, in single-peril insurance disputes, if the damage would have occurred anyway because of another unstoppable factor, courts find coverage lacking.[822]

2.2 Recognised Limitations in The Literature

However, scholars have cautioned that a purely mechanical but-for analysis can produce anomalies. Stapleton warns that 'if private law maintains that for a factor to be recognised as a "cause", it must bear a "but-for" relation … it will, in an important set of cases, fail to identify … factors that made a positive contribution … but were unnecessary.'[823] This over- and under-inclusiveness emerges where multiple factors operate. Song similarly emphasises that while the but-for test remains a basic factual inquiry, it is unfit for concurrency: it fails in oversubscribed losses, as each factor alone might be unnecessary yet still a real cause.[824]

In short, though but-for logic supplies a baseline, there has long been a stream of criticism that it breaks down in the face of multiple concurrent perils, one or more of which might be

[817] ibid 13 (Denning LJ)
[818] R Merkin and R Colinvaux (eds), *Colinvaux's Law of Insurance* (13th ed, 2024) para 5–067.
[819] Marine Insurance Act 1906, s 55(1).
[820] Merkin and Colinvaux (n 818) para 5–067
[821] M Song, *Causation in Insurance Contract Law* (2nd ed, Informa Law from Routledge 2024) 71.
[822] Cork v Kirby (n 816); Leyland Shipping Co v Norwich Union Fire Insurance Society [1918] AC 350 (HL); Orient-Express (n 814)
[823] J Stapleton, 'Unnecessary Causes' (2013) 129 LQR 39, 39.
[824] Song Causation (n 821) 70–76.

excluded. Such shortcomings prompt courts to turn to proximate cause or to interpret policy wording to avoid common sense defying outcomes,[825] as we shall examine.

3. Multiple Concurrent Causes and The Limits of The But-For Approach

3.1. The *Wayne Tank* Principle and *The Miss Jay Jay* Rule

Where multiple events cause or contribute to the same loss, insurers or insureds often dispute which peril should be designated the proximate cause. English law has evolved two well-known concurrency principles:

- Suppose one cause is insured, and another is an excluded peril, both being approximately equal in efficacy. In that case, coverage is defeated, known as the *Wayne Tank* principle, from *Wayne Tank & Pump Co Ltd v Employers' Liability Assurance Corp Ltd*.[826]
- If one cause is insured and another is merely uninsured (rather than excluded), the insured can recover, known as the Miss Jay Jay principle from *JJ Lloyd Instruments Ltd v Northern Star Insurance Co Ltd* (*The Miss Jay Jay*).[827]

The difference turns on whether the second cause is actually excluded from coverage. A As Song observes, the distinction between excluded and merely uninsured causes reflects not just causation logic but also judicial interpretation of contractual intent.[828]

3.1.1. The *Wayne Tank*

In *Wayne Tank*, a factory was destroyed by a fire that arose from two interdependent causes: a negligent supply of defective equipment (excluded) and its unattended operation (covered). The Court of Appeal found that coverage was defeated because the excluded cause joined with an insured cause to produce the loss.[829] *Wayne Tank* thus solidified a concurrency rule in English insurance: if an insured peril and an excluded peril both operate interdependently to cause the loss, the exclusion wins.[830]

---

[825] M Song, 'Revisiting concurrent causation and principles in English insurance law: a legal fiction?' (2021) 6 JBL 457, 457–465.
[826] Wayne Tank (n 813)
[827] Miss Jay Jay (n 813)
[828] Song Revisiting (n 825) 457- 458
[829] Wayne Tank (n 813); Merkin and Colinvaux (n 818)
[830] Wayne Tank (n 813); Song Revisiting (n 825) 462 - 464

3.1.2. *The Miss Jay Jay*

In *Miss Jay Jay*, a yacht was damaged by two concurrent causes: a design defect (uninsured) and bad weather (covered). The Court of Appeal held that because the design defect was merely uninsured rather than excluded, its presence did not defeat coverage. *Miss Jay Jay* confirmed that if an insured peril and an uninsured (but not excluded) peril operate together to cause the loss, coverage remains.[831]

While straightforward in concept, concurrency law can become mired in interpretive nuance. *Wayne Tank* was decided on the premise that both perils were concurrent and interdependent. These concurrency principles can turn out to be ambiguous when multiple wide-area factors converge, such as war perils or large-scale natural disasters.[832] For instance, if a property is damaged by looting during a riot (a covered peril) that occurs simultaneously with a state of civil war (an excluded peril), it may be unclear whether the damage stems from the riot itself or from the broader war conditions, raising questions about how the concurrency rules should apply.

The courts, however, rarely start an extended inquiry into concurrency, sometimes labelling one cause dominant to avoid concurrency altogether.[833]

This highlights how concurrency complicates a purely factual but-for analysis, prompting courts to consider whether policy wording or proximate cause doctrine better captures the true effect of multiple perils.

3.2. Stapleton's Unnecessary Cause Theory and Its Relevance

Jane Stapleton offers a helpful concept called unnecessary causes, highlighting that certain factors may meaningfully contribute to a loss even though they were not required under the strict but-for test. She writes:

> '… private law may well be normatively interested in a factor which made a positive contribution to the relevant step in the mechanism by which an

[831] Miss Jay Jay (n 813)
[832] Merkin and Colinvaux (n 818) para 5 - 096
[833] Song Revisiting (n 825) 464 - 465

> indivisible injury occurred even though its relation to the occurrence of the injury was unnecessary'.[834]

Put simply, Stapleton contends that if a factor genuinely contributes to the final outcome, the law ought to treat it as causal, even if it was not essential on its own. This view challenges the orthodox but-for approach, particularly in multi-cause scenarios.[835] Applying this idea to English insurance law helps us see potential flaws in decisions that hinge on identifying a single dominant cause or reject coverage based solely on a but-for analysis of one peril. Ultimately, Stapleton's analysis prompts insurers and courts to look at how each factor operates rather than automatically excluding a loss if any peril could alone have caused it.

3.3. Reappraising *Wayne Tank* and *Miss Jay Jay* Through Stapleton's Lens

3.3.1. The *Wayne Tank* Principle

In *Wayne Tank*,[836] the Court of Appeal denied coverage where two concurrent, interdependent perils operated: an excluded peril (defective equipment) and a covered peril (negligent operation). Since both worked together to destroy the factory, the presence of the excluded peril nullified any claim. Under Stapleton's approach, one might ask whether the excluded defective equipment peril was truly essential or whether it simply added to the outcome. If it was not crucial, it may have been an unnecessary but still real contributing factor, something the law should still consider in deciding coverage.

3.3.2. *The Miss Jay Jay* Principle

In *Miss Jay Jay,* [837] two concurrent perils also operated: a covered peril (bad weather in the form of rough seas) and an uninsured peril (a design defect or inherent vice). The insured could still recover because the second peril was uninsured rather than excluded. In Stapleton's view, the design defect might have been unnecessary in a strict but-for sense, perhaps the storm alone was enough to cause damage, but it still contributed. If so, excluding recovery simply because the design defect also played a part would ignore the fact that the policy specifically insured damage from bad weather. Thus, *Miss Jay Jay* aligns better with

[834] Stapleton (n 823) 42
[835] Stapleton (n 823) 45
[836] Wayne Tank (n 813)
[837] Miss Jay Jay (n 813)

Stapleton's argument: a non-essential but positive contribution from an uninsured factor need not ruin coverage if there is also a valid insured peril.

