THE

OXFORD UNIVERSITY UNDERGRADUATE LAW JOURNAL

~

ISSUE EIGHT

TRINITY TERM

2019

THE EIGHTH EDITORIAL BOARD

THE HONORARY BOARD

Sir Nicolas Bratza

Professor Michael Bridge

Donald Findlay QC

Professor Christopher Forsyth

Ian Gatt QC

The Rt Hon. the Lord Judge

The Rt Hon. the Lord Kerr of Tonaghmore

Michael Mansfield QC

The Rt Hon. the Lord Neuberger of Abbotsbury

The Rt Hon. the Lord Phillips of Worth Matravers

The Lord Pannick QC

His Honour Judge Gordon Risius CB

Dinah Rose QC

Sir Konrad Schiemann

The Rt Hon. the Lord Wilson of Culworth

CONTENTS

FOREWORD (PRIVATE LAW)

Dr Sandy Steel
Fellow in Law of Wadhman College
Associate Professor of Law in the Oxford Faculty of Law

The articles in this volume examine complex topical issues in private law. Each article marshals interesting arguments and sheds light on the issue examined. Even those very familiar with the areas under scrutiny will find themselves with something to think about.

Two qualities of the articles stand out. First, they are appropriately critical: they follow lines of reasoning where they lead, even if this means full on disagreement with the views of judges and academics. Mehleen Rahman thoughtfully probes the validity of the UKSC's reasoning in *Prudential* on whether the value of the opportunity to use money received by a defective transfer is received at the payor's expense. Julia Brechtelsbauer analyses the current status in tort law of the framework for the illegality defence set out by the majority in *Patel v Mirza*; she is critical of the lower courts' treatment of *Patel* and persuasively doubts the cogency of differing approaches to the illegality defence across private law areas. Second, the depth of the analysis in each article is impressive. Antonia Kendrick's contribution weaves together close doctrinal analysis of an important UKSC decision on patent infringement, comparing the court's approach to that in other EU jurisdictions, with insightful discussion of its broader implications. Zach Pullar draws on tort theory and authorities from other common law jurisdictions to mount an interesting case for a mental incapacity defence in the tort of negligence.

I hope that the valuable contributions that comprise this volume of the journal encourage other student authors to think

deeply about private law, and law more generally, in the future. And then to write about it.

FOREWORD (PRIVATE LAW)

Professor Laura Hoyano
Senior Research Fellow in Law at Wadham College
Associate Professor of Law in the Oxford Faculty of Law
Barrister, Red Lion Chambers
Fellow of the Honourable Society of the Middle Temple

Sometimes, the deeper we academics and practitioners delve into our subject, the less we see; as we tread the same ground over the years, pondering the same issues, our feet wear a deep trench which allows only a glimpse of a small patch of sky. The intelligent, perceptive student, like those published here, interrogates the issues anew, acknowledging, but unconstrained by, precedent and scholarly argument. The Oxford undergraduate syllabus and tutorial system are designed to encourage that kind of thinking. Over my years of teaching and research, I have several times formally credited some undergraduates (two of them in their first year of studies) with ideas proposed and developed in our tutorial discussion (three of them still refer to "our footnote").

Accordingly, it is my pleasure and honour to have been invited to introduce the four private law essays in the eighth issue of the Oxford University Undergraduate Law Journal. Whilst the journal this year has opened up its submissions to undergraduates in law from all leading English research universities, the vast majority of the articles selected for inclusion in this issue were written by Oxford undergraduates reading for a three or four-year degree. They attest to their authors' intellectual ambition and to the incisiveness of their education.

I have learned much in reading the private law essays in this volume, which are all extended critiques of judgments from British courts, two from the Court of Appeal Civil Division, and two from the UK Supreme Court. Two essays are in my 'home

territory' of tort law. One, by Julia Brechtelsbauer, has tested my thinking about the illegality defence with an extended critique of refusal of the Court of Appeal to apply the reasoning in the contract case of *Patel v Mirza* to the tragic negligence case of *Henderson*, which is now on its way to the UK Supreme Court. The other, by Zach Pullar, analyses the objective test for the standard of care, highlighting apparent inconsistencies in approach between two Court of Appeal judgments, which can only be resolved by the Supreme Court. Whilst the topic of the recoverability of compound interest for wrongly paid taxes might seem arid to some – at least to those who have not tried to recoup monies from Her Majesty's Revenue – Mehleen Rahman, author of "Is Time up for Recovery of Time Value?", demonstrates the implications of the judgment of the UK Supreme Court in *Prudential Assurance* for the concept of 'enrichment' across the entire field of restitution for wrongs. The fourth essay, by Antonia Kendrick, demonstrates how a judgment of the UK Supreme Court widening patent protection through 'equivalents' confers excessively wide protection on patent-holders, potentially stripping researchers in many scientific fields of the protection of their 'inventive step' if it can be considered as falling within a broader conception of the original patent than originally applied for. This encourages 'gaming' of the patent system in a way which is seemingly condoned by the British courts but not by other European jurisdictions. The author suggests that this ruling may jeopardise incremental research in the industries where we need it most, especially the pharmaceutical industry.

I commend all of these essays to readers from across the spectrum of legal experience. They offer fresh perspectives on important issues.

FOREWORD (PUBLIC LAW)

Professor Paul Craig
Professor of English Law, St John's College, Oxford

Legal reasoning is foundational to the study of law at university. It is central to the legal education undertaken by law undergraduates during their three years or more at University. The skills in this respect are what you learn during your University years, and it is what we, the academic community, attempt to imbibe, through lectures, tutorials, seminars and our own legal writing. This understanding of legal reasoning is key, irrespective of the subject matter being studied. It is equally important in the context of contract and constitutional law, tort and administrative law, and land law and EU law, or indeed any other subject that is part of the law syllabus.

Skills in terms of legal reasoning play out in the determination and assessment of the positive law, the identification of what the law actually is on any particular topic. This may well require analysis of various primary sources, including statute, treaty provision, secondary legislation, codes and case law. The inter-relation between these discrete elements is itself an expertise honed over time. We can all look back with an admixture of amusement and humility at our first efforts to read a statute, or a complex judgment of the Supreme Court or Court of Appeal. It is a testimony to the learning capacity of university law students how these skills develop expeditiously, such that the sophistication that is brought to bear on University exams towards the end of the first year are unrecognizable as compared to those that existed but six months earlier. This learning curve continues to rise over time, the trajectory bearing testimony to the growing ease with the materials.

Legal reasoning is not, however, confined to discernment of the positive law. It is equally significant in the assessment of the normative dimension of legal study. Law is properly regarded as falling within the social sciences. There is a policy underlying every legal rule. There are values that underpin all legal norms. Understanding of the positive law is but the first step in evaluating the end product. The good lawyer should always be cognizant of this and bring a critical eye to the statute, case law or any other legal norm. The policy may be readily discernible, or it may, by way of contrast, be opaque. The values that the legal rule seeks to serve may be readily identifiable, and unitary, or they may be plural and contestable. This is but the start of the normative inquiry, not the end thereof. The good lawyer may need to test provisional conclusions against the purposes and values served by other legal rules in closely linked terrain, in order to assess the fit between the two.

The skills adumbrated above are fashioned over the entirety of a person's legal career. Hackneyed although it may sound, lawyers never stop learning. I vividly recall my initial years as an academic, when a senior colleague told me that I would, in time, develop a 'feel for my subject'. I had no idea what that meant at the time, nor when it would happen, or indeed what it would mean. It did happen after the effluxion of some considerable time, but it was worth waiting for.

The Oxford University Undergraduate Law Journal is important in fashioning skills in legal reasoning. It is valuable for those undergraduates minded to undertake more extensive analysis of a legal problem than is capable of being done within the confines of a normal essay assignment, or even a longer term-length paper. It requires them to tackle the plethora of issues adumbrated above, including articulation of the legal status quo, identification of the normative values that are enshrined therein and critical evaluation thereof. It is entirely proper that law

undergraduates undertake such studies. They thereby enhance their legal skills, and contribute to the development of legal scholarship on the relevant topic. I have been impressed with the quality of the contributions that I have read, which display good knowledge of the primary and secondary literature in the relevant area, combined with perceptive insight as to the way in which the law should be developed hereafter. I wish the editors and contributors, present and future, well in their endeavours in this regard.

INTRODUCTION TO THE EDITION

Kenneth Chong and Anna Yamaoka-Enkerlin
Editors-in-Chief

It is with great pleasure that we introduce the 8th edition of the Oxford University Undergraduate Law Journal.

The amount and variety of law that an undergraduate law degree demands be covered can bring with it the feeling that there is limited room to become a specialist – to immerse yourself in an area, to discover and pursue niches of interest, and to develop detailed expertise on difficult questions.

It is the OUULJ's mission to provide this opportunity. The authors featured in this edition are to be commended for seeking out such adventure.

Julia Breschtelsbauer critically analyses *Henderson v Dorset Healthcare University NHS Foundation Trust*, arguing that the Court of Appeal was mistaken not to apply the decision in *Patel v Mirza* across the whole of tort law. With the Supreme Court recently granting permission to appeal, this timely, cogently argued analysis should be of wide interest.

Antonia Kendrick examines the vexed question of equivalents following the recent Supreme Court decision in *Actavis v Elli Lilly* and its progeny. She tackles the wide-ranging implications of the decision - from its relevance to the anticipated Unified Patent Court to the decision's potential to disincentivise innovation, undermining a key aim of the patent system.

Zach Pullar evaluates, and rejects, the *status quo* that mentally ill defendants to a negligence action are held to the standard of the ordinarily careful man, arguing moreover that the case law he examines betrays the law's underlying commitment to fault as its 'organising idea.' He does not stop there but takes on the

challenge of proposing a solution, suggesting a clear path to legislative reform.

Mehleen Rahman applies a close reading of the Supreme Court's decision in *Prudential Assurance v Revenue & Customs Commissioners* to a systematic exploration of the latest in a long line of overpaid tax litigation cases. She ultimately submits that the Court effected a significant change in the law of unjust enrichment on the basis of reasons which, once scrutinised, are largely unsound.

Abe Chauhan's engaging piece explores the tension between the principle of universality upon which human rights are founded, and the reality whereby states seek to limit the universal protection of rights where this would be self-implicating. Critical of existing models of extraterritorial application of human rights treaties, the article comprehensively sets out the case for a causal model to act as a needed replacement.

Oliver Pateman, in another topical submission, considers the Mental Health Capacity (Amendment) Bill 2019, which is set to come into force in October 2020. Importantly, he argues that cause for concern persists, as in neglecting to consider the human rights implications the Bill may fail to stand up to scrutiny in the courts, or be insufficiently resourced to give meaningful human rights protection to cared-for persons.

Tatiana Podstolna assesses the changes brought to the interpretation of Art. 102 TFEU by the emergence of novel digital markets, bringing a rigorous approach to an examination of how the features of these markets have required the Commission and the CJEU to reassess its approach in order for the law to remain fit to achieve its own goals.

Congratulations to Antonia Kendrick, winner of the prize for Best Private Law Submission, and Tatiana Podstolna, winner of the prize for Best Public Law Submission.

Many thanks are owed. The OUULJ depends on an entire team of Associate Editors. They often take on extensive additional reading on top of regular coursework in order to bring a standard of excellence to their collaborations with the authors whose work is featured here. Special thanks to Senior Associate Editors Tim Koch and Oskar Sherry, who, as well as contributing their invaluable experience and expertise, assisted with selecting this year's editorial board. The work of Sponsorship Officer Isadora Janssen and Publicity Officer Kulsimran Sidhu has also been key to advancing the OUULJ's profile over the past year.

Finally, as an outgoing Editor your highest hope is that your successors will rise to the challenge of stewardship, advancing the OUULJ's mission while surpassing what legacy you manage to leave behind.

Niamh Kelly and Adrian Burbie have done just that. We could not have been more impressed by their energy, organisation, and dedication. They also innovated – opening up the Journal to submissions from undergraduates from all Russell Group universities, aligning the OUULJ's formatting with that of the Oxford Journal of Legal Studies, and securing a spot for the OUULJ on bookshelves in university and college libraries across Oxford. Finally, they deepened the OUULJ's connection with the Faculty of Law, whose continued support has been gratefully appreciated. The number, quality, and diversity of the subject matter of the articles featured in this edition are a testament to Niamh and Adrian's exceptional leadership.

Throughout our tenure we felt inspired and challenged by our peers. We hope that readers come away from this 8th Edition of the Oxford University Undergraduate Law Journal with the same feeling.

PRIZES

Best Private Law Submission to the Eighth Edition of the Oxford University Undergraduate Law Journal (2019):

Actavis v Eli Lilly: Patent Medicine, Risk of Side Effects
Antonia Kendrick
St Hugh's College, Oxford

The Editors are grateful to Professor Laura Hoyano (Wadham College, Oxford) for adjudicating upon the private law articles.

Best Public Law Submission to the Eighth Edition of the Oxford University Undergraduate Law Journal (2019):

Recognising Human Irrationality: Consumer Behaviour and Assessing Commercial Strategies in Digital Markets under Art. 102 TFEU
Tatiana Podstolna
Worcester College, Oxford

The Editors are grateful to Professor Paul Craig (St John's College, Oxford) for adjudicating upon the public law articles.

PRIVATE LAW ARTICLES

Henderson and Illegality in Tort: A Needless Retreat into Opacity

Julia Brechtelsbauer*

Abstract—This article discusses the decision of *Henderson v Dorset Healthcare University NHS Foundation Trust*. It argues that the Court of Appeal was mistaken not to apply the decision in *Patel v Mirza* across the whole of tort law. This article advances the view that *Patel* represented a step forward for illegality in tort law because it forced public policy decisions to be addressed more explicitly in judgments. The test formulated in *Patel* is also asserted to be clearer and transparent, leading to greater predictability. It is for these reasons that the decision in *Henderson* is regrettable.

1. Introduction

It is a great shame that the Court of Appeal in *Henderson v Dorset Healthcare University NHS Foundation Trust* [1] made an unnecessary

* Wadham College, Oxford. I am grateful to Sandy Steel and to the OUULJ editorial team for comments made on earlier drafts. All errors remain my own.

[1] *Henderson v Dorset Healthcare University NHS Foundation Trust* [2018] EWCA Civ 1841, [2018] 3 WLR 1651.

regression in choosing not to apply *Patel v Mirza* in a tort law context.[2] Refusing to address policy concerns of consistency explicitly will have consequences not only for illegality in tort law but potentially English law as a whole. Yet again, judges refuse to transparently address public policy concerns.

In *Henderson* the claimant argued that the decision of the unjust enrichment case *Patel v Mirza* should apply to tort law. The *Patel* judgment included a new test to be applied to cases involving illegality. The issue in *Henderson* was whether the new test in *Patel* had replaced the previous rules for illegality in tort found in *Gray* and *Clunis*. [3] The Court of Appeal held that *Patel* did not apply where there is previous authority. Thus, the claimant was held to be barred from compensation due to her illegality, using the test from *Gray*.

This article contends that *Patel* should be applied regardless of previous authority because the test formulated in *Patel* is preferable to previous formulations. The test proposed by Lord Toulson in *Patel* improves on the previous law, which was complex, uncertain and lacked transparency. The *Patel* test is more open and transparent, addressing the different underlying principles of illegality more explicitly, thus making it a predictable test.

The first part of this article will argue that *Patel* should be interpreted, as a matter of law, to apply across private law. The second part of this article will concern *Henderson* and the reasons

---

[2] *Patel v Mirza* [2016] UKSC 42, [2017] AC 467.

[3] *Gray v Thames Trains Ltd* [2009] UKHL 33, [2009] 1 AC 1339; *Clunis v Camden HA* [1998] QB 978 (CA).

given by the Court of Appeal for not applying *Patel*; this article will argue that the reasoning of the Court of Appeal demonstrates their misinterpretation of *Patel*. Key to this assertion is that in *Patel*, the Supreme Court used previous case law to influence the new test as proposed, meaning that the decision is rooted in the previous jurisprudence and is a development of it rather than a completely new invention. The third part of this article will demonstrate why *Patel* is preferred as an alternative test to the test employed in *Henderson*.

2. Patel v Mirza - Applicable across private law

It is better to start by determining what *Patel* decided, and then to evaluate the decision of *Henderson*, which necessarily concerned the decision of *Patel*. It is interesting here to analyse the academic predictions made prior to *Henderson* and evaluate whether these predictions are reflected in the decision.