3.4. *Leyland Shipping* and The Search for The Real Effective Cause

In *Leyland Shipping Co v Norwich Union Fire Insurance Society*, a ship was torpedoed by a German submarine during wartime. Though damaged, it remained afloat and was taken to port, where it was delayed and ultimately sank due to the worsening of the damage. The insurer argued the proximate cause was not the torpedo but the decision to remain in port.,[838] a decision that highlights how courts look beyond factual causation, focusing on which cause is truly operative. The House of Lords distinguished between immediate and proximate causes, emphasising that while a torpedo strike triggered the sinking, other factors like sheltering in port also played a role. However, the submarine attack prevailed as the real effective cause of loss. Here we see alignment with the broader principle in multi-cause scenarios that the but-for test alone may not isolate the pivotal peril. It may not strictly be a concurrency issue, but it complements the concurrency rules advanced in *Wayne Tank* and *The Miss Jay Jay*, reinforcing the view that mechanical logic is still inadequate.

4. *Orient-Express* – The High-Water Mark of Strict But-For Analysis

An oft-cited modern controversy is *Orient-Express Hotels Ltd v Assicurazioni Generali SpA*.[839] The claimant's hotel in New Orleans suffered physical damage from Hurricanes Katrina and Rita, yet the rest of the city was also devastated, meaning tourism collapsed regardless of the hotel's state. Under the but-for approach, the High Court (upholding an arbitral tribunal) concluded that because there would still have been a drastic reduction in business even if the hotel had not been physically damaged, the insured peril (the property damage) did not cause' the revenue losses.[840] In effect, the wide-area damage overshadowed the specific property damage, leading to minimal recovery.

4.1. Broader Criticisms and Commercial Impact

The *Orient-Express* decision was widely criticised for defying commercial sense. If physical property damage was at least partly responsible for the business interruption, it appeared incongruous to deny recovery solely because another uninsured factor also contributed.

[838] Leyland Shipping Co v Norwich Union Fire Insurance Society [1918] AC 350 (HL).
[839] Orient-Express (n 814)
[840] Orient-Express (n 814) Song Causation (n 821) 71–72

Meggitt observes that insurers drew on *Orient-Express* as 'a second bite of the cherry … by the method for calculating the sum to be paid,' exploiting the presence of extensive, uninsured damage in the wider area.[841]

In essence, *Orient-Express* was the high-water mark of a purely factual but-for approach in concurrency disputes. Many scholars predicted (correctly) that it would be overturned or revised. As will be seen, later developments, especially *FCA v Arch*, exemplify a shift away from a rigid but-for analysis when multiple perils overlap.

4.2. Critiquing Orient-Express Through Stapleton's Lens

Stapleton's perspective helps explain criticism of *Orient-Express*.[842]There, the court said the hotel's loss of business would have happened anyway due to widespread hurricane damage, so the specific property damage did not matter for insurance. However, Stapleton's view asks whether the local physical damage still contributed meaningfully. If so, ignoring it as unnecessary may be too narrow. Courts that rely solely on the but-for test risk missing real, though not strictly essential, causes. This broader notion, recognising that unnecessary causes can matter, suggests *Orient-Express* might have overlooked or discounted a cause that was genuinely at play. Stapleton might label *Orient-Express* a prime example of unnecessary cause analysis done incorrectly: the unnecessary wide-area damage was enough to fail the but-for test, even though the property damage contributed meaningfully to lost custom.[843] This reflects Stapleton's emphasis on recognising any factor contributing to the outcome, even if it is not strictly decisive under a mechanical but-for analysis.

5. The Supreme Court's Rejection of Rigid But-For Logic: *FCA v Arch*

5.1. The Supreme Court's Intervention

In early 2021, the Supreme Court delivered *Financial Conduct Authority v Arch Insurance (UK) Ltd*, ruling on business interruption (BII) claims arising from the COVID-19 pandemic.[844] Many policy wordings required proof that the insured peril, often phrased as a local outbreak of disease, was the cause of the interruption. Insurers contended that even if the disease existed within a 25-mile radius, the nationwide pandemic overshadowed it, so

[841] G Meggitt, 'Business not as usual - The Financial Conduct Authority v Arch and Others' (2022) 4 JBL 257, 271.
[842] Orient-Express (n 814)
[843] Stapleton (n 823) 45
[844] FCA v Arch (n 815)

but-for that local outbreak, the losses would still have occurred from the broader national lockdown.

The Supreme Court unanimously rejected this argument, thus overruling *Orient-Express*. The Court explained that the but-for test, while relevant in most ordinary cases, 'fails to exclude a great many cases … [and] excludes some cases where one event could or would be regarded as a cause of another event'.[845] Indeed, Lords Hamblen and Leggatt observed that but-for is simply an inadequate threshold test, especially where multiple concurrent events each can be said to cause the loss. If one local outbreak is an equal, effective cause alongside a global outbreak, it cannot be dismissed solely because it fails but-for. The result: coverage was triggered if there was a local outbreak within the radius, even though there would have been business interruption from the wider pandemic anyway.[846] This outcome signalled a decisive departure from the narrow analysis in *Orient-Express* and underlined the limitations of but-for reasoning in multi-cause disputes.

5.2. Overruling *Orient-Express*

The Supreme Court addressed *Orient-Express* directly, holding that the tribunal and High Court had erred by applying a strict but-for approach. Lord Briggs and Lord Hodge emphasised that if two proximate causes arise from the same underlying fortuity (e.g. the hurricanes in *Orient-Express* or COVID-19 in *FCA v Arch*), the but-for test is too blunt an instrument.[847] Indeed, circumstances with the same underlying or originating cause' must be stripped from the counterfactual.[848] In other words, one cannot imagine a hypothetical city unaffected by hurricanes while simultaneously acknowledging that the same hurricanes destroyed the subject matter insured.

As *Colinvaux's Law of Insurance* summarises, The Supreme Court refused to countenance that possibility and overruled *Orient-Express* 'The correct outcome, in that case, should have been that the hurricanes led to business interruption proximately caused by material damage … The existence of the latter [wide-area damage] did not detract from the former.'[849] This reversal demonstrates unequivocally that the highest court in English law has rejected a

[845] FCA v Arch (n 815) 181–182
[846] FCA v Arch (n 815) 182, 319
[847] FCA v Arch (n 815) 310
[848] FCA v Arch (n 815) 309 - 310
[849] Merkin and Colinvaux (n 818) para 5 - 101

purely mechanical but-for test in multi-cause insurance disputes.[850] Consequently, the overruled decision now stands as a clear example of how courts should not assess causation in cases involving overlapping perils.

6. The Two Pillars – Over/Under-Inclusiveness and Policy Interpretation

The *FCA v Arch* reasoning rests on two pillars. Firstly, the but-for test is over- and under-inclusive. As the Court put it, the but-for test cannot be treated as the exclusive test because 'it is likely to throw up false positives and false negatives'.[851] Two independently sufficient causes each fail 'but-for' if the other cause would still have produced the same damage, yet common sense insists each is a real cause. This is well illustrated in *Cook v Lewis* [1951], where two hunters negligently fired their rifles at the same time, and a third person was injured by one of the bullets. It was impossible to prove whose shot caused the injury, so neither shooter satisfied the but-for test individually. Nonetheless, the Supreme Court of Canada held that one or both must be liable, recognising the inadequacy of strict but-for reasoning in such overdetermined scenarios.[852]

Secondly, policy interpretation outranks a strict but-for test. *FCA v Arch* holds that 'the question whether particular consequential harm to a policyholder is subject to indemnity is … a matter of interpreting [the] bargain.'[853] If an insured peril stands as a proximate cause, coverage is not negated simply because an uninsured cause also contributed. Where a policy frames an outbreak within twenty-five miles as insured, the concurrency with external outbreaks is overshadowed by reading them as part of a single underlying event, thereby preserving coverage.

Meggitt comments that the most significant implication for UK-based insurers is a 'change in the former's response to pandemic-related claims' and that many 'previously rejected claims might now become valid or expanded in value.'[854] This underscores how the Supreme Court's concurrency approach yields practical changes: insurers cannot disclaim liability merely by pointing to additional uninsured events if the insured cause is also proximate. The

[850] FCA v Arch (n 815) [181]
[851] FCA v Arch (n 815), para 181
[852] *Cook v Lewis* [1951] SCR 830.
[853] FCA v Arch (n 815) para 320
[854] Meggitt (n 841) 269

upshot is that *FCA v Arch* affirms the essay's proposition: the but-for test alone is inadequate for deciding multi-cause insurance losses under English law.

7. Scholarly Perspectives and Critical Assessment

7.1. Is The But-For Test Dead or Merely Curtailed?

The question remains whether the but-for test is altogether dead. Plainly, the courts continue to use it in single-cause scenarios or basic factual inquiries. If a building were destroyed only by an excluded peril or if it would have burned down irrespective of the insured peril, the but-for test excludes coverage. As the Supreme Court stated, the but-for test is 'often … described as a minimum threshold test of causation', albeit inadequate in multi-cause contexts,[855] meaning that if a cause fails but-for, it almost certainly cannot be regarded as the main cause. *FCA v Arch* merely acknowledges that, in concurrency contexts, the but-for test cannot alone decide matters.

Song observes that although English courts often treat the but-for approach as a straightforward way to identify factual causation, that method 'has … not been effectively defined and developed' and can yield inadequate results where more than one peril operates on equal terms.[856] Such concurrency questions are often sidestepped by focusing on a single proximate cause or re-casting the matter as an interpretive exercise.[857] Indeed, *Orient-Express* arguably illustrated those shortcomings by applying a strict but-for analysis to a multi-factor scenario, limiting coverage because the same hurricane destroyed the surrounding area.

Stapleton illustrates the weakness of the but-for test through an example involving three people, each negligently leaning on a car parked on a mountainous lookout. Individually, none exerts enough force to push the car over a small curbstone, but any two combined can do so, thereby causing the car to roll down the mountain. Under a strict but-for requirement, no single individual would be deemed a factual cause, even though, collectively, all three precipitated the car's destruction and ought to be recognised as causal contributors.[858]

[855] FCA v Arch (n 815815) para 181
[856] Song Revisiting (n 825) 458 - 465
[857] Song Revisiting (n 825) 458 - 465
[858] Stapleton (n 823) 39 - 40, 45

Despite these criticisms, I disagree with any suggestion that the but-for test has no continuing function. For straightforward claims lacking concurrency, it remains the first factual filter. Clarke notes, ' In England the proximate cause is said to be the efficient, effective or dominant cause.'[859], and if a single event stands out, the but-for approach typically suffices to confirm or refute it. Indeed, the Supreme Court in *FCA v Arch* acknowledged that 'in the vast majority of … cases … if Y would have occurred anyway irrespective of X, then X cannot be said to have caused Y'.[860] That is a direct nod to but-for logic. The real judicial recognition is that but-for is insufficient in a narrower but crucial class of multi-cause disputes.

7.2. The Inescapable Role of Judicial Interpretation

A deeper reflection arises on how judges approach concurrency. Lord Hoffmann famously argued that courts focus less on abstract causation doctrines than on interpreting legal rules 'involving the limits of liability,' which Song discusses in her examination of his views.[861] She observes that, according to Hoffmann, 'judges follow a natural process … the interpretation of the rule goes first, before deciding whether the facts satisfy those requirements'.[862] This emphasis on interpretation resonates with concurrency disputes, where clauses on exclusions or trends ultimately decide coverage rather than a purely philosophical stance on causation.

In that sense, *Wayne Tank* and *The Miss Jay Jay* can be read primarily as exercises in policy interpretation rather than concurrency doctrine. Indeed, Song contends that '*Wayne Tank* and *The Miss Jay Jay* seem to have been misinterpreted … Instead of making principles, [they] are merely cases exemplifying judicial interpretation of contractual terms and the intentions of the contracting parties'.[863] There is merit to this viewpoint because *Wayne Tank* turned on an exclusion, but it also yielded concurrency 'rules' still invoked by courts. They are not wholly misread; both do illustrate the core concurrency approach, yet more recent judgments, notably *FCA v Arch*, reaffirm that interpretation of the policy, rather than a rigid concurrency formula, is ultimately what governs how English courts reconcile multiple causes.

[859] M Clarke, *Law of Insurance Contracts: Chapter 25 Causation* (Maritime Insights & Intelligence Ltd) 25–3.
[860] FCA v Arch (n 815) para 181
[861] Song Revisiting (n 825) 459 - 460
[862] Song Revisiting (n 825) 459 - 460
[863] Song Revisiting (n 825) 473 - 474

7.3. The Interplay of Judicial Interpretation and Evolving Doctrine

A striking example of how legal doctrine can both anchor and adapt judicial reasoning is the Supreme Court's reversal of *Orient-Express* by two judges who had originally endorsed it. While Meggitt laments that 'such judicial backtracking … does little to reinforce one's faith in the consistency of judicial reasoning,'[864] it also highlights the law's capacity to evolve. In *FCA v Arch*, those same judges explicitly noted, in wonderful prose, their willingness to 'gracefully and good naturedly' surrender 'former views to a better considered position'.[865]

Far from destabilising the law, this reflects a necessary flexibility. Rigid adherence to tests like but-for can yield commercially absurd results, especially amid unprecedented events such as a global pandemic. As Lord Denning once wrote: 'I will do justice, not I will do law,'[866] implying that doctrine must remain attuned to practical outcomes rather than exist as a mere formula.

Yet critics worry that too much judicial discretion risks decisions turning on 'what the judge ate for breakfast'.[867] The key is preserving core principles, ensuring judges do not rely on unstructured common sense while allowing doctrines to be refined over time. *FCA v Arch* illustrates that a purely mechanical but-for approach can fail to capture parties' commercial intentions. The Supreme Court's contextual reading of policy wordings shows how established doctrine and contractual construction must converge to yield coherent, equitable results.

7.4. Evaluating Song's Thesis: Largely Theoretical Concurrency?

While Dr Song suggests that *Wayne Tank* and *The Miss Jay Jay* have been overly generalised in English insurance law, her assertion that concurrent causation remains 'a largely theoretical analysis … [such] that courts have … not been effectively defining and applying it'.[868] The Supreme Court's decision in *FCA v Arch* shows that concurrency has evolved beyond mere theory, prompting judges to grapple directly with multi-factor scenarios and to overrule *Orient-Express*.[869] If concurrency were still purely theoretical, one would expect

[864] Meggitt (n 841) 281
[865] FCA v Arch (n 815) [312]
[866] Rt Hon Lord Denning, The Road to Justice (Sweet & Maxwell 1988) 4
[867] Meggitt (n 841) 281; see also Dan Priel, 'Law Is What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea' (Osgoode Hall Law School, York University, SSRN abstract no 2982716) https://ssrn.com/abstract=2982716 accessed 18 December 2024
[868] Song Revisiting (n 825) 464 - 465
[869] FCA v Arch (n 815)

the Court to avoid redrawing the causal boundaries of business interruption coverage. Instead, *FCA v Arch* underscores that the notion of multiple operative causes, whether interdependent or independent, can prove critical in deciding actual policy outcomes, particularly where a covered peril combines with other events (pandemic or otherwise). The Court's firm rejection of the but-for approach in *Orient-Express* indicates that concurrency analysis is no longer sidelined by simply classifying one cause as dominant or recasting the inquiry as interpretation alone. Consequently, while concurrency might once have been underdeveloped*, FCA v Arch* has propelled it into a more concrete doctrine that obliges insurers and policyholders alike to consider how multiple concurrent perils interact in the real world.