A. What did Patel decide?

Mr Patel paid Mr Mirza a total of £620,000 following an agreement under which the defendant would use the money to bet on shares using inside information. Use of inside information for this purpose is a criminal offence under s. 52 of the Criminal Justice Act 1993. The agreement consequently could not be carried out since the inside information was not supplied and so the claimant sought repayment of the £620,000 through a claim for unjust enrichment.

The Supreme Court held that the claimant was entitled to restitution, despite the illegality involved. Lord Toulson, speaking for the majority of six (out of nine judges), rejected the rule-based approach, namely the reliance rule found in *Tinsley v Milligan*. [4] Lord Toulson said:

> '[O]ne cannot judge whether allowing a claim which is in some way tainted by illegality would be contrary to the public interest, because it would be harmful to the integrity of the legal system, without (a) considering the underlying purpose of the prohibition which has been transgressed, (b) considering conversely any other relevant public policies which may be rendered ineffective or less effective by denial of the claim, and (c) keeping in mind the possibility of overkill unless the law is applied with a due sense of proportionality. […] That trio of necessary considerations can be found in the case law'. [5]

The three rules found in Lord Toulson's *dicta*, labelled (a), (b) and (c), are to be referred to as the new *Patel* 'test' for illegality. The three considerations can be summarised as (a) purposive evaluation of prohibition, (b) possible effect on relevant public policies and (c) proportionality.

---

[4] *Tinsley v Milligan* [1994] 1 AC 340 (HL), rejected in *Patel* (n 2) at [110].

[5] *Patel* (n 2) [101].

B. How is this different to the law prior to Patel?

Comparing the *Patel* test with previous authority is inherently difficult because of the uncertainty and complexity which riddled this area of law. The law prior to *Patel* was confused by a variety of different doctrines. Although the law was confused, the concerns beneath the doctrines were not conflicting and this is what informed, and was crystallised, in the subsequent *Patel* test. Where *Patel* excels is that it is a test derived from a multitude of cases; it manifests the common threads from the cases in one test resulting in greater clarity in the law. [6]

In the decisions prior to *Patel*, two concerns arise as to the illegality defence: that the law remains consistent and is not self-defeating; and that the blameworthiness of the claimant is considered. Consistency is reiterated in a variety of forms but all demonstrate the desire of the courts to not undermine the criminal law through allowing private law recovery. Lord Hoffmann in *Gray* makes two arguments as to the justification behind the illegality defence, but this article argues they broadly demonstrate the same concern for consistency in the law. First, he says 'it is offensive to public notions of the fair distribution of resources that a claimant should be compensated (usually out of public funds) for the consequences of his own criminal conduct.'[7]

Second, he makes explicit reference to inconsistency: 'The inconsistency is between the criminal law, which authorizes the damage suffered by the plaintiff in the form of loss of liberty

[6] As argued by Andrew Burrows, 'Illegality after *Patel v Mirza*' (2017) 70 CLP 55; and Lord Toulson in *Patel* (n2) at [101]

[7] *Gray* (n 3) [51].

because of his own personal responsibility for the crimes he committed, and the claim that the civil law should require someone else to compensate him for that loss of liberty.' [8]

It may be argued that Lord Hoffmann's two justifications do in fact portray different concerns, but this article notes that upon closer inspection, they both relay the need or desire for consistency in the law. If one attempted to separate the two, one may argue the second point seems to be communicating a principle well known to the defence of illegality; namely, that criminals should not benefit from their own wrongdoing. But, in reality, the root of this principle is still *consistency*. How do we ensure that a criminal does not benefit from their own wrongdoing? By denying them the benefits of private law associated with the wrongdoing, that is to say, by promoting *consistency* in the law.

The first justification of fair distribution of public resource can similarly be reduced to a concern of consistency. Why are we concerned about fair distribution of resources? Should criminal and private law clash as a result of inconsistency, this would effectively result in a waste of public money. This demonstrates the practical necessity for consistency in the law, whereas the explicit reference to inconsistency demonstrates a more theoretical concern, namely, that it makes little sense to take liberty away for one reason and compensate for the loss of liberty for another.

This consistency concern is not just found in *Gray*. This policy motivation for the illegality defence is also found in

[8] ibid [37].

National Coal Board v England where the court examined the legislation to see whether its policy was affected by the illegality. [9] This is similarly advocating consistency in the law; it would be contrary to harmony in the law should the policy of the legislation be undermined through recovery in private law. There even seems to be a concern for the moral message which lack of the illegality defence could create: several cases (such as *Cross v Kirkby*) question whether allowing a claim would condone the claimant's illegal act. [10] This is rooted in a reluctance to contradict the 'oughtness' or the normative property of the criminal law.

Second, many of the cases, though contrasting in the means by which they achieve it, seek to address the blameworthiness of the claimant. Courts have taken notice of the seriousness of the claimant's illegality: for example, Lord Rodger proposed in *Gray v Thames Trains Ltd* that the decision may vary if the offence for which the claimant was convicted was trivial. [11] Similarly, Lord Sumption in *Les Laboratoires Servier v Apotex* comments that 'what constitutes turpitude for the purpose of the defence depends on the legal character of the acts relied on'. [12] Further, in *Clunis*, Lord Justice Beldam said that public policy only requires the court to deny assistance to the claimant if they were implicated in the illegality. [13] This particular concern has led to courts to balance interests. Lord Justice Ward said in *Hewison v Meridian Shipping*: 'the disproportion is between the claimant's

[9] [1954] AC 403 (HL).

[10] Cross v Kirkby [2000] EWCA Civ 426 (CA)

[11] *Gray* (n 3) [83].

[12] [2014] UKSC 55, [2015] AC 430 [28].

[13] [1998] QB 978 (CA) 987.

conduct and the seriousness of the loss he will incur if his claim is not allowed.' [14] The importance of the seriousness of the crime, implication in illegality and a need for proportionality all are veins of the same idea; namely, that the claimant should not be treated unfairly in comparison to their blameworthiness, hence leading to some cases adopting this proportionality approach.

The *Patel* test crystallises the substantive concerns of previous case law, but it does so by producing a different test. The *Patel* test broadly examines (a) the purpose behind the prohibition, (b) other public policies which may be affected and (c) proportionality. The *Patel* stages capture the two common threads of consistency and blameworthiness prevalent in the reasoning of the case law. In stages (a) and (b), the court is concerned about ensuring consistency in the law; requiring that the purpose is not undermined and allowing recovery that would not be contrary to other public policy. This captures the consistency concerns of Lord Hoffmann in *Gray*; purposive evaluation requires that the law does not, for example, take away liberty for one reason, and compensate it for another. One could also argue that stage (b) captures more precisely the practical motivation for consistency. Arguably, if an investigation costs a lot of money for public authorities, this may be a contrary to a general policy to have efficient public spending should recovery for an illegal act then be allowed. Proportionality in the third limb addresses the blameworthiness theme prevalent in the case law, regarding the *seriousness* of the offence, and how implicated the claimant was in the illegality.

[14] [2002] EWCA Civ 1821, [2003] ICR 766 [72].

This argument that case law was crystallised into the three limb *Patel* test is further exemplified by the Supreme Court's decision to overrule *Tinsley v Milligan.* [15] Lord Toulson suggests that the reliance principle of *Tinsley* 'could produce different results according to procedural technicality which had nothing to do with the underlying policies.' [16] Namely, it did not address the two threads of consistency and blameworthiness prevalent in the case law; most significantly the concern about the blameworthiness of the claimant. The reliance principle did not involve the weighing-up predominant in various cases, as previously argued. It therefore led to unfair and arbitrary decisions. Thus, through the *Patel* test, the majority of the Supreme Court were interested in developing and furthering both the common threads within the case law; proportionality and consideration of blameworthiness.

This view that previous law is subsumed into the *Patel* test is supported by *Patel* itself; Lord Toulson in his judgment refers to a variety of existing tort case law, for example *Gray*, but does not opine that they are wrong, and thus, not to be applied. Instead, they are discussed prior to his introduction to the test, demonstrating that this case law has informed the test produced. Lord Toulson does not *create* the test; it is observed through accumulated case law. Lord Toulson, the author of the *Patel* test, says himself, 'that trio of necessary considerations can be found in the case law.' [17] That is not to say that nothing changed through the production of the *Patel* test; the existing law was morphed into

[15] *Tinsley* (n 4).

[16] *Patel* (n 2) [87].

[17] *Patel* (n 2) [101].

another form. This form was a new clear and transparent test fusing the case law principles into one authority, instead of the blurred considerations before *Patel*. To assert that the *Patel* test is an invention is simply misguided; it is innovative unification of principles previously considered in the case law of illegality. It is a restructuring – but utilising the same resources as previously. This restructuring was the Supreme Court tailoring the law, to be greatly more clear, concise and more explicitly principled.

C. What predictions were made as to the application of Patel?

The applicability of the *Patel* decision was uncertain; the *ratio* of the decision involves unjust enrichment illegality, but Lord Toulson, leading the majority, speaks widely as if the *Patel* test were to apply to the whole of private law. This article argues, in agreement with Burrows, that 'it is clear […] that the majority saw itself as laying down a new approach to illegality across civil law. So, the new approach is applicable to the enforcement of a contract or a tort claim or a claim based on property law.' [18] Goudkamp argues that there are three 'clues' in the judgment of *Patel* that the majority intended the test to be applied across private law: [19]

> (i) The claim in *Patel* was unjust enrichment, but nothing was suggested by the judges to be particular about this context so far as the illegality doctrine is concerned;

[18] Andrew Burrows (n 6)

[19] James Goudkamp, 'Does *Patel v Mirza* apply in tort?' [2017] PILJ 1, 2.

(ii) The justices of the Supreme Court drew on case law from other parts of private law; and

(iii) *Tinsley* is rejected, a trusts case (applied in tort) although the facts of *Patel* were unjust enrichment.

(i) One may attempt to rebut this contention by demonstrating that, in this case, Lord Toulson observes the doctrinal specificity of unjust enrichment. [20] But it would be misguided to regard this as a strong criticism; Lord Toulson is specific when *applying* the test to the area of unjust enrichment. By saying that there are additional applicative considerations unique to unjust enrichment is not to render the test inapplicable more widely, for example, in tort. It suggests there are *additional* considerations when using the test for unjust enrichment. In fact, it would be wrong for the law not to mould to the niches of the type of claim to which it has been applied. The *Patel* test, in its (a), (b) and (c) format is clearly general and not exclusive in these elements. Only upon application, does the exclusivity and specificity of the law in hand come into play.

(ii) Judges often draw on other areas of the law to inspire their decision, but do not intend for their decision to impact those areas of the law. What is important, therefore, is *how* the case law is used. This article argues that the case law is used not only to inspire the *Patel* test, but to form it. The case law is not merely inspiration from which the test is drawn; it is its root – this is a more direct and active use of the case law. The *previous* case law is utilised to *propel* the *future* law. This article contends that the

[20] See eg. *Patel* (n 2) [44].

majority judgment makes use of the current mismatch of case law across private law, in an attempt to find one common theme within it. This view of the judgment is particularly supported by Lord Toulson's comment that:

> 'Looking behind the maxims, there are two broad discernible policy reasons for the common law doctrine of illegality as a defence to a civil claim. One is that a person should not be allowed to profit from his own wrongdoing. The other, linked, consideration is that the law should be coherent and not self-defeating condoning illegality by giving with the left hand what it takes with the right hand.' [21]

As noted, there are underlying concerns found in the case law which are then translated into the new *Patel* formulation. It would be nonsensical to criticise the general law applicable to all private law, analyse the cases, propose a new formulation, and then have the new formulation only apply to unjust enrichment. It appears in *Patel* that the majority are trying to find a solution within the whole of private law, and this should therefore be applied across private law. Throughout Lord Toulson's judgment he makes references to the criticism of the current state of the law of illegality in general, speaking of, for example, the four areas of critique by the Law Commission: 'complexity, uncertainty, arbitrariness and lack of transparency'. [22] This would be as a result of the case law prior to *Patel*.

---

[21] Ibid [99].

[22] *Patel* (n 2) [23]; Law Commission, *The Illegality Defence* (n 1) [3.50]–[3.60].

(iii) Goudkamp observes that, as there is no obvious reason why the policy-based test should be extended to the law of trusts and stop there, the consequent inference should be that *Patel* applies across private law, including to the law of torts. Further, the fact that *Tinsley* is expressly rejected, whilst no other cases are, is further evidence of the *Patel* test being rooted in this case law and thus not expressly overruling it or being incompatible with it. [23] This suggests, again, that *Patel* is a different *form of test* rather than one which completely *replaces* all of the prior case law.

In addition to Goudkamp's clues, this article's (and Burrows') contention about the wide applicability of the *Patel* test can be observed from the language and theme of the judgment. This is most poignant in the judgment of Lord Toulson. As previously highlighted, Lord Toulson speaks generally about the criticism of the current state of the illegality principle through reference, for example, to Law Commission reports. Further, Lord Toulson outlines how his three considerations are 'found in the case law'; [24] it is clear that the 'proportionate' limb of considerations is found from *ParkingEye Ltd v Somerfield Stores Ltd*, [25] a contract law case, whereas a multi-factor, flexible approach is observed by Lord Toulson to be in *Hounga v Allen* (a tort case). [26] Additionally, it is clear that Lord Toulson's (a), (b) and (c) are in no way specific to unjust enrichment, and constructed to be

[23] E.g., *Gray* (n 3) is not explicitly rejected in *Patel* (n 2)

[24] *Patel* (n 2) [101]

[25] *ParkingEye Ltd v Somerfield Stores Ltd* [2012] EWCA Civ 1338, [2013] QB 840.

[26] *Hounga v Allen* [2014] UKSC 47, [2014] 1 WLR 2889 [81].

compatible across private law. This combined with the context of the judgment following a call from Lord Neuberger for greater clarity in the whole area of unjust enrichment, leads to a fair assumption that the *Patel* judgment would extend across private law. [27]

The justification for the illegality defence given by Lord Toulson *informing* his consequent judgment is general to the whole of private law:

> 'Looking behind the maxims, there are two broad discernible policy reasons for the common law doctrine of illegality as a defence to a civil claim. One is that a person should not be allowed to profit from his own wrongdoing. The other, linked, consideration is that the law should be coherent and not self-defeating, condoning illegality by giving with the left hand what it takes with the right hand.' [28]

This wide justification is promoting harmony in the law, which should be encouraged. Harmony is preferable because it makes the law clearer, simplified and consequently more predictable and easier to follow. Thus, in promoting harmony, we help to uphold a fundamental rule of law principle. That is not to say it would be contrary to the rule of law should there be different rules for tort and unjust enrichment, for example, but that the rule of law would be upheld to a *greater* extent should there be congruence through private law in the illegality doctrine.

---

[27] *Jetivia SA v Bilta (UK) Ltd* [2015] UKSC 23, [2016] AC 1 [15].

[28] *Patel* (n 2) [99]

If we think harmony is beneficial, we may think it strange to limit this harmony to unjust enrichment and other areas of law, and to not extend it to tort law. Lord Justice McCombe highlights in *XX v Whittington Hospital Trust* that to argue that *Patel* is limited and thus not applicable to tort is contrary to the theme of the decision of *Patel*: 'the case itself emphasises the importance of cohesion and consistency in the law.' [29] It may be asserted that *Clunis* and *Gray* meet this consistency concern better. This may be true, but the rules in those cases do not adequately address the second thread of the case law: blameworthiness and balancing of interests to produce a proportionate response.

Goudkamp comes to a slightly different but not necessarily incompatible forecast:

> 'The policy-based test that was endorsed by *Patel* now applies throughout private law, but that test does not supersede (although it may perhaps operate as a cross-check on) more specific tests that have been developed in relation to particular types of claim. If this analysis is correct, cases concerned with the law of illegality decided in the law of torts remain relevant in the wake of *Patel*, including *Gray,* save for any cases that embrace the discredited reliance test.' [30]

What Goudkamp seems to suggest is that the courts are bound to apply *Patel* but that the court may be informed by previous judgments in application of the *Patel* considerations. This does not mean that *Patel* would not apply to the whole of

[29] *XX v Whittington Hospital NHS Trust* [2018] EWCA Civ 2832 [65]

[30] Goudkamp (n 17)

private law; it means that *in addition* to the *Patel* test the courts should be bound by previous decisions in that part of illegality, where there are more specific tests. This is not incompatible with the idea that the *Patel* test merely clarifies existing law – it just means past law is still relevant. So, for example, *Gray* would still be binding if the specific claim was familiar to that case. Goudkamp says that *Patel* does not supersede existing law; it is argued in agreement with him to the extent that *Patel* does not *overrule* previous law and thus does not supersede it in this way.