Conclusion

1. Judicial Recognition of But-For's Inadequacy and Synthesis of Scholarly Perspectives

The but-for test remains a fundamental baseline in simpler single-peril contexts, although an unqualified reliance on it is indeed inadequate once multiple concurrent or overdetermined events enter the frame. This inadequacy was dramatically illustrated in *Orient-Express*, where the strict but-for approach essentially negated coverage for property damage on the basis that the wider, uninsured cause (city-wide devastation) would have reduced business anyway. The Supreme Court in *FCA v Arch* explicitly overruled this logic, clarifying that the but-for test can produce over and under-inclusive outcomes. The new stance is that concurrency is addressed by reading the policy's language and intentions, ensuring that if an insured cause is proximately involved, coverage is not automatically lost just because another uninsured cause is also sufficient to produce the damage. As the Supreme Court emphasised, 'the question whether particular consequential harm to a policyholder is subject to indemnity … [is] a matter of interpreting their bargain.'[870]

Legally and conceptually, concurrency challenges illustrate the weakness of the but-for test when multiple events each can cause the same harm.

Academic consensus, seen in works by Stapleton, Song, Clarke and Meggitt, favours a proximate cause approach shaped by contractual definitions rather than an unconditional

[870] FCA v Arch (n 815) para 320

but-for method. Although some scholarly views differ on the breadth of *FCA v Arch's* practical effect, the Supreme Court's explicit rejection of *Orient-Express* is no small development. It cements the principle that it is insufficient to show the insured peril fails the but-for test once concurrency arises. Otherwise, an insurer might escape liability in nearly every wide-area or pandemic scenario.

Ultimately, judicial recognition of the but-for test's inadequacy for multi-cause insurance losses is clear. The test persists as an elementary tool in single-cause or straightforward factual scenarios. Still, English law now unequivocally holds that it cannot be the final arbiter of causation in overdetermined claims.

This emphasis on proximate cause has long-standing roots, apparent from *Leyland Shipping,* where the House of Lords favoured the 'real effective' cause over a purely factual analysis.[871] This lineage, confirmed by *FCA v Arch*, ensures that coverage disputes are resolved in line with the parties' commercial expectations, something pure but-for logic too often failed to achieve.

2. Looking Ahead – Significance and Future Implications

This article aims to contribute to the ongoing review of concurrency in insurance with its focus on how concurrency disputes underscore the limitations of the but-for test in modern insurance practice. The stakes are high: global catastrophes such as the COVID-19 pandemic have highlighted scenarios where multiple overlapping causes operate, prompting courts to adopt a more nuanced, proximate-cause framework. There recent developments have immediate practical relevance. Insurers may now recognize the need for clearer exclusions or additional policy requirements specifically addressing wide-area damage or pandemic-driven losses, while policyholders and brokers can better negotiate insurance terms that reflect genuine commercial risks. Policymakers, too, have a need to examine how concurrency affects the stability of entire industries, potentially guiding new regulatory measures to ensure fair underwriting and consistent coverage outcomes.

The Supreme Court's stance in *FCA v Arch* has potentially wide-ranging implications. Insurers might redesign policy language to clarify how concurrent or wide-area events impact coverage, and we will wait to see if courts see further disputes testing the boundaries

[871] Leyland (n 838)

of concurrency, particularly as climate-related or systemic losses grow in frequency and scale. By moving away from rigid but-for logic, English law has progressed toward a more commercially realistic understanding of causation, though the shape of that evolution remains open to further judicial refinement. Ultimately, this realignment underscores not just the theoretical importance of concurrency doctrine, but its ongoing and practical relevance to everyday insurance disputes.

BIBLIOGRAPHY

Cases

Cook v Lewis [1951] SCR 830

Cork v Kirby Maclean Ltd [1952] 2 All ER 402 (CA)

Financial Conduct Authority v Arch Insurance (UK) Ltd [2021] UKSC 1, [2021] Lloyd's Rep IR 63

JJ Lloyd Instruments Ltd v Northern Star Insurance Co Ltd (The Miss Jay Jay) [1987] 1 Lloyd's Rep 32 (CA)

Leyland Shipping Co v Norwich Union Fire Insurance Society [1918] AC 350 (HL)

Orient-Express Hotels Ltd v Assicurazioni Generali SpA [2010] EWHC 1186 (Comm), [2010] 1 CLC 847, [2010] Lloyd's Rep IR 531

Wayne Tank & Pump Co Ltd v Employers' Liability Assurance Corp Ltd [1974] QB 57 (CA)

Legislation

Marine Insurance Act 1906

Articles and Books

Clarke M, *Law of Insurance Contracts: Chapter 25 Causation* (Maritime Insights & Intelligence Ltd)

Denning A T, *The Road to Justice* (Sweet & Maxwell 1988)

Meggitt G, 'Business not as usual—The Financial Conduct Authority v Arch and Others' (2022) 4 JBL 257

Merkin R and Colinvaux R (eds), *Colinvaux's Law of Insurance* (13th ed, 2024)

Priel, Dan, 'Law Is What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea' (Osgoode Hall Law School, York University, SSRN abstract no 2982716) <https://ssrn.com/abstract=2982716> accessed 7 January 2025.

Song M, ‘Revisiting concurrent causation and principles in English insurance law: a legal fiction?’ (2021) 6 JBL 457

Song M, *Causation in Insurance Contract Law* (2nd ed, Informa Law from Routledge 2024)

Stapleton J, ‘Unnecessary Causes’ (2013) 129 LQR 39

Fraudulent Claims in Marine Insurance: From Strict Forfeiture to A Balanced Legal Approach

Xinyue Zhang

Abstract

Fraudulent claims are a key factor in marine insurance, often determining whether the insured can receive compensation. The article explores fraudulent claims in marine insurance, emphasizing their legal impact and evolving judicial attitudes. Initially, courts took a strict stance, leading to full claim forfeiture. However, the *DC Merwestone* introduced the concept of 'collateral lies', distinguishing them from outright fraud. This marked a shift toward a more balanced approach, as reflected in the Insurance Act 2015, ensuring fairness while maintaining strong anti-fraud measures.

Introduction

A fraudulent claim is one of the most important factors in marine insurance, which usually influences whether the insured could obtain insurance compensation. Ideally, insurance should be a valid tool for persons and businesses to manage risk. However, with the development in the insurance business, fraud in insurance brings more instability and uncertainties to some extent. The Insurance Fraud Task Force pointed out the following statement in January 2016: Insurance fraud is a serious problem, estimated to cost £50 to each insurer and over £3 billion to the country each year. The living cost for everyone has increased since the influence of fraudulent insurance claims, which caused the price increase in essential products, for example, car insurance, health insurance, and home insurance.[872] It is believed that fraud is an important problem that needs to be solved, and that is the reason why policymakers continue to reform the law and provide more clarifying explanations. This essay will explain why the majority decision in *the DC Merwestone* invalidate the words that Lord Hobhouse said in *the Star Sea*: 'The fraudulent insured must not be allowed to think: if the fraud is successful, then I will gain; if it is unsuccessful, I will lose nothing' by the first focus on the true nature of the utmost good faith and the law on fraudulent claims, then identify what is a collateral lie and the difference between the fraudulent claims and

[872] Introductory words from the Foreword of the Final Report of the Insurance Fraud Task Force, January 2016: The Report is available at <https://www.gov.uk/government/publications/insurance-fraud-taskforce-final-report>(accessed on 5 January 2024).

collateral lie. The attitude of fraud by courts in many cases will be mentioned to prove that this change meets the needs of the marine insurance business development.

1. Fraudulent Claim: Theoretical Starting Point

The good faith and fraud are like the two sides of a coin. The premise of understanding the true meaning of fraud is to clarify the definition of the good faith. The Marine Insurance Law 1906, section 17, does not attempt to give an explanation of the "utmost good faith". However, it is believed that this cannot be regarded as a legislative defect: good faith is a concept that evolves with the times, and it should have flexibility and space for interpretation, rather than being defined by specific words.[873] If the case law tries to shed light on the meaning of utmost good faith, it should start with the pre-contractual duty to inform, due to the decision of *the Carter v Boehm* on the duty of utmost good faith was primarily concerned with the duty to disclosure and misrepresentation by the insured to the insurer.[874] After that, there is a long debate about the meaning and scope of the duty of utmost good faith. However, it is known that the duty of utmost good faith under English law is not limited to the pre-contractual stage, although there is more discussion required on the different meanings of it in the pre-contractual and post-contractual stages. The following part will focus on the post-contractual duty of utmost good faith.

1.1. Good Faith and Fraudulent Claim

In *The Star Sea*,[875] Lord Clyde made it clear that good faith did apply to the post-contract stage. He stated that while the special provisions immediately following section 17 could refine the general rules applicable to the stage of contract formation, they could not exhaust them. In this case, an alternative approach could be chosen, which means that a flexible interpretation of the scope of utmost good faith could be used in the post-contractual stage. The case of *Goshawk v Tyser*[876] concretized good faith and credit into an absolutely non-actionable obligation at the level of principle.[877]

[873] Rhidian Thomas (ed), *The Modern Law of Marine Insurance Volume 2* (LLP 2002) 30.
[874] (1766) 3 Burr. 1905.
[875] *Manifest Shipping Co Ltd v Uni-Polaris Insurance Co Ltd (The Star Sea)* [2001] UKHL 1, [2001] 1 Lloyd's Rep 389.
[876] *Goshawk Dedicated Ltd and others v Tyser & Co. Limited and another* [2006] EWCA Civ 54, [2007] Lloyd's Rep IR 224, per Rix LJ at [53].
[877] Hjalmarsson Johanna, 'Fraudulent insurance claims: legal definition and judicial consequences' (PhD thesis, University of Southampton 2016).

The explanation of the utmost good faith is required in many cases with the development of the marine insurance business. The opinion that the duty of utmost good faith should have different meanings and scopes in the stages before and after the conclusion of the contract has been widely approved.[878] Section 17 of the Marine Insurance Act 1906 was not retained in the Insurance Act 2015, the latter clarifying the good faith to a clause that has no function other than an interpretation principle. Therefore, under the Insurance Act 2015, good faith has become an interpretive principle. Peter explained this change in the following statement in his article: 'The pre-contractual and post-contractual obligations applicable to claims will have the same objective, which is to prevent fraud.'[879] Here, the fraud comes to the stage.

1.2. The Strict Attitude Toward Fraudulent Claims

The attitude towards fraudulent claims was undergoing a significant change. In *Britton v. Royal Insurance Co.*,[880] the court ruled that where a person fraudulently exaggerates his claim, not only the exaggerated part will be denied, but the whole of his claim will be forfeited. The court held that if allowed a party to attempt such fraud in the post-contractual stage, it would be a very dangerous signal. Allowing individuals to engage in deceptive practices during insurance claims poses a significant threat to the system.[881] If such behaviour is overlooked and they are still able to receive compensation equivalent to the actual value of the items lost or damaged, it would set a dangerous precedent. Furthermore, if it is discovered that there has been a deliberate act of deception or fraudulent activity in filing the claim, the individual's right to any insurance payout should be completely voided.[882]

This kind of zero-tolerant attitude to fraudulent claims was confirmed in *the Star Sea*, Lord Hobhouse states, 'The fraudulent insured must not be allowed to think: if the fraud is successful, then I will gain; if it is unsuccessful, I will lose nothing.' These words emphasize a strict approach to insurance fraud, implying that any fraudulent act by the insured, regardless of its material impact on the claim, should lead to the claim being forfeited. The

[878] Gerald Swaby, 'A critical examination of the disproportionate rights and duties of insurers and insured visà-vis good faith, fraud and the settlement of insurance claims' (Doctoral Thesis, University of Huddersfield 2016).
[879] Peter MacDonald-Eggers, 'Remedies for the failure to observe the utmost good faith' [2003] LMCLQ 249, 254.
[880] (1866) 4 F & F 905.
[881] Ibid.
[882] (1866) 4 F & F 905, 909.

rationale behind this principle is to deter fraudulent behaviour by making the consequences of such behaviour unequivocally negative for the insured.

1.3. Different Opinions on Fraudulent Claims

Although the forfeiture rule has rationality, it differs from the principles under the Marine Insurance Act 1906, which ruled that a fraudulent claim would result in the contract being avoided by the other party. Lord Longmore in the Mercandian Continent specified that good faith after the conclusion of the contract applies in four circumstances, the fraudulent claim is one of them.883 What's more, in the Star Sea, the House of Lords clarified limitations on the application of section 17 under the Marine Insurance Act 1906.884 The courts determined that the insured's duty to observe utmost good faith stopped with the commencement of legal proceedings. Furthermore, the obligation of the assured to disclose material risk facts to the insurer ends once the insurance contract is finalized and binding. Following this, the strict requirements for presenting a case and the admission of evidence in situations of fraudulent claims are no longer tied to the principle of utmost good faith. In English judicial practice, there was a notable shift regarding the application of the principle related to fraudulent claims initially set out in *the Star Sea*. In the subsequent *Aegeon (No1)* case, it was determined by Lloyd Manns that fraudulent claims should be subject to common law rules rather than section 17 of the Marine Insurance Act 1906.[885] This meant that the assurer could not simply void the contract based on a fraudulent claim. Differing opinions emerged on this topic. The authors of McGillivray on the Law of Insurance suggested that insurers could choose between the "forfeiture rule", where the claimant loses the claim, and the "avoidance rule", where the contract can be avoided by the assurer entirely.