But this article would disagree with Goudkamp should he have meant that the *Patel* test does not *take the place of previous law*; rather it is argued that this was the very intention of the majority in the *Patel* case. This article wishes to emphasise that Goudkamp suggests that the *Patel* test could function as a 'cross-check' on the previous law. This seems to suggest that previous authority, such as *Gray* would be used in tandem with the *Patel* test. In itself, this observation appears to indicate that *Patel* offers something that the prior tests do not; this article argues that the *Patel* test offers clarity, structure and explicit principle which previous tests such as that used in *Gray* cannot provide.

3. Henderson v Dorset Healthcare - A misinterpretation of Patel

A. What did Henderson decide?

Following an alleged negligent release from mental health care, the claimant killed her mother during a serious psychotic episode as a result of her schizophrenia. The claimant was convicted of

manslaughter by reason of diminished responsibility. The question was therefore whether the claimant was barred from recovering her claim for compensation in negligence from the Health Board because of her consequent manslaughter conviction.

One of the arguments made by the claimant was that the precedent set in *Patel* applied to all cases of illegality. This was rejected by the Court of Appeal; it was held that *Patel* had not changed the precedent set in *Gray* and *Clunis*. Although mentioned previously, it is worth acknowledging the tests of those cases and how they were applied for comparison purposes. In *Gray*, a proximity test was used – it was held by the House of Lords that Gray was not entitled to compensation for loss of earnings as this resulted from his own criminal act, notwithstanding that the tortious act of the defendant had led the claimant to commit that offence. In *Clunis* it was held that considerations of public policy prevented the claimant from relying on his own criminal act unless it could be said that he did not know that the nature or quality of his act was wrong. The court followed these decisions in their judgment that the claimant was barred from recovery of damages due to the doctrine of illegality.

Looking back upon academic observations following *Patel,* Burrows' prediction was that *Patel* would apply to all private law, whilst Goudkamp added a gloss to this: previous cases with specific tests would remain binding should they be applicable. Since *Gray* was binding in *Henderson*, and there was no mention of *Patel* in the application of illegality in finding the outcome, the academic prediction of Burrows is contrasted. Goudkamp,

without his suggestion of *Patel* as a 'cross-check', seems to have been accurate in his prediction.

B. Were the judges right in Henderson not to apply Patel?

This article argues that the judges were mistaken not to apply *Patel*; this mistake is founded on the assumption implicated from the reasoning of the Court of Appeal that the *Patel* test is incompatible with the other case law such as *Gray*. This article has observed that the test is in fact based on previous case law. This view is echoed by Burrows, who argues that the approach by the High Court judge, Jay J, (which was upheld by the Court of Appeal) was wrong: 'What [the judge] should have done was to apply the 'trio of considerations' approach in *Patel v Mirza* in a way that, so far as possible, was consistent with the results in *Clunis* and *Gray*. That would have led him, by different reasoning, to the same result that he reached, namely that illegality was a defence on the facts of that case.' [31] This echoes the idea that it is not necessary to depart from *Gray* in order to apply *Patel*.

This article will now demonstrate how the misinterpretation of the function of the *Patel* test is evident from the reasoning of the Court of Appeal. The principal reasons given in the joint judgments of Sir Terence Etherton MR, Ryder and Macur LJJ as to why *Patel* did not apply can be condensed into two justifications.

Firstly, the judgment observes that there is absence of discussion of tort in the original *Patel* decision: 'It is impossible to

[31] Burrows (n 6) fn 28

discern in the majority judgments in *Patel* any suggestion that *Clunis* or *Gray* were wrongly decided or to discern that they cannot stand with the reasoning of *Patel'*. [32] This argument is that simply, tort is not mentioned in any detail. This observation ignores the general and obvious theme of the judgment. It is clear from the *Patel* judgment that the intention is widespread reform. This is particularly prevalent when Lord Toulson talks about the responsibility of the courts to reform when it is clear that the legislature will not, and that is the job of the courts to cure 'defects in the common law' which has occurred in other instances such as the infamous decision of *Jogee*. [33]

Further, this contention seems to be missing the point of why illegality needed a Supreme Court judgment in the first place; the issue, as observed also by the Law Commission, is not with the results of the decisions but the means by which they are found. This probably also explains why cases such as *Gray* were not directly overruled. The previous methods before *Patel* were muddied and confusing, whereas Lord Toulson's suggestion has been commended for its encouragement of transparent consideration of public policy. This is further developed by the fact that had *Patel* been applied, it is likely that the same outcome would have been met in *Henderson*, as noted by Burrows. [34]

The second justification for not applying *Patel* in *Henderson* is that *Clunis* and *Gray* are never said to be wrongly decided, and there is no implication that they cannot stand with

---

[32] *Henderson* (n 1) [88]-[89]

[33] *R. v Jogee (Ameen Hassan)* [2016] UKSC 8, [2017] AC 387 [114]

[34] Burrows (n 6), fn 28

the reasoning in the *Patel* case. [35] This reasoning demonstrates where the Court of Appeal may have been mistaken: there is no evidence in the judgment of *Patel* saying *Gray* and *Clunis* are wrong because the judgment was exactly doing the opposite of this. The judgment endorses previous principles underlying the case law, but distils the case law and these principles into a clearer and more concise test. That is not to say it is consistent with *all* previous case law; that would be to deny the obvious change which *Patel* makes completely. That the judgment endorses previous principles of cases also explains why there is no suggestion that the test is incompatible with these previous decisions. This is because this case law is what the *Patel* test is borne out of; the judges are finding the 'common denominators' of previous law in an attempt to find unifying principles. These unifying principles are then crystallised *into* the *Patel* test.

The contention that the judges were influenced by previous case law is also supported by decisions subsequently applying the *Patel* test. In *Tchenguiz & ors v Grant Thornton UK LLP & ors* Judge Knowles comments that since a previous case 'the law on illegality has been explained or developed'. [36] This certainly suggests that *Patel* is a development of previous law, rather than a completely new and foreign invention. Later judgments are also consistent with the idea that the main difference between the *Patel* test and previous authority is the form of the test; in *XX v*

[35] *Patel* (n 2) [89]

[36] *Tchenguiz & ors v Grant Thornton UK LLP & ors* [2016] EWHC 3727 (Comm) [39].

Whittington Hospital Trust Lord Justice McCombe talks of a 'new formulation of the law of illegality.' [37]

The arguments for restriction in *Henderson* are unconvincing and there are strong observations of wide-reaching reform intentions of the judges in *Patel*. This article therefore argues that the judges in *Henderson* were wrong not to apply *Patel* and misguided in their assumption that it was a black and white choice between previous law or the *Patel* test; they are in fact compatible. With application being made to appeal to the Supreme Court by the claimant in *Henderson*, it can only be hoped that the application will be granted and that the decision of the Court of Appeal will be overturned; it would not be a complete reversal of the decision because the same outcome would have been achieved, the means employed by the Court of Appeal, however, should be corrected.

The rest of this article will explore possible implications of the Court of Appeal choosing to apply *Gray* and *Clunis* in place of the *Patel* test. It will be argued that this is an undesirable step law, re-introducing a multitude of issues found with such tests: namely, lack of transparency and abundance of confusion. It is asserted *Patel* effectively resolves these issues. Thus, the Court of Appeal's decision is regrettable.

[37] *XX* (n 23) [57]

4. Henderson and its Implications: A Return to Confusion

A. Unnecessary Regression

The practical consequence of *Henderson* is that where there is pre-existing binding authority, which does not concern unjust enrichment, this will apply *instead of* the *Patel* test. This article argues that the decision of *Henderson* is a mistake for tort law because it means that the *Patel* test is not applied. The *Patel* test *should* be applied not only as a matter of interpretation from the judgment of *Patel* but because there are attractive reasons for preferring the *Patel* test over other existing case law such as *Gray*. If it is argued that the *Patel* test has its origins in case law such as *Gray*, one cannot argue that the principle behind *Patel* is better than that of *Gray* because they necessarily are the same. This aligns with previous arguments that the 'threads' of consistency and blameworthiness are found in previous case law, and reiterated in the *Patel* test. If the *substance* is not different, then it must be the *form* which makes the test better than the previous case law. This is because the form cures the defect otherwise produced by the case law: confusion and uncertainty.

This article has further observed that *Henderson* represents an *unnecessary* regression. There is simply no need for this confusion and uncertainty, which *Patel* had cured. This was argued through highlighting the two common threads of consistency and blameworthiness in previous case law and demonstrating how they were translated into the *Patel* test. *Patel* changes the form, not the substance of the law.

B. Defects in previous law now reintroduced as a result of Henderson

The issue with the law prior to *Patel* is best explained by the Law Commission:

> 'the conceptual basis on which the judges make their decisions is uncertain. Different judges have analysed the defence in different ways. A whole range of tests has been suggested as appropriate, in some cases to the exclusion of any other. As a result, it is difficult to predict an outcome or to explain the outcome in terms of the apparent rationales behind the illegality defence.' [38]

Notably, this was because of the variety of considerations at play, for example, seriousness of the offence, as opposed to the self-defeating danger for allowing the recovery for an illegal act. Consequently, the law is hard to predict. This is concerning since predictability of the law is an essential rule of law concept allowing true autonomy. This article contends this is because the common threads informing and influencing the *Patel* test are again masked behind these different headings in cases (such as *Gray* and *Clunis*), when, they were just derivatives or a species within the two main threads: consistency and blameworthiness.

C. How Patel cures these defects

The *Patel* test makes the law clear. The court is to apply the three strands, (a) (b) and (c) in reaching their outcome as to whether

[38] Law Commission, *The Illegality Defence* (n 1) 142.

illegality should deny the claimant of a claim. This in turn makes the law more certain because it is obvious what factors are relevant to the outcome. It is may seem paradoxical to argue that *Patel* encourages predictability, when it is a discretionary test, and discretion is usually criticised for being unpredictable. It is argued that this is mitigated by three considerations.

First, the decision in *Patel* does not give *no* guidance to judges, or potential claimants; it is not full discretion. Lord Toulson highlights what can be considered, as noted above.

Second, *Patel* only seems to be uncertain because it the first test to use such policy concerns. More solid rules will be distilled through development in the common law, on a case by case basis. Just like established relationships giving a duty of care, there may eventually be crimes that can never result in compensation in tort.

Third, the law would be more predictable as a result of *Patel* because judges can show more openly and honestly which policy motivations are influencing their decision, rather than hiding behind the rules of *Gray,* for example. If judges face the real concerns of illegality – namely consistency in the law and blameworthiness of the claimant – then the law is much more transparent and predictable.

Addressing these concerns is enabled by the *Patel* test. This third contention is not a novel idea and has been argued in other areas of tort and private law. For example, discretion around policy has been advocated by Stapleton in relation to pure

economic loss, [39] and Peden, in relation to implied terms in law in contract law. [40] This suggests that the *Henderson* decision may have been a digression for tort but also a restriction and pulling back on the law. The acceptance of the *Patel* test could symbolise a wider acknowledgment of the need for judges to address policy issues and discuss them more explicitly.

5. Conclusion

Burrows' prediction of *Patel*'s applicability was not a niche interpretation of the judgment. This article has argued that it is in fact the most natural and obvious reading from the judgment that the *Patel* test should apply across private law. This article has highlighted that the *Patel* test was not an entirely novel concept; built upon case law it is a declaration of clarity, of quite simply, what was already there. It is not incompatible with the tort decisions of *Gray* or *Clunis,* nor does it seek to overrule them. In fact, it endorses them and the principles behind them. This article has argued that the Court of Appeal made a mistake in assuming that the *Patel* test and previous authority were choices of black and white; they are, in reality, the same shade.

Further, since *Patel* solved issues highlighted with the previous law – particularly as to clarity, and predictability of the

[39] Jane Stapleton, 'Duty of Care and Economic Loss: A Wider Agenda' (1991) 107 LQR 249

[40] Elisabeth Peden, 'Policy Concerns Behind Implication of Terms in Law' (2001) 117 LQR 459

law – *Henderson* represents a digression of tort law. Additionally, it represents a step back for the law as a whole; policy considerations are to remain in the shadows. The *Patel* test would have enabled judgments to be more open thus unmasking the true policy reasoning behind many cases of illegality. *Henderson* represents a missed opportunity for tort law, and perhaps for other areas of English law as a consequence.

Actavis v Eli Lilly: Patent Medicine, Risk of Side Effects

Antonia Kendrick*

Abstract—This article discusses the recent Supreme Court decision *Actavis v Eli Lilly*, which reformulates the UK's approach to assessing patent claim infringement and introduces a doctrine of equivalents to the UK. Three arguments are made: (1) The decision brings UK law in convergence with that of other European jurisdictions, demonstrating a structurally flawed judicial attempt to prepare for the Unified Patent Court; (2) The UK's recognition of doctrine equivalents facilitates overly broad patentee protection, particularly if one considers numerical claims and prosecution history; and (3) Overbroad patentee protection disincentivises innovation, undermining a key aim of the patent system.

1. Introduction

The recent UK Supreme Court decision *Actavis UK Ltd v Eli Lilly and Co* [1] reformulates the UK courts' approach to assessing patent claim infringement and introduces a doctrine of equivalents to the

* St Hugh's College, University of Oxford. I would like to thank Luis Henrique Porangaba and the OUULJ editorial team for their insightful comments on earlier drafts. All errors are my own.

[1] *Actavis UK Ltd v Eli Lilly and Co* [2017] UKSC 48.

UK. In this article, *Actavis v Eli Lilly* will be examined from three perspectives: (1) European harmonisation, (2) breadth of patentee protection and (3) effect on innovation.

(1) The decision brings UK law on patent infringement in convergence with the law of other European jurisdictions. This is part of judicial efforts to align national approaches and thereby ease the transition to the Unified Patent Court ('UPC'). However, harmonisation via national courts is structurally flawed due to the lack of a common and supreme authority. Therefore, it lacks the legal certainty and low litigation costs that the UPC may facilitate.

(2) The UK's recognition of a doctrine of equivalents is problematic, since the doctrine facilitates overly broad patentee protection. *Actavis v Eli Lilly* may allow a patentee to draft a very narrow claim at the patenting stage and later claim a significantly larger monopoly at the infringement stage. The problems arising from overbroad patentee protection are most apparent when one considers *Actavis v Eli Lilly*'s effect on numerical claims and prosecution history.

(3) Overbroad patentee protection disincentivises innovation, thus undermining one of the key aims of the patent system. This possibility is exacerbated by the fact that under the current approach, even equivalents requiring an inventive step are capable of infringing patent claims.

2. A brief word on patents

Before engaging with the complexities of *Actavis v Eli Lilly*, a brief word on patents is useful. [2] A patent is a time-limited monopoly granted in return for a novel and inventive technical subject matter disclosed in a patent specification comprising claims and a description. [3] Patent disputes are usually resolved with reference to a person or team skilled in the art, i.e. a notional addressee with common general knowledge of the technical field in question. Harmonised prosecution pathways exist at national and European levels. An applicant can apply for a patent via her national office (for example the UK Intellectual Property Office) or through the European Patent Office ('EPO'), which offers a bundle of national patents covering those European states in which the applicant seeks recognition of her monopoly. The scope of a patent's monopoly is interpreted by the national courts via Article 69 of the European Patent Convention ('EPC').

---

[2] A detailed description of patents is beyond the scope of this essay. For an introduction to the subject and an overview of the key issues, see William Cornish's 2002 Clarendon Law Lectures, 'Intellectual Property: Omnipresent, Distracting, Irrelevant?', in particular Lecture 1: Inventing. The most up to date account of intellectual property law, including patent law, in a textbook is Lionel Bently, Brad Sherman, Dev Gangjee and Phillip Johnson, 'Intellectual Property Law' (OUP, 5th ed. 2018).

[3] For examples of EPO patent applications concerning different technical fields, see: <https://www.epo.org/applying/european/Guide-for-applicants/html/e/ga_aiii.html> accessed 23 November 2018.

3. An overview of Article 69 EPC and Actavis v Eli Lilly

Article 69 and its Protocol are key to understanding *Actavis v Eli Lilly*. The provision states that the extent of European patent protection is determined by the claims, which are interpreted using the description and drawings. [4] Article 1 of the Protocol on the Interpretation of Article 69 ('the Protocol') elaborates further that the extent of protection is not confined to the 'strict, literal meaning of the wording of the claims'. [5] Nonetheless, the claims are not to serve as a mere 'guideline' either. Instead, Article 69 'is to be interpreted as defining a position between these extremes which combines a fair protection for the patent proprietor with a reasonable degree of legal certainty for third parties.' [6] Article 69 therefore embodies a compromise between what have been termed the 'fence post' and 'sign post' approaches. [7] The 'fence post' approach marks the outer limits of the patentee's monopoly, while the 'sign-post' approach indicates, but does not determine, the monopoly's scope. [8]

More controversially, Article 2 of the Protocol states that 'due account shall be taken of any element which is equivalent to that specified in the claims' when assessing the patent's scope of protection. [9] This provision was adopted with the aim of

---

[4] Article 69 EPC.
[5] Article 1 (General Principles) of the Protocol on the Interpretation of Article 69 EPC.
[6] ibid.
[7] Hugh Laddie, 'Kirin-Amgen – The End of Equivalents in England?' (2009) 40 IIC 3 [13]. For a slightly different view, see Lord Hoffmann in *Kirin-Amgen Inc v Hoechst Marion Roussel Ltd* [2004] UKHL 46 [25].
[8] ibid.
[9] Article 2 (Equivalents) of the Protocol on the Interpretation of Article 69 EPC.

furthering a harmonised European approach to the treatment of equivalents. [10] The term equivalents refers to variants of protected inventions that still fall within the scope of the patent. The notion is illustrated well by the facts of *Catnic Components Ltd v Hill & Smith Components Ltd.* [11] In this case, a patent directed to a builders' lintel specifying that the plates should 'extend vertically' [12] was infringed by lintels with plates 6° and 8° off the vertical. [13] The infringer had drawn on the patentee's observation that the plates bear the greatest load when vertical. Therefore, the infringer produced lintels that were roughly vertical and could still bear heavy loads. In recognising that patents should be given a 'purposive construction', [14] *Catnic* demonstrates that patent claims, which are supposed to set out the patentee's monopoly, should not be construed too rigidly upon interpretation. Indeed, Lord Hoffmann, in *Kirin-Amgen Inc v Hoechst Marion Roussel Ltd,* compared the approach in *Catnic* with that used to interpret commercial contracts, frequently also referred to as purposive construction. [15]

Actavis v Eli Lilly centred on the chemical pemetrexed, an antifolate which inhibits cancerous tumour growth. However, if taken on its own, the chemical carries potentially lethal side effects. Lilly discovered that the side effects can be largely avoided if pemetrexed disodium is administered with vitamin B12. They subsequently applied for a patent claiming the manufacture of this new medicament. The specification contained both general statements about antifolates and references to pemetrexed disodium in particular. [16] During prosecution at the EPO, the

[10] Basic Proposal for Revision of the EPC 59.
[11] *Catnic Components Ltd v Hill & Smith Components Ltd* [1982] RPC 163.
[12] ibid 188 line 29.
[13] ibid 244 lines 26-29.
[14] ibid 243 lines 3-5.
[15] *Kirin-Amgen* (n 7) [34].
[16] *Actavis v Eli Lilly* (n 1) [17].

claims were narrowed to pemetrexed disodium. [17] Later, Actavis proposed generic products involving pemetrexed compounds combined with vitamin B12 for cancer treatment but using the variants pemetrexed dipotassium or pemetrexed ditromethamine (i.e. compounds not expressly referred to in Lilly's patent claims). Delivering the unanimous judgment, Lord Neuberger held that Actavis' proposed products would directly infringe Lilly's patent in the UK, France, Italy and Spain. [18] In doing so, Lord Neuberger set out what the Court of Appeal has recently termed a 'markedly different' [19] approach to infringement than that put forth by Lord Hoffmann in the earlier House of Lords decision *Kirin-Amgen*. [20] Lord Neuberger held in *Actavis v Eli Lilly* that infringement should now be approached via two questions: [21]

> 1. Does the variant infringe any of the claims as a matter of normal interpretation; and, if not
>
> 2. Does the variant nonetheless infringe because it varies from the invention in a way or ways which is or are immaterial?

Question 1 raises the issue of interpretation, whereas question 2, which would reach so-called equivalents, must be answered with reference to facts and expert evidence. Lord Neuberger argued that Lord Hoffmann had conflated these separate issues in *Kirin-Amgen*, treating them incorrectly as a single question of interpretation. [22] Furthermore, Lord Neuberger opined that the Protocol questions formulated by Hoffmann J (as he then was) in the 1990 judgment *Improver Corp v Remington*

---

[17] ibid [80].
[18] ibid [112].
[19] *Icescape Ltd v Ice-World International* [2018] EWCA Civ 2219 [59].
[20] *Kirin-Amgen* (n 7).
[21] *Actavis v Eli Lilly* (n 1) [54].
[22] ibid [55].

Consumer Products Ltd [23] are helpful in assessing question 2, but only if reformulated. Thus, the Protocol questions now read: [24]

> 1. Notwithstanding that it is not within the literal meaning of the relevant claim(s) of the patent, does the variant achieve substantially the same result in substantially the same way as the invention, i.e. the inventive concept revealed by the patent?
>
> 2. Would it be obvious to the person skilled in the art, reading the patent at the priority date, but knowing that the variant achieves substantially the same result as the invention, that it does so in substantially the same way as the invention?
>
> 3. Would such a reader of the patent have concluded that the patentee nonetheless intended that strict compliance with the literal meaning of the relevant claim(s) of the patent was an essential requirement of the invention?

There is an important shift between the original *Improver Corp* formulation of the Protocol questions [25] and the *Actavis v Eli Lilly* approach. In *Improver Corp* itself, the Protocol questions led Hoffmann J to conclude that a patent for an epilator which worked by trapping hairs in a rotating coiled helical spring was not infringed by an epilator that used a slotted rubber rod to achieve the same effect. [26] In light of the new approach, Lord Neuberger has suggested extra-judicially that the case may have been decided differently today. [27] This is because the slotted

---

[23] *Improver Corp v Remington Consumer Products Ltd* [1990] FSR 181.
[24] *Actavis v Eli Lilly* (n 1) [66].
[25] See *Improver Corp* (n 23) 7 for the original Protocol questions.
[26] ibid 12.
[27] See the UCL Faculty of Laws Institute of Brand and Innovation Law discussion of *Actavis v Eli Lilly* titled 'K=Na', available at

rubber rod epilator would now fall under Lord Neuberger's second question, namely whether the variant infringes because it varies from the invention (the rotating coiled helical spring epilator) in an immaterial way.

Actavis v Eli Lilly also establishes a 'sceptical, but not absolutist' [28] approach to correspondence between the patent applicant and the patent examiner during the examination process – in other words the prosecution history. In *Actavis v Eli Lilly* itself, the correspondence revealed that Lilly's claims were narrowed from antifolates to pemetrexed and finally to pemetrexed disodium, [29] but Lord Neuberger considered this irrelevant to his finding of direct infringement. [30] The prosecution history is, however, relevant in two scenarios: [31] first, if the point in question is truly unclear when confined to the specification and claims of a patent, and the file unambiguously resolves the point; second, if it would be contrary to the public interest to ignore the file's contents.

4. European harmonisation

A. Judicial alignment

Actavis v Eli Lilly signals the UK's broad alignment with a Continental European approach to Article 2 of the Protocol, as seen through a comparison with Germany, the Netherlands, Italy

<https://www.ucl.ac.uk/laws/news/2017/nov/equivalents-k-na-genie-out-bottle> accessed 16 November 2018.

[28] *Actavis v Eli Lilly* (n 1) [87].

[29] ibid [76]-[80].

[30] ibid [89].

[31] ibid [88].

[32] and Switzerland. Lord Neuberger explicitly discussed the first three jurisdictions in his judgment, and although Switzerland was not mentioned, it is relevant as an EPC state that has recently followed the UK's ruling in *Actavis v Eli Lilly*.

The German Federal Court of Justice (Bundesgerichtshof, or 'BGH') set out its approach to equivalents in *Schneidmesser I* [33], citing the UK decision *Improver Corp* as an example of 'harmonised law'. [34] The BGH held that an equivalent may fall within the scope of patent protection, setting out three questions that are similar but not identical [35] to Hoffmann J's Protocol questions. The *Schneidmesser I* questions are summarised lucidly by Meier-Beck (one of the BGH judges who decided *Schneidmesser I*) as follows: [36]

> 1. Does the modified embodiment solve the problem underlying the invention by means which have objectively the same effect? If the answer is no, the equivalent falls outside the patent's scope. If yes, the second question is:
>
> 2. Was the person skilled in the art enabled by his expertise on the priority date to find the modified means

[32] Lord Neuberger discusses these three jurisdictions as well as Spain and France in *Actavis v Eli Lilly* (n 1) at [44]-[52].
[33] *Schneidmesser I* BGH Case No. X ZR 168/00. Translated judgment available at [2003] ENPR 12.
[34] ibid [23]. This is an odd conclusion, since the product in question, an epilator, was held to infringe in Germany: see *Improver Corp* (n 23) 12-13. This divergence may be due to Germany recognising a doctrine of equivalents at the time, whereas the UK did not.
[35] Nicholas Pumfrey, Martin J. Adelman, Shamnad Basheer, Raj S. Dave and Peter Meier-Beck, 'The Doctrine of Equivalents in Various Patent Regimes – Does Anybody Have It Right?' (2009) 11 YaleLJ&Tech 261, 291 and 294.
[36] ibid 291-293.

> as having the same effect? Again, if the answer is no, there is no infringement. If the answer to both questions is yes, the third question is:
>
> 3. While answering question 2, are the considerations that the person skilled in the art applies drawn from the technical teaching of the patent claim (so that the person skilled in the art took the modified embodiment into account as being an equivalent solution)?

Actavis v Eli Lilly aligns the UK approach more closely with that of Germany, since both now recognise and use similar tools to assess equivalents. As Widera notes, [37] although the *Schneidmesser I* and Protocol questions are similar, one distinction remains: under Lord Neuberger's approach, equivalents requiring an inventive step are able to infringe, [38] whereas this is not the case in Germany. [39] It is argued below that this is a significant difference, and that the UK stance hinders innovation.

The current UK approach is also similar to that of the Netherlands, which applies the doctrine of equivalents if (i) the variant is foreseeable at priority date, (ii) the inventive concept covers the variant, (iii) the variant makes use of the inventive concept and (iv) reasonable legal certainty is not unduly compromised. [40] Factors (ii) and (iii) in particular align with Lord Neuberger's first reformulated Protocol question.

In Italy, a variant infringes if (i) it reproduces the inventive core of the patent and (ii) it is obvious, although (iii) the fact that the variant includes some non-obvious modifications or

---

[37] Phillip Widera, 'Has Pemetrexed revived the Doctrine of Equivalence?' (2018) 13 JIPLP 238, 245.
[38] *Actavis v Eli Lilly* (n 1) [64].
[39] ibid.
[40] ibid [51].

does not include all elements of the patent claim does not prevent infringement. [41] Factors (i) and (ii) again resemble the corresponding *Actavis v Eli Lilly* Protocol questions.

Lord Neuberger clearly attempted to align the UK approach with that of other European jurisdictions. He discussed the approach of other EPC states at length [42] and referred to Germany, the Netherlands and Italy in particular in connection with the reformulated Protocol questions. [43] Furthermore, the UK Supreme Court judgment purports to decide the issue of infringement not just as regards the UK patent, but also as regards the French, Italian and Spanish designations of the patent. [44] This has had a significant effect, as *Actavis v Eli Lilly* has indeed been followed by the Court of Milan [45] and the Swiss Federal Supreme Court, [46] despite the fact that Lord Neuberger did not comment on Switzerland.

The Court of Milan found infringement by clarifying the role of prosecution history, stating that it can only be 'completely secondary and ancillary, and may at most provide merely circumstantial evidence as to the patent holder's willingness to exclude or not certain solutions from the scope of protection.' [47] Thus, Lilly's amendments did not restrict the scope of protection

---

[41] ibid [48].
[42] ibid [44]-[52].
[43] ibid [62]-[64].
[44] ibid [92]-[102].
[45] *Eli Lilly & Company and Ors v Fresenius Kabi Oncology Plc and Ors* Case no. 45209/2017. Translated judgment available at <http://eplaw.org/wp-content/uploads/2018/11/IT-Court-of-Milan-Lilly-v.-Fresenius-pemetrexed-EN.pdf> accessed 15 January 2019.
[46] *Eli Lilly & Company v Actavis Switzerland AG* Case no. R11301CH00. Translated judgment available at <http://eplaw.org/wp-content/uploads/2017/11/CH-Judgment-Swiss-Federal-Supreme-Court-ENG.pdf> accessed 15 January 2019.
[47] *Eli Lilly v Fresenius Kabi* (n 45) 9.

so as to exclude equivalents of pemetrexed disodium, meaning that Actavis' proposed products fell within the patent. Similarly, the Swiss Federal Supreme Court's finding of infringement related closely to its conception of prosecution history, which the Court held is 'generally not determinative of the patent claims' [48] and therefore does not determine the scope of protection either.

This demonstrates judicial alignment across the EPC states in anticipation of the UPC. The UPC will constitute a centralised authority that has exclusive competence over European patents [49] and facilitates the granting of such patents with unitary effect [50] (as opposed to the bundle of national patents currently issued by the EPO). As Fisher notes, 'by ultimately tethering the interpretation of the grant to a single judicial body much of the twisting and turning, the conceptual gerrymandering and inconsistent treatment that are cited as reasons for the revision of Art.69 and the Protocol, may have been avoided.' [51] This is because once the UPC is in place, a European patent can be enforced in multiple states via a single infringement action before the UPC. It will therefore be less expensive for patentees to enforce their rights: instead of having to bring separate infringement actions before national courts, the patentee can rely on one judgment issued by the UPC. Furthermore, since all participating states will be bound by the judgment of the UPC, [52] national courts will be prevented from reaching different conclusions. The UPC will therefore both lower costs and protect

[48] *Eli Lilly v Actavis Switzerland AG* (n 46) 4.3.
[49] Articles 1, 32 and 34 UPCA.
[50] Article 1 UPCA.
[51] Matthew Fisher, 'New Protocol, same old story? Patent claim construction in 2007: looking back with a view to the future' [2008] IPQ 133, 7.
[52] Article 32 UPCA.

legal certainty. [53] There is currently no legislative provision which requires alignment, but the national courts are nonetheless choosing to follow each other's reasoning. Such judicial efforts to harmonise will ease the transition to a unified regime.

B. Challenges facing the UPC

However, the UPC currently faces two significant challenges. First, a constitutional complaint has been filed against German ratification of the UPC Agreement ('UPCA') on the grounds that it breaches a constitutional requirement that adoption of legislation amounting to a transfer of sovereign power to European institutions must be decided by a two-thirds majority in the *Bundestag* (federal parliament) and the *Bundesrat* (council representing the sixteen federal states). [54] Second, the UK's imminent withdrawal from the EU poses a challenge to its membership of the UPCA, which only extends to EU Member States [55] (although the UK ratified the agreement whilst still within the EU). [56] These are both serious obstacles, since the UPCA requires the participation of a minimum of thirteen Member States to come into force, including the three Member States with the highest number of European patents in effect in 'the year preceding the year' in which the UPCA was signed. [57]

---

[53] For discussion of these and other benefits, see <https://www.withersrogers.com/knowledge-bank/unitary-patent-package/ten-reasons-welcome-upc/> and <https://www.jakemp.com/en/knowledge-centre/upc-and-unitary-patent/unified-patent-court-the-benefits> accessed 9 February 2019.
[54] This requirement is derived from Articles 23(1) and 79(2) of the *Grundgesetz* (Basic Law for the Federal Republic of Germany).
[55] Article 2 UPCA.
[56] In April 2018, after the Brexit referendum but before formal withdrawal from the EU.
[57] Article 89 UPCA.

The Protocol to the UPCA clarifies that these three Member States are Germany, France and the United Kingdom. [58]

There are differing views on whether the UK could remain party to the agreement after withdrawing from the EU. Some legal practitioners argue that the UK's membership is essentially a political question dependent on the will of other Member States to amend the UPCA in the UK's favour. [59] Others suggest that there is a pragmatic interest in retaining UK involvement, since this could prevent strategic patenting effects between the UPC and the UK courts. [60] Whichever view one takes, Germany's non-participation is the more immediate challenge, as its refusal to ratify the UPCA could bring the project to a halt even if the UK is allowed to participate.