On the other hand, Professor Clarke advocated for a singular approach: a fraudulent claim should fail entirely, and the assurer should have the right to rescind the contract.[886] However, any honest claims made prior should be upheld, and the insurer should not be able to reclaim sums paid on past honest claims. Judge Mance echoed and supported Clarke's view in *Axa General Insurance Ltd. v. Gottlieb*, affirming that collateral lies fall under a special common law rule.[887] It shows that the methods of dealing with fraudulent claims in the context of

[883] *K/S Merc-Scandia v Certain Lloyd's Underwriters and Others* [2001] 2 Lloyd's Rep 563.
[884] *Manifest Shipping Co Ltd v Uni-Polaris Insurance Co Ltd (The Star Sea)* [2001] UKHL 1, [2001] 1 Lloyd's Rep 389.
[885] *Agapitos v Agnew (The Aegeon)* [2002] EWCA Civ 247; [2002] 2 Lloyd's Rep 42.
[886] Malcolm Clarke, 'Lies, Damned Lies and Insurance Claims' (2000) NZLRev 233, 233-261.
[887] *Axa General Insurance Ltd v Gottlieb* [2005] EWCA Civ 112.

insurance contracts have become more sophisticated. Another perspective comes from European civil law jurisdictions, where insurance fraud is often addressed with more proportional penalties. In Germany and France, for instance, fraudulent claims may result in partial denial rather than full forfeiture, particularly if the insured party can prove that a significant portion of the claim remains valid. These systems focus on the principle of proportionality, ensuring that penalties align with the severity of the fraudulent act.[888]

The debate on the meaning and scope of fraudulent claims and the obligation of utmost good faith reflects a shift in attitudes toward fraudulent claims. In *the Star Sea*, the Lords clarified that the utmost good faith applies not only before but also after the conclusion of an insurance contract. Nowadays, the fraudulent claim rule in common law is more balanced. While outright fraudulent claims are still treated harshly, there's a clearer distinction made between fraudulent claims and collateral lies. The law now acknowledges the need for proportionality and fairness, ensuring that minor misrepresentations or irrelevant falsehoods do not lead to disproportionately severe consequences. Thus, the debate on fraudulent claims continues to evolve, reflecting a broader shift toward a more balanced and equitable legal framework in marine insurance. The tension between deterring fraud and ensuring fairness remains a central concern for courts and policymakers worldwide.

2. Collateral Lies: When Lies are not Fraud

With the development of the insurance business, there are increasingly more complex situations in practice. If all fraud is regarded as fraudulent claims, it may lead to more unequal consequences. In the case of S*hah v UlHaq*,[889] there was an incident where S accidentally collided with W's vehicle. Subsequently, W and his wife made a legitimate insurance claim for the actual injuries sustained. They also prompted K to pretend to have been in the car during the accident. K, however, was not actually present in the vehicle at the time of the collision, leading to suspicions that W was instigating a fraudulent claim.[890] The Court of Appeal noted that the fraudulent claims rule applies only to both parties to the insurance contract, which means that W's claims are not affected by any fraudulent acts

[888] Richard Aikens, 'When is a "fraudulent claim" only a "collateral lie"?'[2023] Lloyd's Maritime & Commercial
Law Quarterly 340 < https://www.i-law.com/ilaw/doc/view.htm?id=379392#CLQ:20170339.21 > accessed 25 December 2023.
[889] [2009] EWCA Civ 542.
[890] John Lowry and Philip Rawlings, 'Insurance Fraud: The 'Convoluted and Confused' State of the Law' (2016) The Law Quarterly Review 132 (January) 96-119.

made by W in respect of K's claims.[891] Whether or not the law is good at deterring such conduct and whether or not civil courts aim to punish it, the inflexible nature of fraudulent claims rules has been cited.

The Law Commission acknowledged the importance of the judicial stance that clear penalties are necessary to prevent policyholders from engaging in fraudulent activities. However, they also noted the current laws were overly complex and unclear.[892] They advocated for solutions that would offer insurers remedies that are grounded in principle, proportionate to the issue, and reliable. In a similar vein, in *the DC Merwestone,*[893] Popplewell J. observed in the initial ruling that the existing rules could lead to excessively severe and unfair outcomes for the assured, which could inadvertently benefit an insurer that does not deserve such an outcome.[894] What kind of a claim can be regarded as related to the claim, and what is not related to the claim requires more analysis and explanation. Therefore, how to clarify a collateral lie has become an important part of marine insurance.

2.1. *The DC Merwestone*

In the case of *the DC Merwestone*, the trial judge determined that the sinking of the vessel was a covered event under the marine insurance policy, attributing the loss to a sea peril. The shipowners were found to have fulfilled their duty of due diligence, negating the insurer's argument that the vessel was 'unseaworthy' and thus not covered under the terms provided by the Marine Insurance Act 1906. Despite this, an untruth was told by the insured's representative about how water entered the ship during the claims process.[895] Gilman and others have defined what constitutes a fraudulent device, describing it as any untruthful statement or falsified evidence, such as a fake document, used to bolster an otherwise valid claim. The discussion that follows will delve into the concept of the fraudulent device, also known as a collateral lie.[896]

[891] Julie-Anne Tarr, 'Grappling with fraudulent insurance claims and 'collateral lies': Comparative insurance law developments in the United Kingdom and Australia' (2019) Journal of Business Law 2019(1) 43-61.
[892] John Lowry and Philip Rawlings, 'Insurance Fraud: The 'Convoluted and Confused' State of the Law' (2016) The Law Quarterly Review 132 (January) 96-119.
[893] [2013] EWHC 1666 (Comm)
[894] [2013] EWHC 1666 (Comm).at [167]
[895] Julie-Anne Tarr, 'Grappling with fraudulent insurance claims and 'collateral lies': Comparative insurance law developments in the United Kingdom and Australia' (2019) Journal of Business Law 2019(1) 43-61.
[896] Jonathan Gilman, Robert Merkin, Claire Blanchard, and Mark Templeman: *Arnould's Law of Marine Insurance and Average* (1st supp, 17th edn, Sweet & Maxwell 2008).

2.2. Judgement of The United Kingdom Supreme Court

In a surprising turn, the UK Supreme Court, except Lord Mance, showed empathy towards the insured party in their judgment.[897] The majority redefined the concept of fraudulent devices, terming them as 'collateral lies', which implies that such lies do not undermine a valid claim. Lord Sumption distinguished between a fraudulent claim and a collateral lie by stating that in a fraudulent claim, the dishonesty is aimed at gaining something the claimant is not entitled to, whereas in a collateral lie, the claimant seeks nothing beyond what is legally due, and the lie does not pertain to the claim's validity or value. Lord Hughes provided what could be considered a definition of collateral lies in paragraph 92 of his judgment in *the DC Merwestone* case: 'A collateral lie is told with the intention of improving the liar's position, but it does not, in fact or law, affect the legitimacy of the claim, whether the lie is believed or exposed.'[898]

However, Lord Mance differed in opinion from the majority in the UK Supreme Court regarding the use of collateral lies in insurance claims, maintaining that forfeiture should be the consequence of employing such deceit.[899] He emphasized the role of the fraudulent devices rule in promoting honesty and deterring fraud within the claims process. He pointed out that fraud in claims often involves either (a) the fraudulent pursuit of a non-existent or invalid claim, (b) the fraudulent exaggeration of a valid claim, or (c) the pursuit of what might be a dubious claim, which the insured believes or fears to be questionable.[900] Lord Mance articulated that the fraudulent devices rule prevents a claimant from embellishing a false or exaggerated claim with lies, and then, upon discovery, simply switching to another deceitful tactic or lying in court to advance their spurious claim. In situations where the claim is questionable, this rule also discourages the insured from manipulating the claims process and hinders them from stopping the insurer from exploring or challenging the claim's weaknesses—prompting the lie in the first place. He argued that in any of these cases, using a fraudulent device that is material to the claim acts as a definitive bar to its continuation, eliminating the need for further investigation into the underlying facts and serving as a strong

[897] [2016] UKSC 45
[898] Poomintr Sooksripaisarnkit, 'Marine insurance–collateral lies: when lies are not fraud' (2017) Maritime Business Review 2.1, 52-56.
[899] James Dalton, 'Supreme Court ruling is setback in fight against fraudulent claims' (Insurance Times, 21 July 2016), <https://www.insurancetimes.co.uk/supreme-court-ruling-is-setback-inspurious-claim-fight-abi/1419111.article> accessed 18 January 2023.
[900] *Eagle Star Insurance Co. Ltd v Games Video Co. (GVC) S.A. (The Game Boy)* [2004] EWHC 15 (Comm).

deterrent against dishonesty. These protective measures, Mance contended, align well with the foundational principle of mutual trust in insurance.