In light of this uncertainty, it could be argued that individual EPC Member States should continue taking all the steps they can to unify their approaches independently. However, such judicial harmonisation will remain a flawed effort. It cannot ensure full harmonisation, as noted by Lord Neuberger in *Schütz (UK) Ltd v Werit (UK) Ltd*: 'complete consistency of approach between different national courts of the EPC states is not a feasible or realistic possibility at the moment.' [61] There is no ultimate arbiter to ensure consistency, meaning that the benefits that a UPC would usher in (legal certainty and lower litigation costs) can only be partially realised through piecemeal harmonisation. The distinction between the UK's recognition of

---

[58] Article 3 of the Protocol to the UPCA on provisional application.
[59] See <https://www.lawyer-monthly.com/2018/06/uk-and-the-upc-what-will-happen-post-brexit/> accessed 17 November 2018.

[60] See <http://ipkitten.blogspot.fr/2016/07/guest-post-is-brexit-breaking-unitary.html> accessed 17 November 2018.
[61] *Schütz (UK) Ltd v Werit (UK) Ltd (Nos 1 to 3)* [2013] UKSC 16 [40].

inventive equivalents and Germany's rejection of them [62] comes to mind as an irreconcilable difference that is unlikely to ever be settled by national courts. The current mode of alignment is therefore necessarily limited.

5. Breadth of patentee protection

A second issue emerging from *Actavis v Eli Lilly* is breadth of patentee protection. Broad patent scope is best understood as protection outside of what Lord Neuberger terms the 'normal interpretation' of a patent. [63] The danger of this is indicated by Lord Neuberger's extra-judicial comments. He has argued that Lilly's patent would only have been limited to the claims if they had been phrased 'for pemetrexed disodium and nothing else.' [64] The onus is therefore on the patentee to explicitly limit her monopoly at the drafting stage. If she does not do so, there is always the possibility of claiming a wider monopoly at the infringement stage. This is problematic, as becomes clear when one considers *Actavis v Eli Lilly*'s effect on numerical claims and prosecution history in particular. In both cases, *Actavis v Eli Lilly* privileges patentees at the expense of legal certainty for third parties.

It is conceded that the finding of infringement in *Actavis v Eli Lilly* itself may be justified on the basis that the person skilled in the art would appreciate that Actavis' products would work in the same way as pemetrexed disodium when combined with vitamin B12. Furthermore, the person skilled in the art would appreciate that salt screening tests to determine whether a variant

[62] *Actavis v Eli Lilly* (n 1) [64].
[63] ibid [54].
[64] The UCL Faculty of Laws discussion (n 27).

of pemetrexed disodium would work were routine. [65] However, the UK's recognition of a doctrine of equivalents still allows patentees to draft a very narrow claim at the patenting stage and claim a significantly wider monopoly at the infringement stage, thus resulting in overbroad patentee protection. This facilitates the creation of monopolies that do not correspond to the scope of the claim.

A. Numerical claims

One situation in which patentees actively discount areas of patent protection is in making strict numerical claims and thereby excluding other numerical values. At a recent panel event, [66] Lord Neuberger conceded that he was 'attracted' by the point that where a patentee 'bookends' her claims with an upper and lower numerical limit, there is a clear indication that she does not want her monopoly to extend beyond the chosen values. Other speakers at the panel event agreed that, in general, there is not much room for a doctrine of equivalents in relation to strict numerical claims. Germany's BGH judge Meier-Beck approved [67] Jacob LJ's comment in *Rockwater Ltd v Technip France SA* that 'if the patentee has included what is obviously a deliberate limitation in his claims, it must have a meaning. One cannot disregard obviously intentional elements.' [68] Meanwhile, O'Malley, a judge on the US Court of Appeals for the Federal Circuit, noted that the patentee often has a very good reason for including numerical limitations in her application, namely to avoid the prior art. [69]

---

[65] *Actavis v Eli Lilly* (n 1) [69].
[66] The UCL Faculty of Laws discussion (n 27).
[67] ibid.
[68] *Rockwater Ltd v Technip France SA* [2004] RPC 46 [41].
[69] The UCL Faculty of Laws discussion (n 27).

The difficulty in reconciling strict numerical claims with the UK's doctrine of equivalents has come to the fore in *Regen Lab SA v Estar Medical Ltd*, [70] a recent Patents Court decision applying Lord Neuberger's new approach. The patent in question concerned a method of producing platelet rich plasma ('PRP'). Hacon J explained the technical background as follows: 'platelets are small unnucleated cells contained in blood plasma, responsible for blood clotting and tissue repair. PRP is plasma in which the platelet count is higher than would be found in the plasma of untreated blood. It is used by clinicians to promote the healing of wounds. There are also non-clinical diagnostic uses of PRP.' [71] At the priority date, different centrifugation methods were used to separate the platelet fraction to obtain the PRP. [72] Regen, the patentee, claimed a process for centrifugation using a thixotropic gel. Regen's method can be summarised in three steps: [73] (1) centrifuging whole blood from a patient in a separator tube containing a thixotropic gel and an anticoagulant; (2) removing half of the platelet poor plasma from the top of the tube; and (3) re-suspending the remaining plasma and buffy layer to produce PRP. Claim 1 of the patent mentioned two strict numerical limits: it specified that the separator tube be either a glass tube containing thixotropic gel and 0.10M sodium nitrate (an anticoagulant) or a polyethylene terephthalate separator tube containing a highly thixotropic gel formed by a polymer mixture and an anhydrous sodium citrate at 3.5mg/ml. [74]

---

[70] *Regen Lab SA v Estar Medical Limited, Estar Technologies Limited, Medira Limited, Lavender Medical Limited, Antoine Turzi* [2019] EWHC 63 (Pat).
[71] ibid [8].
[72] ibid [10].
[73] See <http://ipkitten.blogspot.com/2019/02/applying-actavis-questions-to-numerical.html> accessed 10 February 2019.
[74] *Regen Lab v Estar* (n 70) [25].

Although the patent was found invalid for lack of novelty and inventive step, [75] Hacon J went on to consider the issue of infringement. The alleged infringer, Estar, used an anticoagulant molarity of 0.136M, instead of the 0.10M specified in Regen's patent. [76] The relevant question for present purposes is whether Estar's chosen molarity nevertheless infringed Regen's patent. Applying Lord Neuberger's reformulated Protocol questions per *Actavis v Eli Lilly*, Hacon J found, with respect to question 1, that the inventive concept of claim 1 was 'the preparation of PRP for solely therapeutic use by employing a thixotropic gel wherein (a) there is only one centrifugation and (b) after centrifugation about half the supernatant is removed and the platelets are then re-suspended in the enriched plasma.' [77] Moving on to question 2, Hacon J concluded that it would be obvious to the skilled person that Estar's variant achieved substantially the same result as Regen's invention in substantially the same way. [78] However, most of the discussion centred on question 3. Hacon J held that the Regen would only intend that strict compliance with the literal meaning of claim 1 would be essential to the inventive concept if 'there had been a sufficiently clear indication to the skilled person that strict compliance with the figure of 0.10M was intended.' [79] The Patents Court therefore found that if Regen's patent had been valid, it would have been infringed.

Regen Lab v Estar indicates that following *Actavis v Eli Lilly*, strict numerical limitations will not necessarily demarcate a patent. This conclusion is problematic, since it allows a patentee to rely on numerical limits at the drafting stage as evidence of novelty, only to later claim infringement by a variant that falls

[75] ibid [197].
[76] ibid [212].
[77] ibid [235].
[78] ibid [242].
[79] ibid [252].

outside the patent's supposed limitation. The patentee is thereby able to have her cake and eat it: she takes the benefit of numerical claims (assist in establishing novelty) without conceding the burden (limited scope of monopoly). Meanwhile, the third party is left in an uncertain situation: what will amount to a 'sufficiently clear indication' [80] that the patentee intended strict compliance with the numerical values included in the patent? Litigation is required to establish its meaning, and inspiration from other jurisdictions is lacking, since remarks from senior judicial figures suggest that the UK approach is unorthodox. [81] Additionally, the UK's ruling on infringement in *Regen Lab v Estar* conflicts with the German ruling on the same dispute. In the German equivalent of the case the Regional Düsseldorf District Court dismissed Regen's claim of infringement. [82] These conflicting outcomes reinforce the point that judicial harmonisation has inevitable limits unless subject to a common and supreme authority.

Unfortunately, *Actavis v Eli Lilly* does not seem to allow for a more balanced approach to numerical limitations: Lord Neuberger has conceded extra-judicially that his personal view of numerical claims may be inconsistent with the judgment. [83] Given the tension between Lord Neuberger's extra judicial opinion and his ruling in *Actavis v Eli Lilly*, it seems that he overlooked the impact his judgment would have on numerical claims in patent law. This oversight is leading English law in a worrying direction that is based wholly on patentee-sided considerations.

---

[80] ibid.
[81] The UCL Faculty of Laws discussion (n 27).
[82] <https://www.prnewswire.com/il/news-releases/uk-high-court-revokes-regen-lab-prp-patent-for-lack-of-novelty-and-inventive-step-830262338.html> accessed 10 February 2019.
[83] The UCL Faculty of Laws discussion (n 27).

B. The role of prosecution history

The risk of overly broad patent monopolies is increased by *Actavis v Eli Lilly*'s stance on prosecution history. Clearly wary of an American-style prosecution history estoppel, which is of general applicability, subject to the limitations set out in *Warner Jenkinson Co v Hilton Davis Chemical Co* [84] and *Festo Corp v Shoketsu Kinzoku Kogyo Kabushiki Co*, [85] the UK Supreme Court only allows a prosecution history estoppel in two exceptional circumstances. This reinforces the suggestion outlined above that a patentee may accept a narrow monopoly during prosecution only to argue for a broader patent during infringement. Thus, all patentees who do not fall within one of Lord Neuberger's exceptions will be able to reach back and claim something during infringement that they have expressly denounced at the drafting stage. Again, such patentees get the best of both worlds: accepting a narrow monopoly during prosecution makes it more likely that a patent will be granted, but the patentee can then claim a much wider right during infringement.

If alleged infringers wish to rely on one of Lord Neuberger's recognised exceptions, they may face an uphill battle in attempting to do so. The first exception in particular may require much litigation to flesh out, despite Lord Neuberger considering this category 'self-explanatory'. [86] When exactly is a point 'truly unclear' on examination of the specification and claims, and when does the file 'unambiguously resolve' the issue? Since no examples are given in *Actavis v Eli Lilly*, the solutions to these questions are less obvious than Lord Neuberger suggests.

---

[84] *Warner Jenkinson Co v Hilton Davis Chemical Co* 520 US 17 (1997).
[85] *Festo Corp v Shoketsu Kinzoku Kogyo Kabushiki Co* 535 US 722 (2002).
[86] *Actavis v Eli Lilly* (n 1) [88].

6. Effect on innovation

Overly broad patentee protection disincentivises innovation. This is a serious drawback given that patents are very often justified on the grounds that they promote innovation. [87] Merges and Nelson discuss the effects of overly broad protection, noting that 'when a broad patent is granted or expanded via the doctrine of equivalents, its scope diminishes incentives for others to stay in the invention game, compared … with a patent whose claims are trimmed more closely to the inventor's actual results.' [88] Extending patent protection further makes it difficult for competitors to predict how the claims will be construed, resulting in legal uncertainty and thereby discouraging inventions which build on the teachings of a pre-existing patent. This is an especially relevant concern for the pharmaceutical sector. Since it costs an estimated £1.15 billion and twelve years to research and develop one new medicine, [89] many innovations in this area depend on pre-existing drugs, i.e. incremental innovation. Such innovation is valuable [90] for four reasons. First, response to a particular drug can vary significantly across patients, so having many therapeutic alternatives allows physicians to decide which

[87] For a more detailed discussion, see Dominique Guellec and Bruno van Pottelsberghe de la Potterie, 'The Economics of the European Patent System: IP Policy for Innovation and Competition' (OUP, 1st ed. 2007).

[88] Robert Merges and Richard Nelson, 'On the Complex Economics of Patent Scope' (1990) 90 Colum. L. Rev. 839, 916.

[89] See the Association of the British Pharmaceutical Industry's 2014 publication 'Delivering Value to the UK' at page 30, available at <https://www.abpi.org.uk/media/1673/delivering_values_dec2014.pdf> accessed 15 January 2019.

[90] Steven Globerman and Kristina M. Lybecker, 'The Benefits of Incremental Innovation – Focus on the Pharmaceutical Industry' (2014) 25.

treatment will best suit an individual patient. [91] Additionally, the availability of different therapeutic options means that patients are more likely to comply with their treatment regimen. [92] Third, an entirely new drug rarely constitutes the best version of that drug: [93] as Lybecker points out, 'subsequent development allows for improvements in a therapeutic class that would otherwise never occur.' [94] Finally, incremental innovation encourages competition, resulting in favourable prices for consumers. Unfortunately, the likelihood of incremental pharmaceutical innovations infringing patents may have increased following *Actavis v Eli Lilly*. Researchers must now decide whether they are willing to run the risk of infringement in developing treatments that build on existing drugs. In a worst-case scenario, *Actavis v Eli Lilly* may deter incremental innovation.

This possibility is exacerbated by the fact that under Lord Neuberger's approach, even equivalents requiring an inventive step may infringe. His position diverges from that in German law, which reaches the opposite conclusion. Lord Neuberger argued that adopting the German position was unnecessary since the existence of an inventive step may entitle the infringer to a new patent. [95] This raises the immediate question of which perspective a court would use to address the issue of inventive equivalents. One assumes that the judge would adopt the perspective of a person skilled in the art, but such a hypothetical addressee is famously unimaginative (i.e. lacks the capacity for creativity), as noted by Lord Reid in *Technograph Printed Circuits Ltd v Mills & Rockley (Electronics) Ltd* [96] and reaffirmed by Jacob LJ in *Rockwater*

[91] ibid 26.
[92] ibid 27.
[93] ibid 28.
[94] ibid.
[95] *Actavis v Eli Lilly* (n 1) [64].
[96] *Technograph Printed Circuits Ltd v Mills & Rockley (Electronics) Ltd* [1972] RPC 346, 355.

v Technip. [97] The person skilled in the art would thus not consider an inventive equivalent to fall within a patentee's monopoly. It can therefore be doubted whether 'inventive equivalents' should even be categorised as 'equivalents' – the fact that they are inventive suggests a distinctive rather than equivalent feature. Given this conceptual confusion, it may be difficult to even recognise an inventive equivalent in practice. Lord Neuberger's recognition of inventive equivalents undermines his claim that *Actavis v Eli Lilly* limits the UK's doctrine of equivalents to 'immaterial variations'. [98] The German position is to be preferred: it sets a clear limit to the doctrine of equivalents, meaning that competitors can more easily assess whether their products infringe.

A second objection to the UK approach is Lord Neuberger's view that an infringer would be able to obtain a new patent. The infringement trial necessarily takes place some time *after* the filing date or validly claimed priority date but the assessment as to the inventive step of the infringed patent is made by reference to the state of affairs *at* the priority date. This means that even if the infringer might have been entitled to a new patent at the priority date, between the priority date and the infringement trial other prior art may have become available which would render the variant non-inventive. The grant of a new patent might at this stage be unfeasible. Lord Neuberger's stance on inventive equivalents introduces a confusing element into the doctrine that may hinder innovation by weakening the position of competitors.

The arguments concerning breadth of patentee protection and effect on innovation are particularly pertinent to cancer research, which depends on the wide-ranging expertise of clinicians, molecular biologists, computational biologists,

[97] *Rockwater v Technip* (n 68) [7].
[98] *Actavis v Eli Lilly* (n 1) [34].

statisticians, nanotechnology experts and chemical engineers. [99] This means that there are many potential infringers who may be hesitant of investing in projects that could, following *Actavis v Eli Lilly*, face legal challenges. The decision may therefore deter further discoveries in this area, which is unfortunate given the inherent social aspect of intellectual property relating to cancer research.

7. Conclusion

Actavis v Eli Lilly is an example of judicial attempts across EPC Member States to align their approaches in preparation for the UPC. However, the doctrine of equivalents allows for overbroad patentee protection. The UK's conception of the doctrine is especially problematic in this respect, as becomes clear when one examines *Actavis v Eli Lilly*'s effect on numerical claims and prosecution history. Furthermore, overbroad patentee protection undermines innovation, which may have very detrimental consequences, particularly in the pharmaceutical sector – at worst, it may hinder incremental innovation.

[99] Neil Savage, 'Collaboration is the key to cancer research' 10 April 2018 Nature (International Journal of Science).