3. Movement away from the Absolute Stance

An absolute stance usually leads to less flexibility, which may bring more questions when a rule is used in some facts. Lord Hobhouse's words in *the Star Sea* convey the judge's strict attitude towards fraudulent claims and the firm crackdown on fraudulent claims.[901] The fraudulent claims rule was designed to embed a punitive aspect within the legal framework, aiming to discourage the submission of dishonest claims by imposing strict penalties on such actions, also expected to make the insured become particularly careful when they make a claim. In this circumstance, it seems logical to take a draconian remedy to punish minor infractions. Over time, the broadening definition of fraud to encompass fraudulent means and devices has tipped the balance in favour of insurers. However, it's conceivable that expanding the fraudulent claims rule too far disrupts this balance, rendering a strict, punitive approach inadequate for all instances of fraudulent claims. This concern may trouble insurers, but the UK Supreme Court, through a majority decision, has clarified that the rule does not apply to collateral lies, signalling a move away from a one-size-fits-all penalty for all types of fraudulent behaviour in claims.

3.1. Proportionality Principle

The treatment of "collateral lies" in *the DC Merwestone* demonstrates respect for the principle of proportionality. Involving the collateral lies into the scope is a movement away from the strict stance that Lord Hobhouse mentioned in *the Star Sea.* This new trend was approved by most of the Lords. Lord Hughes highlighted a crucial distinction in insurance law between two types of dishonesty. On one hand, there are those who commit fraud with the intention of obtaining a payout to which they are not entitled, which is outright fraud. On the other hand, there are those who may tell a lie or present false evidence in support of a claim but do not stand to gain any more than what they are legally entitled to receive.[902] Furthermore, Lord Sumption provided insight into the motivations behind collateral lies, suggesting that they are often a result of the assured's frustration with the insurer's delay in

[901] Naraya Lamart, 'Certainty vs. Equity: A Case for Reform of the Duty of Utmost Good Faith?' (2018) 32 Austl & NZ Mar LJ 59 < https://www8.austlii.edu.au/au/journals/ANZMarLawJl/2018/10.pdf > accessed 15 Jaunary 2024.
[902] *Versloot Dredging* [2016] UKSC 45; [2017] A.C. 1 at [100].

acknowledging and settling claims. These lies are used to reinforce the claim's validity or to hasten the payment process, rather than to claim something beyond what is rightfully due.[903]

In *the DC Merwestone*, the Supreme Court distinguished between "fraudulent claims" and "incidental lies". In a fraudulent claim, the insured acts fraudulently to obtain compensation for which he or she is not actually entitled. In the case of incidental lies, the insured person is actually entitled to compensation, and the lies do not affect this right. The Supreme Court held that in the latter case, even if the insured had lied during the claim process, as long as the lie was not relevant to the legality of the claim, the right to claim should not be completely forfeited. This ruling reflects the principle of proportionality, which states that the legal consequences should be commensurate with the nature and severity of the insured's dishonest conduct. Previously, even non-critical lines could result in the insured losing all claims, which was considered too harsh. The Supreme Court's ruling means that in some cases, the law's penalties should be limited to ensure that penalties do not exceed the actual impact of dishonest conduct.[904] Therefore, the attitude toward solving the fraudulent claim has changed after the Versloot, since it is believed that in this case it is justified that the fraudulent claims rules do not apply. In practical terms, *the DC Merwestone* ruling has prompted the Marine insurance industry to reassess and modify the way it handles claims, especially when dealing with legitimate claims that contain incidental lies. The ruling spurred the development of Marine insurance law to combat fraud while also treating claims that are legitimate in nature more fairly.

3.2. Balancing of Interests

The common law's fraudulent claim rule, providing for the voiding of an insurance contract in cases of fraud, has been criticized for potentially being excessively harsh in some circumstances. This criticism stems from the realization that applying the rule rigidly can result in unfair outcomes, particularly when the dishonesty does not materially impact the claim's validity. This has prompted a shift towards a more discerning application of the rule, making a distinction between wholly fraudulent claims and minor untruths that do not compromise the overall claim.[905]

[903] ibid [3].
[904] *Versloot Dredging* [2016] UKSC 45; [2017] A.C. 1 at [100]; see also Lord Toulson at [106]–[108].
[905] Julie-Anne Tarr, 'Grappling with fraudulent insurance claims and 'collateral lies': Comparative insurance law developments in the United Kingdom and Australia' (2019) Journal of Business Law 2019(1) 43-61.

In Australia, the policymaker has made an effort to balance the interests of the assurer and the assured, which has been perceived as a tolerance for minor infractions. Australia's law of Insurance Contracts Act 1984, specifically section 56, allows the court to mandate that the insurer pay an amount it deems fair in cases where only a minor portion of the claim is fraudulent and withholding payment for the rest would be unduly harsh. This approach indicates an adjustment in defining what constitutes a fraudulent claim.[906] The Australian Law Reform Commission, while advocating for this legislative change, acknowledged that while fraud must be deterred, a rule that allows one fraudulent claim to contaminate other unrelated claims under the same policy may result in inconsistent outcomes between insured parties with multiple policies and those with a single policy covering several risks.[907]

In conclusion, attitudes toward fraud are evolving in both Australia and the UK, demonstrating that a more flexible approach could better serve the principles of marine insurance law. In the future, further refinements in marine insurance law may be necessary to ensure that fraudulent claims are penalized appropriately without disproportionately harming honest policyholders. Courts may increasingly consider factors such as the intent behind the fraudulent act, the financial impact on the insurer, and the overall fairness of claim denial when adjudicating disputes. This evolving perspective underscores the need for a dynamic legal framework that adapts to industry changes while maintaining a fair and effective approach to fraudulent claims.

3.3. Remedy

The approach to remedies for fraudulent claims in insurance law has developed over time. Originally, as stated in *Britton v Royal Insurance Co,*[908] the rule was that any willful fraud in a claim would result in the assured losing any right to the policy. This is based on the principle articulated by Lord Hobhouse in *the Star Sea*, emphasizing that fraud in an insurance claim should lead to the claim being forfeited. Later, in *the Mercandian Continent*, Aikens J clarified that the duty of utmost good faith under section 17 of the Marine Insurance Act 1906 does not continue after the formation of the contract unless there is a question of entering into a new contract.[909] Additionally, it was held that a forged document, which was

[906] Insurance Contracts Act 1984 (Australia) s.56(3).
[907] Australian Law Reform Commission, 'Report on Insurance Contracts, No.20' (1982), para. 243.
[908] (1866) 4 F&F 905.
[909] *K/S Merc-Scandia XXXXII v Lloyd's Underwriters (The Mercandian Continent)* [2001] EWCA Civ 1275; [2001] 2 Lloyd's Rep 563

not material to the assurer's decision on the claim, should not affect the ultimate liability of the assurer. The Court of Appeal further refined the insurer's remedy to the right to end the insurance contract in the event of breach by the assured, as opposed to refusing all claims under the policy. Then, in *Axa General Insurance Ltd v Gottlieb*,[910] the court concluded that insurers are not entitled to reclaim payments made on earlier, unrelated claims. This prevents insurers from revoking coverage retroactively for past claims that were not related to the fraud.