Of Unsound Mind: The Eternal Triangle of Negligence, Fault, and Mental Impairment

Zach Pullar*

Abstract—This article evaluates, and rejects, the *status quo* that mentally ill defendants to a negligence action are held to the standard of the ordinarily careful man. It critically considers the tenability of the Court of Appeal's 'elimination of responsibility' reasoning in its *Dunnage v Randall* decision in light of its previous *Mansfield v Weetabix* judgment and the relaxed standard applicable to relevantly analogous child defendants. It is argued these examples betray the law's underlying commitment to fault as its 'organising idea', despite contrary assessments. The article proposes a legislative solution by which judges may exclude certain mentally ill defendants from negligence liability.

1. Introduction

The standard of care in negligence is the benchmark against which the propriety of the conduct of a defendant who owes a duty of

---

* University of Exeter. I am grateful to the OUULJ editorial team for comments on earlier drafts. All errors remain my own.

care to another is measured. Where this benchmark is set is a question of law, which this article addresses. The objective approach to this standard is of such long standing [1] that any modern case whose outcome apparently turns on the rule's applicability could be, in the realm of fiction, susceptible to a 'premature adjudication' in its favour. [2] However, the critically diverse reaction to the Court of Appeal case of *Dunnage v Randall* [3] may prove an important testament to the contrary, betraying an antipathy for the status quo. [4]

This article seeks to demonstrate a fundamental inconsistency between a number of critical judgments on incapacity and responsibility for the purposes of negligence law, and a *de facto* dichotomy between physical and mental impairment despite their formal identity *de jure*. [5] This article proffers a legislative solution: a decisive statement by Parliament affording the mentally impaired the legal protection their situation

[1] *Blyth v Birmingham City Waterworks* (1856) ER 1047, 1049 (Alderson B).

[2] See, e.g., John Mortimer, 'Rumpole and the Old, Old Story' in *The Second Rumpole Omnibus* (Penguin 1988) 499.

[3] *Dunnage v Randall* [2015] EWCA Civ 673, [2016] QB 639 (CA).

[4] E.g. John Spencer, 'Dunnage v Randall: Liability - Personal Injury - Negligence' (2015) 4 JPI Law 200; J Goudkamp and M Ihuoma, 'A Tour of the Tort of Negligence' (2016) 32 PN 137; Maria Orchard, 'Liability in Negligence of the Mentally Ill: A Comment on *Dunnage v Randall*' (2016) 45(4) Common Law World Review 366; John Fanning, 'Mental Capacity as a Concept in Negligence: Against an Insanity Defence' (2017) 24(5) Psychiatry, Psychology and Law 694.

[5] *Dunnage* (n 3) [104], [127], [159].

objectively demands, a clarion call which other common law jurisdictions have previously answered. [6]

2. The current law

It is trite law that 'the standard of care in the law of negligence is the standard of an ordinarily careful man'. [7] However, whilst objective, this standard is not a unitary benchmark; the particular circumstances in which putatively tortious activity occurs are selectively considered. [8] Where the defendant's activity imports some specialism or professional skill, 'the test is the standard of the ordinary skilled man exercising and professing to have that special skill'. [9] Similarly to the professional's reflexive standard—and, therefore, not uniquely [10] —children are held to a standard that is age-sensitive: [11] the variable standard thus does not *merely* shift upwards. [12] However, the law decisively refuses to countenance the defendant's *inexperience* when measuring his

[6] At common law: see below n 113 and n 114.

[7] *Bolton v Stone* [1951] AC 850 (HL) 863.

[8] Goudkamp and Ihuoma (n 4) 140-41. See, e.g., *Bolam v Friern Hospital Management Committee* [1957] WLR 582 (QB) 586.

[9] *Bolam* (ibid).

[10] Goudkamp and Ihuoma (n 4) 140-41.

[11] *Mullin v Richards* [1998] WLR 1304 (CA); *Orchard v Lee* [2009] EWCA Civ 295. The significance of this in particular is developed below.

[12] As is at least implicit perhaps in *Bolam*, which, in discussing 'what in law we mean by "negligence"', seems to contrapose the 'ordinary case' only against the case involving 'some special skill or competence': see *Bolam* (n 8).

conduct; the appropriate standard of skill is determined by that reasonably to be expected of an ordinary person undertaking the relevant non-professional [13] or professional [14] activity. Whilst the standard varies in these instances to reflect the defendant's circumstances, it remains, it is said, strictly independent of the ordinary person's 'idiosyncrasies'. [15] The relaxation of the standard for children, the court reasons, is merely commensurate with the reduced 'degree of sense and circumspection' [16] which is an objective 'characteristic' of a child; [17] likewise defendants must meet objective professional demands.

However, the unassailability of the objective standard—that the law governing children and professionals might already indicate is more apparent than real—is further undermined by the position as regards physical/mental impairment, and any resultant confusion is only exacerbated by the internal inconsistency of this area, developed below. [18] Thus, whether the impairment merits a relaxed standard pivots on whether the defendant was (reasonably) *aware* of that impairment. Only where he is not aware of it will a relaxed standard apply. [19] Suppression of the standard in consequence of some peculiar impairment betrays a basis that

---

[13] *Nettleship v Weston* [1971] QB 691 (QB).

[14] *Wilsher v Essex AHA* [1988] AC 1074 (HL).

[15] *Glasgow Corp v Muir* [1943] AC 448 (HL) 457.

[16] *McHale v Watson* (1966) 115 CLR 199 [9].

[17] ibid [6].

[18] John Fleming, *An Introduction to the Law of Torts* (2nd edn, Clarendon Law Series 1985) 27.

[19] *Mansfield v Weetabix* [1998] 1 WLR 1263 (CA); Cf *Roberts v Ramsbottom* [1980] 1 WLR 823 (QB). This is pithily known as the 'doctrine of prior fault'.

is 'idiosyncratic' to the possessor; otherwise the latter would simply be, for example, an 'ordinary' lorry driver, [20] against whose 'ordinary' standard of conduct the impaired person would be found negligent. [21] Crucially, relaxation is restricted to where the impairment 'entirely eliminates responsibility', [22] as in the event of severe hypoglycaemic attack where the defendant's actions are 'involuntary'. [23] There will be no suppression, it follows, where the defendant suffers schizophrenic attacks, since his mind 'still directs his actions'. [24] The law accordingly rejects this idiosyncrasy: it considers the schizophrenic's actions to be voluntary, and the schizophrenic defendant a legally responsible agent. [25]

3. Case for reform

The 'elimination of responsibility' reasoning of the Court of Appeal in *Dunnage v Randall* cannot support a finding of liability in that case consistent with its binding predecessor *Mansfield v*

[20] As in *Mansfield*.

[21] *Nettleship* (n 13).

[22] *Dunnage* (n 3) [114].

[23] *Mansfield* (n 19); ibid [48].

[24] *Dunnage* (n 3) [145].

[25] Similarly, schizophrenia does not *per se* prevent a patient from being legally capacitous to make decisions about his treatment: see, e.g., Clare Dyer, 'Woman with Schizophrenia is Judged to Have Mental Capacity to Decide on Amputation' (2014) 348 BMJ <https://www.bmj.com/content/348/bmj.g1737> accessed 7 April 2019.

Weetabix, [26] nor with the case law, canvassed above, more generally. In *Dunnage*, a Mr Randall, the claimant's schizophrenic uncle, doused himself in petrol during an altercation with the claimant in the latter's home. He set himself alight in protest and resultantly perished. The claimant, in an unsuccessful preventive attempt, was likewise 'engulfed in flames' and sustained serious burns. [27] Of significance is that the case was brought in the context of a household insurance scheme from which the claimant stood to benefit if he could establish that Randall's actions caused him '*accidental* bodily injury'. [28] Accepting that Randall owed a duty of care to the claimant, the issue for the court was thus whether his actions were negligent; or, in other words, whether they fell below the requisite standard of care. The predecessor appeal case, *Mansfield*, disputed Weetabix's vicarious liability for the damage to a shop, owned by the Mansfields, caused by the company's lorry driver, a Mr Tarleton. The driver, who (unbeknownst to him) suffered from insulinoma—resulting in excessive insulin production—crashed into the shop during the onset of a hypoglycaemic attack. [29]

According to the court in *Dunnage*, since the mind of a conscious, schizophrenic defendant yet 'directs his actions', [30] those actions are 'voluntary' and the defendant accordingly susceptible to liability in negligence. [31] However, this article will seek to demonstrate that whether the defendant's volition was

---

[26] Goudkamp and Ihuoma (n 4) 138-40.

[27] *Dunnage* (n 3) [3].

[28] ibid [4] (emphasis added).

[29] *Mansfield* (n 19) 1265.

[30] *Dunnage* (n 3) [145].

[31] ibid [131].

wholly or 'grossly' [32] impaired, the significance of the impairment for his legal responsibility *should*, in the appropriate case, [33] be no different. The apparent justification for the 'involuntariness exception' to liability is that fault is 'entirely eliminated' if a defendant 'did nothing himself to cause the injury' [34] because of physical/mental incapacity. Importantly, the success of the exception in *Mansfield* followed from the court's recognition that the law cannot hold a defendant who is *incapacitated* from acting to the same standard as a capacitous person. To rule otherwise would be tantamount to a finding of strict liability, since a capacitous defendant reasonably would not have crashed, and 'that is not the law' [35]—confirmed, on the face of it, in *Dunnage*. [36]

By this same inductive process, the court in *Dunnage* should have considered that, given his *non*-rationality—'[w]hat he *could not do was to decide* to take any alternative action that would obviate the risk' created by his actions [37]—the defendant was *incapable* of rational control, and so, like his *Mansfield* counterpart, of conformity with reasonable standards of conduct otherwise attainable by the ordinary man. On the *Dunnage* court's own admission, the named reason for that incapability is not material; the *effect* of it, however, is so. [38] Indeed, whilst *Dunnage* has revitalised this debate, the view that incapacity (associated with

---

[32] ibid [26]. That is, owing to mental illness.

[33] See below, 'Proposal for Reform'.

[34] *Dunnage* (n 3) [115], [132]. Much, therefore, rests on the appropriate construction of the reflexive pronoun.

[35] *Mansfield* (n 19) 1268. Given exceptions are, arguably, the doctrines of vicarious liability and non-delegable duties of care.

[36] *Dunnage* (n 3) [129].

[37] ibid [143] (Arden LJ) (emphasis added).

[38] *Dunnage* (n 3) [104] (Rafferty LJ).

mental illness) should be capable of excusing liability is not novel. [39] A 16th century articulation of first principles by Montaigne remains particularly compelling: 'We cannot be held responsible beyond our strengths and means, since the resulting events are quite outside our control, and, in fact, we have power over nothing except our will; which is the basis upon which all rules concerning man's duty must of necessity be founded.' [40]

Contrary to Vos LJ's view, the situation of the *Dunnage* defendant is incomparable to the qualitatively different 'matter of regret that even the most intelligent in our society sometimes do act irrationally', in which case the relevant benchmark is correctly unvaried. [41] Such defendants *would* 'act unreasonably because such persons *do not lack* rational control': [42] it is the incapacity for control intrinsic to defendants with certain mental illnesses that drives the requirement for legal protection. The *Dunnage* defendant, conversely, was 'not at fault' in relevantly the same sense as his *Mansfield* counterpart; it cannot accordingly be said that the two decisions are 'far removed' [43] from each other, since both decisions are concerned with the legal relevance of an incapacitating 'idiosyncrasy'. [44] Once this is accepted, the remaining material distinction between these two cases would appear to consist in the division that all three of the judges in

---

[39] See, e.g., James Barr Ames, 'Law and Morals' (1908) 22 Harv L Rev 97, 99-100; WGH Cook, 'Mental Deficiency in Relation to Tort', (1921) 21 Columb L Rev 333, 344; Francis Bohlen, 'Liability in Tort of Infants and Insane Persons' (1924) 23 Mich L Rev 9, 28, 31-32.

[40] JM Cohen, *Montaigne: Essays* (5th ed, Penguin 1967) 25.

[41] *Dunnage* (n 3) [135].

[42] Goudkamp and Ihuoma (n 4) 144 (emphasis added).

[43] *Dunnage* (n 3) [147].

[44] Goudkamp and Ihuoma (n 4) 139-40.

Dunnage purported to reject as a matter of law: [45] that Mr Randall's impairment was mentally incapacitating, whereas Mr Tarleton's—the *Mansfield* lorry driver's—was physically so.

Fleming is therefore correct to underscore that the opposite outcomes of these cases—one in respect of the consequences of unexpected seizure; the other, of those flowing from schizophrenic attack—are 'difficult to reconcile'. [46] This article accordingly rejects the 'binary capacity/incapacity interpretation' adopted in *Dunnage*, in which any defensible distinction between the two cases under discussion is located. [47] The *Dunnage* defendant's conduct may have been technically 'voluntary' insofar as his mind 'directed the attack'; [48] and, admittedly, *criminal* liability would be obviated only where the defendant can 'show that he was in a state of automatism [importing unconsciousness requisite for criminal liability]. [But] *in civil cases that is not the test*.' [49] *Mansfield* suggests that test—when reduced to a question of principle—is whether his conduct can be ascribed fault: it demonstrates that a loss of capacity for control is excusatory of negligence liability, since the resultant lack of blame *prima facie* excludes negligence; [50] and, importantly, that such a 'loss of control *may or may not be accompanied* by loss of consciousness'.

---

[45] *Dunnage* (n 5).

[46] Fleming (n 18).

[47] Fanning (n 4) 699.

[48] *Dunnage* (n 3) [146].

[49] *Mansfield* (n 19) 1266 (emphasis added).

[50] ibid 1268.

[51] Outside of a *Mansfield*-type scenario, therefore, this might manifest as a loss of *rational* control. [52]

Indeed, it is suggested even a case such as *Nettleship v Weston* [53]—a *prima facie* obstacle to a fault-oriented 'test in civil cases'—can be interpreted such that it does not impugn fault-based reasoning in principle. *Per* Lord Denning MR, '*[m]orally* the learner driver is not at fault; but legally she is liable to be *because she is insured* and the risk should fall on her.' [54] Given the contingent fact of the 'policy of the Road Traffic Acts' [55] that all road users should carry third-party insurance, '[…] the judges [saw] to it that [the defendant] [was] liable' *so that* the claimant may recover. [56] This is plainly 'artificial' [57] and inverted reasoning that puts the cart (the claimant's entitlement to recover) before the horse (the defendant's liability on the facts). But for the contingent facts involved in *Nettleship*, the defendant's liability could and *should* have fallen to a principled, subjective determination according to her fault.

Developing the expanded notions of capacity and voluntariness indicated by the *Mansfield* judgment, this article suggests it would have been quite open to the *Dunnage* court to accept the argument advanced by counsel for the defendant—framed here pithily by David Owen—that 'an actor fairly may be

---

[51] ibid 1267 (emphasis added).

[52] *Dunnage* (n 3) [145].

[53] n 13.

[54] ibid 700 (emphasis added).

[55] ibid 699.

[56] ibid 700.

[57] *Cook v Cook* (1987) 162 CLR 376 [5] (Brennan J).

held accountable … only if he was *at fault* in causing it, [and] *only if his choices that resulted in the [claimant's] harm fairly may be blamed*'. [58] Neither the deluded *Dunnage*, nor the insulinomic *Mansfield* defendant, were 'blameworthy' in their choices; [59] and neither, it follows, could 'fairly' be held to account for any resultant damage. It has been judicially suggested 'the law of negligence is concerned less with what is fair than with what is culpable'; [60] but, on a fault-oriented conception of 'fairness' such as that articulated by Owen and embodied at common law in *Mansfield*, these two norms should drive at the same conclusion.

Mansfield and *Dunnage* demonstrate the materiality of 'idiosyncrasies' to blameworthiness, to which, *Mansfield* reveals, negligence law responds; and 'if no blame can be imputed to the defendant, the action, based on negligence, must inevitably fail'. [61] In this light, the positive statements of commentators like Goudkamp supporting 'the fact' that negligence is exclusively a type of conduct (excluding considerations of mentality), on the premise that the law of negligence 'shuns the most reliable indicators of when the attribution of blame is warranted, such as deliberation and choice', [62] seem problematic. A list of such 'reliable indicators' must also include these legally relevant incapacitating 'idiosyncrasies'. This article supports Goudkamp's

[58] David Owen, 'Philosophical Foundations of Fault in Tort Law' in David Owen (ed), *Philosophical Foundations of Tort Law* (Clarendon Press 1995) 228 (emphasis added).

[59] *Dunnage* (n 3) [156]; *Mansfield* (n 49) respectively.

[60] *Bolton* (n 7) 868 (Lord Radcliffe).

[61] *Snelling v Whitehead* [1998] RTR 386 (HL) (Lord Wilberforce).