Nowadays, in the Insurance Act 2015, section 12 confirms by enactment the forfeiture rule: (1) The assurer is not obliged to pay the claim if the assured makes a fraudulent claim under an insurance contract. The fraudulent claim refers to any act by the insured to deceive the insurer to gain a payout to which they are not entitled. This includes exaggerated claims, fabricated claims, or any deceitful manipulation in the claim process. (2) The assurer may also recover any compensation that has been paid to the insured for the claim. This means the assurer has the right to take back any claim payments that were made based on a fraudulent claim. (3)The assurer has the option to treat the insurance contract as stopped from the time of the fraudulent act. If the insurer chooses to terminate the contract, they are no longer liable for any risk from the time of the fraudulent act. Still, they must return any premiums paid for the period after the termination. The section 12 of the Insurance Act 2015 reflects policymakers' firm stance against insurance fraud, which allows insurers to counter fraudulent activity by refusing to pay fraudulent claims and recovering any funds that have been dispersed under such claims. This change insists the attitude that deterring fraudulent claims and balances the interests of insurer and insured. While the Insurance Act 2015 has clear provisions for the treatment of fraudulent claims, it is supplemented by a more specific judicial interpretation of the treatment of collateral lies.[911]

To summarize, because of the consideration of the importance of fraud and the difference between fraudulent claims and collateral lies, the remedy for the former has been shifted from a strict forfeiture to a more nuanced approach. This evolution of attitudes and rules reflects an attempt to have a balanced and accurate approach to addressing insurance contract fraud.

[910] *Axa General Insurance Ltd v Gottlieb* [2005] EWCA Civ 112.
[911] Katie Richards, 'Time's up for wholly fraudulent insurance claims: the case for new statutory remedies' (2020) 7 Journal of Business Law 580 < https://orca.cardiff.ac.uk/id/eprint/125448/> accessed 17 January 2023.

Conclusion

This article tries to through analyze what the fraudulent claims and collateral lies are and the difference between them, to prove that *the DC Merwestone* has changed the attitude toward fraud in *the Star Sea*. The words of Lord Hobhouse in *the Star Sea,* 'The fraudulent insured must not be allowed to think: if the fraud is successful, then I will gain; if it is unsuccessful, I will lose nothing', emphasized the strict approach to fraudulent conduct by an insured. Under this zero-tolerant attitude, any fraudulent conduct should result in the loss of a claim, irrespective of its effect on the validity of the claim. However, in *the DC Merwestone*, the Supreme Court distinguished between a fraudulent claim and a collateral lie used in the pursuit of a legitimate claim. The majority of the Lords held that while a fraudulent claim should indeed invalidate the claim as a whole, a lie told in the course of pursuing a legitimate claim (which is not itself affected by the lie) should not necessarily result in the forfeiture of that claim. This means that a collateral lie told in the course of making a legitimate claim that does not affect the overall validity of the claim does not automatically invalidate that claim.

The decision of *the DC Merwestone* can therefore be seen as a softening of the harsh position expressed by Lord Hobhouse in *the Star Sea*, introducing a more nuanced approach to the different types of fraud in insurance claims.

Although this article tries to analyse why the fraudulent claims are different from collateral lies, there are still some limitations of this study's methodology. For example, it is hard to analyse every different case of fraud in practice due to the length limitation.

Ultimately, by researching fraudulent claims, we can see that the shift in judicial attitudes and legal frameworks demonstrates that marine insurance law is not static but adapts to economic realities, industry practices, and societal expectations. As legal doctrines continue to evolve, stakeholders—including insurers, policymakers, and courts—must work collaboratively to uphold both the integrity of the insurance system and the rights of honest policyholders.

BIBLIOGRAPHY

Articles and Books

Clarke M, 'Lies, Damned Lies and Insurance Claims' (2000) NZLRev.

Eggers M.P, 'Remedies for the failure to observe the utmost good faith' [2003] LMCLQ.

Gerald Swaby, 'A critical examination of the disproportionate rights and duties of insurers and insured visà-vis good faith, fraud and the settlement of insurance claims' (Doctoral Thesis, University of Huddersfield 2016).

Gilman J, Merkin R, Blanchard C, and Templeman M: *Arnould's Law of Marine Insurance and Average* (1st supp, 17th edn, Sweet & Maxwell 2008).

Introductory words from the Foreword of the Final Report of the Insurance Fraud Task Force, January 2016: The Report is available at https://www.gov.uk/government/publications/insurance-fraud-taskforce-final-report(accessed on 5 January 2024).

James Dalton, 'Supreme Court ruling is setback in fight against fraudulent claims' (Insurance Times, 21 July 2016), <https://www.insurancetimes.co.uk/supreme-court-ruling-is-setback-inspurious-claim-fight-abi/1419111.article> accessed 18 January 2023.

Johanna E, 'Fraudulent insurance claims: legal definition and judicial consequences' (PhD thesis, University of Southampton 2016).

Katie Richards, 'Time's up for wholly fraudulent insurance claims: the case for new statutory remedies' (2020) 7 Journal of Business Law 580
< https://orca.cardiff.ac.uk/id/eprint/125448/> accessed 17 January 2023.

Lowry J and Rawlings P, 'Insurance Fraud: The 'Convoluted and Confused' State of the Law' (2016) The Law Quarterly Review 132 (January).

Naraya Lamart, 'Certainty vs. Equity: A Case for Reform of the Duty of Utmost Good Faith?' (2018) 32 Austl & NZ Mar LJ 59
< https://www8.austlii.edu.au/au/journals/ANZMarLawJl/2018/10.pdf> (accessed on 15 January 2024)

Richard Aikens, 'When is a "fraudulent claim" only a "collateral lie"?'[2023] Lloyd's Maritime & Commercial Law Quarterly 340

< https://www.i-law.com/ilaw/doc/view.htm?id=379392#CLQ:20170339.21> accessed 25 December 2023.

Tarr J, 'Grappling with fraudulent insurance claims and 'collateral lies': Comparative insurance law developments in the United Kingdom and Australia' (2019) Journal of Business Law 2019(1).

Thomas R, *The Modern Law of Marine Insurance Volume 2* (LLP 2002).

Cases

Agapitos v Agnew (The Aegeon) [2002] EWCA Civ 247; [2002] 2 Lloyd's Rep 42.

Axa v Gottlieb [2005] EWCA Civ 112

Britton v Royal Insurance Co (1866) 4 F&F 905, 909

Carter v Boehm (1766) 3 Burr 1905

Eagle Star Insurance Co. Ltd v Games Video Co. (GVC) S.A. (The Game Boy) [2004] EWHC 15 (Comm)

Goshawk Dedicated Ltd and others v Tyser & Co. Limited and another [2006] EWCA Civ 54, [2007] Lloyd's Rep IR 224

K/S Merc-Scandia v Certain Lloyd's Underwriters and Others [2001] EWCA Civ 1275; [2001] 2 Lloyd's Rep 563

Manifest Shipping Co Ltd v Uni-Polaris Insurance Co Ltd (The Star Sea) [2001] UKHL 1

Shah v Ul-Haq [2009] EWCA Civ 542

Versloot Dredging BV and another (Appellants) v. HDI Gerling Industrie Versichering AG and Others (Respondents) [2016] UKSC 45

Versloot Dredging BV v HDI-Gerling Industrie Versicherung AG (The "DC Merwestone") [2013] 2 Lloyd's Rep. 131

Statutes

Insurance Act 2015

Insurance Contracts Act 1984 (Australia), section 56

Marine Insurance Act 1906