[62] James Goudkamp, 'The Spurious Relationship Between Moral Blameworthiness and Liability for Negligence' (2004) Melbourne University Law Review 343, 351. See also Fanning (n 4) 700.

view, however, that *Dunnage*, in dismissing these factors material to blame, is 'seriously deficient'. [63] To the contrary, Spencer has argued that since the *duty* of care in negligence is determined by reference to an 'objective and reasonable standard', [64] whether that duty has been *breached* should be adjudged according to that same standard; [65] and so adjudged 'notwithstanding any mental or physical illness issue pertinent to the alleged tortfeasor'. [66] However, the finding of a duty is a preliminary issue concerned with whether the defendant's negligence is *capable* of being actionable; [67] but whether the defendant was legally negligent is a substantive question of fault—of 'blame', which these 'illness issues' materially inform.

The advantage, therefore—at least in principle—of the prerequisites for liability proffered by Owen is their systematic inclusion of these issues in the negligence inquiry. The prerequisites anticipate a conceptually logical symmetry between provision of compensation for the claimant and of due 'ethical retribution' [68] for the defendant; the latter is liable to compensate only where *at fault*, having made 'blameworthy choices'. [69] Echoing Montaigne, Cane similarly contends that tort liability should be understood as premised on 'ethical principles of personal

---

[63] Goudkamp and Ihuoma (n 4) 138.

[64] *Caparo v Dickman* [1990] 2 AC 605 (HL) 617-18 (Lord Bridge).

[65] Spencer (n 4) 201.

[66] ibid.

[67] *Caparo* (n 64) 629.

[68] Glanville Williams, 'The Aims of the Law of Tort' (1951) 4 CLP 137, 140.

[69] G Williams and BA Hepple, *Foundations of the Law of Tort* (2nd edn, Butterworths 1984) 136.

responsibility' for conduct. [70] If indeed, in Weinrib's terms, correlativity is negligence's 'organising idea', underlying and structuring the claimant-defendant relationship, [71] *Dunnage* (like *Nettleship*, discussed above) suggests an irregularity. The claimant cannot suffer an actionable injustice unless the defendant *commits* an injustice—'the same injustice that the [claimant] has suffered'. [72] If, on account of the arguments presented in this article, a schizophrenic defendant is properly regarded as legally beyond reproach, the entitlement to recover, correlative to a breach of obligation, should not arise. That entitlement might instead be justified against alternative, outcome-based considerations, purposively favourable to the claimant. But, as *Nettleship* demonstrated—and probed further below—these considerations are normatively irrelevant to a correlatively-structured law of negligence. [73]

The conception of justice reflected in Owen's correlative liability conditions, embodied in *Mansfield*, is moreover borne out in the context of child defendants. As cited above, the case law in this vein develops the reasoning that a child's suppressed standard of care is commensurate with, and justified by, his undeveloped 'sense and circumspection'. [74] Materially, this state precludes an 'adult's realisation' of risk and likewise an adult's capacity for *ethical* reasoning, otherwise 'effective as a control' on his conduct. [75] It

[70] Peter Cane, *The Anatomy of Tort Law* (Hart Publishing 1997) 24-5.

[71] Ernest Weinrib, 'Corrective Justice in a Nutshell' (2002) 52(4) UTLJ 349.

[72] ibid 351. See also Fleming (n 18) 179.

[73] Weinrib (n 71) 352.

[74] *McHale* (n 16) [9].

[75] JL Mackie, *Ethics: Inventing Right and Wrong* (Penguin Books 1990) 213.

follows that children are correctly held 'both legally and morally less responsible for what they do'. [76] In every respect pertinent to responsibility, therefore, the cases of the schizophrenic and the child are analogous: both act on impulse, whether 'irresistibly' [77] or otherwise juvenilely, [78] precluding rational self-control; and, as Mackie reflects—achieving a more convincing and precise analogy than that dismissed by the *Dunnage* court that the mentally ill have a 'mental age of five' [79]—the adult psychopath is confined to a *childlike* incapacity for moral reasoning. [80]

Cane, in opposition, suggests the 'bad moral luck' of being endowed with 'deficiencies' in one's capacities is an 'ineradicable feature of the human condition', [81] such that it is morally legitimate to expect a defendant to 'correct' them. However, a likewise ineradicable feature of the human experience is the 'mistakes and failings of childhood … to which everyone is heir', [82] yet there is no corresponding ethical imperative for children to 'correct their deficiencies'. It does not follow that, simply because a deficiency arises by nature, it is morally legitimate to expect a defendant to correct it. At the very least, any such ethical imperative ought to depend on a claimant being aware of his illness—the *Dunnage* defendant was not. [83] On each of these

---

[76] ibid.

[77] *Dunnage* (n 3) [32].

[78] *McHale* (n 16) [9].

[79] *Dunnage* (n 3) [130].

[80] Mackie (n 75) 213.

[81] Cane (n 70) 51.

[82] Fleming (n 18) 26.

[83] *Dunnage* (n 3) [12], [92], [147].

bases, therefore, the acts of neither the schizophrenic nor the child can properly be criticised from an ethical or *legal* perspective: if it would 'offend against principles of justice' to expect of children the capacities of the reasonable man, [84] the same intuitional [85] principles of justice should require that these two classes of defendant receive similar legal treatment.

Supporting objectivity of legal fault, contrarily, is the entirely correct (ethical) argument advocating the comparative treatment of relevantly similar individuals: [86] unequal treatment by the law which otherwise 'appears arbitrary, and for which no sufficient reason can be given, is held to be unjust' [87] and is antithetical to the rule of law [88]—hence the application of a like 'standard of conduct that [the reasonable man] would set for himself *and require of his neighbour*'. [89] However, it must follow that the existence of a 'morally relevant' difference, such as the 'dissimilarities in ability' [90] of both children and the schizophrenic compared to the reasonable person, would by the same token render their identical treatment unjust: it has been demonstrated

---

[84] Irish Law Reform Commission, *Liability in Tort of Minors and the Liability of Parents for Damage Caused by Minors* (Rep 17, 1985) 58.

[85] See, e.g., Steven Pinker, *The Language Instinct* (Penguin Books 1995) 420, who hypothesises on the existence of an innate cognitive 'module' for perceiving and responding to justice and injustice.

[86] William Frankena, *Ethics* (Prentice Hall 1963) 39.

[87] Henry Sidgwick, *The Methods of Ethics*, Book 3, Chapter 5 (first published 1874, Cambridge University Press 2011).

[88] Thomas Bingham, *The Rule of Law* (Penguin Books 2011) 55-59.

[89] *Bolton* (n 7) 869 (Lord Radcliffe) (emphasis added).

[90] Frankena (n 86) 41.

above that physical/mental impairment is legally relevant; [91] and if relevant *similarity* justifies equal treatment, relevant *dis*similarity, 'idiosyncratic' to the defendant, should justify differential treatment. Viewed from a (legal) equality perspective, the threat presented by the status quo to the integrity and legitimacy of the objective standard is recognised even by proponents of objective liability criteria, like Avihay Dorfman. [92] However, Dorfman's suggestion [93] that the 'egalitarian structure' of the objective standard is restored if compliance with its demands is understood as requiring moderation of the *choice* of conduct, rather than how conduct is undertaken—a futility given a mentally ill defendant's insufficient caring skills—is contestable. [94] It may be a consequence of mental impairment, as Arden LJ's *dictum* cited above attests, that even the facility for such decision-making becomes effectively foreclosed.

A subjective standard, then, appears preferable on both ethical and egalitarian grounds. Nonetheless, a purely objective standard *is* defensible—indeed, desirable—if viewed in light of a conception of justice, qualitatively different to that contemplated by Owen, that is rooted in 'the practical necessity of subordinating ethical values to the more exigent demands of public safety'. [95] Negligence law, understood in view of these overriding public policy demands, would thus be justified in 'exact[ing] a degree of care commensurate with the risk' created by the defendant's

---

[91] *Mansfield* (n 19); *Dunnage* (n 3) [104], [127], [159].

[92] Avihay Dorfman, 'Reasonable Care: Equality as Objectivity' (2012) 31(4) Law and Philosophy 369.

[93] ibid 381-85.

[94] Dorfman himself recognises the weakness of this solution, albeit from a different angle, at 386-88.

[95] Fleming (n 18) 22.

activity, [96] rather than with his own capacities. The considerable weight attached to public policy is driven by the impropriety of denying an innocent claimant recovery, solely on account of his transgressor's peculiar idiosyncrasy, despite its incapacitating effect. [97] Where a claimant cannot reasonably appreciate with sufficient foresight the defendant's qualified capacity for taking care, and reasonably modify his conduct accordingly, it cannot be *fair* to shift the loss resulting from the putative negligence onto him. As the law currently recognises, only where the claimant is himself 'responsible', in view of his *own* negligence, is it in principle correct that he should bear a measure of the loss. [98]

Thus, whilst nonstandard capacities may reasonably be expected of children and professionals, this is generally untrue of the mentally ill or the inexperienced: the ability to recover should not depend on the vagaries of chance, defeated by the misfortune of receiving an injury attributable to inexperience; [99] nor, similarly, on what the claimant happens subjectively to know of the defendant's capacities. [100] The enjoyment of coequal individual privilege by members of society to 'move freely' in that society necessarily entails the generic acceptance of coequal mutual obligations of care, legitimating a singular objective standard governing all conduct. [101] As Cane notes, an objective standard achieves 'fairness in loss allocation', thereby justifying the

---

[96] *Read v Lyons* [1947] AC 156 (HL) 173 (Lord Macmillan).

[97] Claimant-centric policy considerations have a considerable length of standing: see *Bessey v Olliot & Lambert* (1681) 83 ER 244.

[98] Law Reform (Contributory Negligence) Act 1945 c.28, s 1.

[99] *Wilsher v Essex AHA* [1987] QB 730 (CA) 750.

[100] *Nettleship* (n 13) 700-01, 709.

[101] Some weight was given to this point in *Dunnage*: see *Dunnage* (n 3) [153].

inevitable 'gap' between legal and ethical responsibility: [102] the loss resulting from the defendant's acts, where neither defendant nor claimant are 'at fault' in the respects considered above, is fairly borne by the defendant who created the risk of loss. As suggested above in connection with *Nettleship*, this is especially so, perhaps, where the tortfeasor—or his estate, as in *Dunnage*—is insured. [103] It is a conventional absurdity to assume every defendant enjoys third party insurance coverage; but by enabling the claimant to recover without personally impoverishing the defendant, it can significantly relieve both the 'admonitory' force of liability and the requirement for 'fault' a punishment presupposes. [104]

Regrettably, in consequence of *Dunnage*, this article does not share the enthusiasm of commentators like Richard Mullender who extol the 'reflexivity' of negligence law, whose 'informing ideals', such as justice, reliably influence how its doctrine is configured. [105] Any true commitment to what Mullender terms 'ethical community', [106] qualifying 'ordinary notions of what is fit and proper' conduct, [107] is doubtful in light of *Dunnage* and its normalisation of a circumscribed involuntariness rule.

---

[102] Peter Cane, 'Justice and Justification for Tort Liability' (1982) 2 OJLS 30, 48.

[103] *Nettleship* (n 13) 699-700; ibid. Speculatively, this may, at least implicitly, have inclined the *Dunnage* court towards a finding of liability in spite of the defendant's lack of fault (as this article contends).

[104] Fleming (n 18) 6.

[105] Richard Mullender, 'The Reasonable Person, The Pursuit of Justice, and Negligence Law' (2005) 68(4) MLR 681, 687-88.

[106] ibid 688.

[107] *McFarlane v Tayside Health Board* [2000] 2 AC 59 (HL) 108 (Lord Millett).

This circumscription would make sense if it can be accepted that 'the criterion of liability in tort is not so much culpability, but on whom should the risk fall' [108] for the claimant-centric reasons considered above; but, it is suggested, there is *no* such singular criterion. This criterion cannot explain the law's mitigation for children; [109] nor, even, why there was no liability in *Mansfield*, since, on the facts, a large corporation is far better equipped to shoulder loss than a small family enterprise. Indeed, as the no-fault compensation practice in New Zealand since 1974 bears out, the logical development of this criterion would appear to favour the abolition of tortious liability altogether—the 'death of tort'. [110] As Orchard comments, [111] given that the reasoning in *Dunnage* 'only becomes plausible' if this criterion of liability prevails, the circumscribed voluntariness rule—a product of that reasoning—is in some doubt. The final section of this article proposes to resolve that doubt, by means of a legislative solution.

4. Proposal for reform

This article submits that the involuntariness rule, having been unduly curtailed by the courts and whose current place in the scheme of negligence is ambiguous, [112] should be reconceptualised and given a statutory basis, furnishing 'incapacity' with a practicable definition. A statutory solution particularly commends

[108] *White v White* [1950] P 39 (CA) 59 (Denning LJ).

[109] Goudkamp and Ihuoma (n 4) 142.

[110] Fleming (n 18) 172.

[111] Orchard (n 4) 370-71.

[112] Goudkamp and Ihuoma (n 4) 145-49.

itself to those who, in light of *Dunnage*, are reasonably sceptical about the capacity of the common law of negligence to be 'reflexive'. Following closely the thread laid by Canadian jurisprudence—in particular, *Buckley v Smith Transport Ltd* [113] and, later, *Fiala v MacDonald* [114]—the proposal mooted here could take the following form:

> A court may relieve a defendant to an action in negligence of liability when, on the balance of probabilities, he is capable of demonstrating, as a result of recognised mental illness:
>
> (1) He had no capacity to understand or appreciate the duty of care owed at the relevant time; or
>
> (2) He was unable to discharge his duty of care as he had no meaningful control over his actions at the time the relevant conduct fell below the objective standard of care.
>
> The statutory discretion does not apply where, at the relevant time, the defendant was, or should reasonably have been aware of his mental illness.

Crucially, this reform does not impugn the objective standard, which, in general, remains a desirable referent for the reasons canvassed above. However, it does empower the court to extend legal protection, hitherto withheld, [115] to the mentally ill

[113] [1946] OR 798 (CA).

[114] [2001] 201 DLR (4th) 680. See in particular [49].

[115] Though, it is submitted, there are strands of English judgments that might support the principle of the reform advocated: see, e.g., *Corr v*

whose impairment genuinely (or, at least, on the balance of probabilities) vitiates the capacity to recognise or discharge a duty of care. Thus, on the facts of *Dunnage*, given the schizophrenic defendant's apparent lack of 'meaningful' ability to take 'any alternative action' other than the course he took, [116] the second limb (at least) is engaged and the statutory discretion applies. It may, after all, be countered that this proposal presupposes *too* great a weight attached to ethical considerations of fault—the main focus of this article—but which are, in fact, of little moment to negligence law. [117] Significantly, however, even the early proponent of compensation-based objectives in tort law, Oliver Wendell Holmes, [118] recognised the differentiable situation of the mentally impaired. Thus, where mental impairment 'manifestly incapacitat[es] the sufferer from complying with the rule which he has broken, good'—and, it may be added, *moral*—'sense would require it to be admitted as an excuse.' [119]

5. Concluding remarks

Dunnage v Randall exemplifies the tension between two competing objectives—one claimant-oriented, one defendant-oriented—that the common law on the standard of care has variously, and with some inconsistency, pursued. The reform this article has advanced

---

IBC Vehicles Ltd [2008] UKHL 13, [2008] 1 AC 884 (HL) [16], [22], [63].

[116] *Dunnage* (n 3) [143].

[117] e.g. Mullender (n 105) 691-92.

[118] Oliver Wendell Holmes Jr, *The Common Law* (first published 1881, Dover Publications 1991) 108.

[119] ibid 109.

is an essential one, for it resolves the apparent inconsistency in the treatment by the law of defendants with relevantly similar incapacities. It is practical and affords real consequence to the observation in *Dunnage* that 'it is the effects of a condition or illness which ought to be in play, not its label,' [120] eliminating the *de facto* line between physical and mental illness as regards negligence liability which was ironically consolidated by *Dunnage*. [121] Beyond its immediate application to mentally incapacitous defendants, such a reform could have broader social ramifications: it strips of currency any lasting anachronistic conceptions of mental illness as 'beset with an atavistic attribution of sin from which modern man has not yet succeeded in emancipating himself.' [122]

[120] *Dunnage* (n 3) [104] (Rafferty LJ).

[121] Orchard (n 4) 371. See also *Clerk & Lindsell On Torts* (22nd edn, Sweet & Maxwell 2017) para 8-161.

[122] Fleming (n 18) 26.

Is Time Up for Recovery of Time Value?

Mehleen Rahman*

Abstract—This article looks at *Prudential Assurance*, the latest judgment in a long line of overpaid tax litigation cases. The Supreme Court, departing from the House of Lords decision in *Sempra Metals Ltd v IRC*, found that the claimants were not entitled to recover compound interest on taxes levied in breach of EU law. This article is critical of the reasoning of the Supreme Court, in particular, of the treatment of 'use value' of money and the narrowing of the 'at the expense of' requirement in unjust enrichment.

1. Introduction

In *Prudential Assurance v Revenue & Customs Commissioners* ('*Prudential*'),[1] Lords Reed, Hodge, and Mance handed down the judgment in the Supreme Court, holding that compound interest was not recoverable in that case.[2] The court held that there is no action in unjust enrichment to recover compound interest

* University College London. I am grateful to the OUULJ editorial team for comments made on previous drafts. All errors remain my own.

[1] [2018] UKSC 39, [2018] 3 WLR 652.

[2] ibid [79].

representing the loss of the use of money ('use value') which the claimant had paid to the defendant. [3] In order to reach this conclusion, the Court notably moved away from the landmark decision of the House of Lords in *Sempra Metals Ltd v IRC* ('*Sempra*'),[4] which had allowed claims in unjust enrichment for the recovery of the time value of money and, additionally, had held that the relevant restitutionary award can be calculated on the basis of compound interest. *Prudential* is significant for two reasons. First, the Supreme Court held that a previous decision of the House of Lords was no longer good law. Secondly, and more broadly, the Court held that the use value of money may not be a relevant enrichment under the law of unjust enrichment.

This article will first briefly discuss the background decisions of the House of Lords in *Sempra* and of the European Court of Justice ('ECJ') in *Test Claimants in the FII Group Litigation v Commissioners of Inland Revenue* ('*FII Litigation*').[5] The article will then argue that the analysis of the Supreme Court in *Prudential* leaves much to be desired: the Court effected a significant change in the law of unjust enrichment on the basis of reasons which, once scrutinised, are largely unsound.

2. Background Decisions

A. Sempra

The claim in *Sempra* arose after the ECJ held that a tax regime allowing resident parent companies, but not non-resident companies, to make an election such that its subsidiaries did not

[3] Ibid.
[4] [2007] UKHL 34, [2008] AC 561.
[5] C-446/04.

have to account for advance corporation tax on dividends paid to it, was contrary to Article 52 of the European Community ('EC') Treaty, which guarantees the freedom of establishment.[6] In *Sempra*, the House of Lords was concerned with whether the calculation of the award required by Community law in the circumstances of this breach was to be on the basis of simple interest or compound interest.[7] The House of Lords held, inter alia, that a claim in unjust enrichment lies for the recovery of the time value of money,[8] in the form of an award of compound interest.

B. FII Litigation

Prudential concerned a test claim brought against HMRC, under a group litigation order made following the decision of the ECJ in *Test Claimants in the FII Group Litigation*. The issue in the *FII Litigation* proceedings arose due to the differential treatment under the UK tax regime of dividends received by parent companies resident in the UK. Under s. 208 of the Income and Corporation Taxes Act 1988, where a UK resident company received dividends from a UK resident subsidiary it was not liable to pay corporation tax in respect of these dividends. In addition to this, the parent company would receive credit equal to the amount of the corporation tax that the distributing subsidiary had paid in advance. On the other hand, UK resident companies, receiving dividends from subsidiaries resident outside the UK, received no such allowances. They were liable to pay corporation tax on the dividends received and, additionally, they received no

[6] *Metallgesellschaft Ltd v Inland Revenue Commissioners and Hoechst AG v Inland Revenue Commissioners* (Joined Cases C-397 and C-410/98) [2001] Ch 620.
[7] *Sempra* (n 4) [12]
[8] ibid [33]

tax credit. The ECJ held that this disparity in the treatment of dividends received fell afoul of Article 56 of the EC Treaty.[9]

3. Questions before the Court in Prudential

While there were several issues to be resolved before the Supreme Court as a result of the decision of the ECJ, this case-note will focus on the Supreme Court's treatment of the question of whether the claimant could claim compound interest in respect of the sum it had unlawfully paid to the Revenue, on the basis that the Revenue was unjustly enriched by the opportunity to use this sum. The Supreme Court held that the claimant could not recover compound interest. Furthermore, their Lordships departed from *Sempra* to the extent that it allowed a claim in unjust enrichment for the recovery of compound interest, representing the 'use value' of money.[10]

4. The Decision of the Supreme Court in Prudential

The Revenue argued that the opportunity to use money mistakenly paid is not a benefit obtained at the expense of the

[9] Prohibiting restriction on the free movement of capital and payments between Member States and between Member States and third countries.
[10] *Prudential* (n 1) [79].

payor. The claimant opposed this, arguing that this line of argument should not be allowed as the Revenue had conceded that the claimant was entitled to recover compound interest on the basis of *Sempra*. The Supreme Court disagreed with the claimant, allowing the Revenue to proceed with their line of argument due to significant developments in the law of unjust enrichment that had occurred since the trial at first instance.[11] The Court found that, in light of these developments, a critical analysis of the reasoning in *Sempra* was necessary. The Court provided 5 reasons that formed the basis of its decision to depart from *Sempra*. As will be explored below, except for the first two considerations of the court, namely, the development of the law since *Metallgesellschaft*, and an apparent omission by the House of Lords in *Sempra*, the majority of the justifications provided are unsound.

A. Development of the Jurisprudence of the CJEU since Metallgesellschaft

The Court of Justice of the EU ('CJEU') in *Metallgesellschaft* held that the tax regime in the UK, specifically s. 247 of the Income and Corporation Taxes Act 1988 ('ICTA'), was contrary to EU law.[12] The regime allowed UK-resident parent companies, but not Member State-resident parent companies, to elect when they paid corporation tax in relation to the dividend they received from their subsidiaries. Parent companies resident in Member States were deprived of this option of election and had no choice but to pay advance corporation tax. On the question of remedy, the CJEU in *Metallgesellschaft* held that the sum due under EU law, pursued under a claim in restitution, would be 'the amount of

[11] ibid [40].
[12] *Metallgesellschaft* (n 6) [76].

interest which would have been generated by the sum, use of which was lost as a result of the premature levy of the tax'.[13]

The Supreme Court stated that, since this pronouncement, the jurisprudence of the CJEU has developed. First, the CJEU has since stated in *Littlewoods* that it is for the internal legal order of Member States to determine the conditions under which such interest arises and the rate at which it is calculated.[14] Second, pursuant to this domestic procedural autonomy, the Supreme Court in *Littlewoods Limited v Commissioners for her Majesty's Revenue and Customs*[15] had held that where the claimant had lost the opportunity to use money, an award of simple interest is an adequate indemnity under EU law.

This justification of the Supreme Court cannot be faulted. It demonstrates that the Court was cognisant of the fact that recovery of compound interest in an action in unjust enrichment was not required by the principle of effectiveness under EU law.[16] This is in fact an important consideration because it means that, in deciding whether the claimant was entitled to recovery under unjust enrichment, the Court rightly held that it was not bound by EU law to provide interest on a compounded basis.

B. Conflict between Sempra and Prior Legislation

The Court questioned the decision of the House of Lords in *Sempra* on the basis of the majority's failure to have regard to

[13] ibid [88].
[14] *Littlewoods Retail Ltd v Revenue and Customs Commissioners* (Case C-591/10) [27].
[15] [2017] UKSC 70, [2017] W.L.R. 1401.
[16] *Prudential* (n 1) [56].

provisions of prior legislation when it decided to award compound interest.[17] The legislative schemes noted by the court were ICTA and the Value Added Tax Act 1994 ('VATA'). S. 78 of VATA contains provisions for the payment of simple interest on overpaid tax. S. 826 of ICTA also provides for the payment of simple interest in the case of overpaid taxes, covering advance corporation tax and mainstream corporation tax.

In allowing a common law claim for the recovery of interest on the basis of mistaken payments, the Supreme Court suggested that *Sempra* eroded the legislative schemes in place for the regulation of payment of interest on overpaid taxes.[18] Per the Supreme Court in *Prudential*, the fact that the court in *Sempra* did so without referring to the relevant legislative schemes diminished the persuasiveness of its reasoning.[19]

Once again, this is submitted to be a relevant consideration. The fact that the House of Lords in *Sempra* had effected a major change in the law without reference to the relevant background legislative schemes leaves the decision vulnerable to criticism. Not only did the court in *Sempra* fail 'to proceed in harmony with Parliament',[20] but in failing to refer to the legislative provisions, the court did not explain why it felt the need to render a judgment that was quite clearly at odds with the existent statutory regime. An omission as conspicuous as this can rightly be used to criticise *Sempra*, on the basis that the court's process of determination was not, or at the very least did not appear to be, as comprehensive as it ought to have been.

---

[17] ibid [59].
[18] ibid [60].
[19] Ibid [59].
[20] *Johnson v Unisys Ltd* [2003] 1 AC 518 [37].

C. Legislation's failure to solve problems arising from Retrospective Effect of Kleinwort Benson

A third consideration for the Supreme Court was the impact of the decision of the House of Lords in *Kleinwort Benson v Lincoln City Council* ('*Kleinwort Benson*')[21] in conjunction with its decision in *Sempra*. In *Kleinwort Benson*, the House of Lords abolished the mistake of law bar, thereby enabling claims in unjust enrichment for recovery of money paid under mistakes of law. [22]

Kleinwort Benson concerned a dispute under the swaps agreement line of litigation in the 1990s. In *Hazell v Hammersmith & Fulham Borough Council*,[23] the House of Lords held that the interest rate swaps agreements entered into by local authorities were ultra vires and, as a result, such agreements were void. This resulted in a line of claims being brought for the restitution of sums paid under the void agreements. The claimant bank in *Kleinwort Benson* wished to ground its claim in restitution on mistake, as this would enable it to rely on s.32(1)(c) of the Limitation Act 1980, which stipulated that the period of limitation in the context of mistake claims does not begin to run until the claimant has discovered the mistake or until the time the claimant would reasonably have been expected to discover the mistake. Problematically however, prior to *Kleinwort Benson*, the law did not allow recovery of money paid under a mistake of law. The House of Lords, by a majority of 3-2, abolished the mistake of law bar. One of the differences between the majority and the minority was whether this decision should have the retrospective effect of rendering *Kleinwort Benson*'s payment, which was lawful *prior* to the

[21] [1999] 2 AC 349.
[22] ibid 373.
[23] [1992] 2 AC 1.

decision in this case, mistaken. The majority held that the decision in *Kleinwort Benson* did have such a 'retrospective effect'.

The Supreme Court in *Prudential* noted that the effect of holding that the decision would have retrospective effect was to enable 'claims to be brought within six years of the mistake being discovered, no matter how long in the past the payment had been made'.[24] The House of Lords in *Sempra*, like the House of Lords in *Kleinwort Benson*, made drastic changes to the law, believing that Parliament would enact legislation to prevent 'serious untoward consequences' for the Revenue.[25] The Supreme Court in *Prudential* noted however, that the legislature had been unsuccessful in addressing the problem arising from the retrospective nature of the decision in *Sempra*.[26] As such, the decision of the House in *Sempra* is questioned by the Supreme Court because it was premised on a belief that legislation would prevent adverse consequences resulting from the operation of s. 32(1)(c) – a belief which has not been substantiated by subsequent enactment of legislation.

The basis of this criticism of *Sempra* is faulty. The Supreme Court is correct in observing that the government has failed to exclude or restrict the operation of s. 32(1)(c) to its own detriment. However, contrary to what the Court seems to suggest initially, this failure of the legislator is not due to the fact that such exclusions or restrictions are impossible: the problems arising from the retrospective effect of decisions in *Kleinwort Benson* are not 'incapable of being fully addressed by legislation'. This is evidenced in the very judgment in *Prudential*, a few paragraphs down, when the Court contradicts itself by acknowledging that the legislator could address the issue by including 'a reasonable

[24] *Prudential* (n 1) [61].
[25] Ibid.
[26] ibid [62].

transitional period'.[27] It may be argued that this deficiency in the legislation forced the Supreme Court to stem litigation being brought against the Revenue for restitutionary awards in the order of billions of pounds. However, this justification cannot stand once we notice that the holding in *Prudential* does not apply solely to claims in unjust enrichment against the Revenue: the impact of this result-led decision-making entails ramifications in cases beyond the tax context. As I argue under the next sub-heading, what is unacceptable here is that in such cases, the blanket narrowing of the scope of claims in unjust enrichment may not be as desirable as it supposedly was in *Prudential*.

In any case, it is submitted that the Court's method of presenting the legislature's failure does not really acknowledge the fact that there are other remedies available to stem the claims being brought against the Revenue. This is problematic because one of the reasons the Court gives for departing from *Sempra* is the impossibility of addressing the flood of tax recovery litigation. But as the judgment itself acknowledges, there is no such impossibility; the legislature is capable of addressing the problem, it had simply failed to do so. One such failed attempt of enacting legislation[28] to this effect was struck down by the House of Lords in *Michael Fleming v Customs & Excise* for lacking the transition arrangements necessary under EU Law.[29] It follows that in the absence of such impossibility, it is reasonable to wonder whether the Court had sufficient reasons to justify its departure.

D. Disruption of Public Finances

The Court noted that the amount the Revenue would be liable to pay in the present case, if recovery of compound interest was

[27] ibid [64].
[28] Value Added Tax Regulations 1995 reg.29.
[29] [2008] UKHL 2, [2008] 1 W.L.R. 195.

allowed, would be in the order of £4-5 billion. In the same vein, the Court noted the sum in question in Littlewoods was £17 billion. These large sums in dispute were yet again the result of the applicable limitation period and the compounding of interest. Because the limitation period did not start running until the discovery of the mistake, claims in this area were capable of being backdated for several decades.[30] This was especially problematic where the principle amount was incurring not only simple interest, but compound interest over this extended period.

It is apparent that the enormity of previously successful claims, potential claims and their potential impact on public finances, was a strong consideration for this Court.

Taking into account the impact of allowing recovery on a particular basis on public finances is not problematic *per se*. This area of restitution is difficult precisely because of the fact that there are underlying policy considerations which pull in opposite directions. More specifically in cases of taxes exacted where they are not due, the constitutional principle that there is to be no taxation without Parliamentary authority seems to go against the protection of the Revenue and the prevention of fiscal chaos. This can be demonstrated with reference to the *Woolwich* ground of recovery for example, where it has been argued that the very nature of claims brought by private citizens against public authorities (such as HMRC) give rise to considerations of a public law nature, which courts hearing a private law claim in unjust enrichment should, nevertheless, pay heed to.[31] However, the problem with result-motivated decision-making is that quite often, the court is focussing on the implications ensuing from the factual matrix of the case in front of them. In turn, this often causes courts to overlook the fact that the law established in this

[30] *Prudential* (n 1) [65].
[31] Williams, *Enrichment and Public Law: A Comparative Study of England, France and the EU* (Hart Publishing, 2010).

specific case will have a knock-on effect on cases across the relevant area of law. Such is the case here, because the policy-motivated decisions in the Revenue cases do and will hold true across the law of unjust enrichment. While there may be justifications for restrictions on recovery where the financial consequence otherwise would be devastating for HMRC, it does not follow that such restrictions, as between two private parties, a mistaken payor and an unjustly enriched payee, are equally justified. Yet, the holding of *Prudential* has precisely this impact. So whilst narrowing the scope of recovery appears to be the desirable decision in the context of a claim against HMRC, effecting change in the law of unjust enrichment on this basis alone ignores that fact that this narrows the scope of recovery in all contexts and may yield undesirable results.

E. Development of 'at the expense of' element in ITC v Revenue

In order to make out a claim in unjust enrichment, a claimant must make out four elements.[32] First, that the defendant was enriched. Secondly, that the enrichment was at the expense of the claimant. Third, there must be an 'unjust factor'.[33] Finally, there must be no defence on which the defendant can depend. The previous considerations outlined above demonstrated that the Supreme Court was ready to effect change in the law of unjust enrichment. In this case, the Justices effected this desired change through the second element of an unjust enrichment claim: by holding that the Revenue had not been enriched at the expense of the Claimant.

---

[32] *Benedetti v Sawiris* [2014] AC 938 [10].

[33] For example, absence of consideration, lack of authority and want of consent etc.

First, the Court noted that in *Investment Trust Companies v Revenue and Customs Commissioners* ('*ITC*'),[34] the Supreme Court had narrowed the scope of the 'at the expense of' element of a claim and consequently, the scope of unjust enrichment. The facts of *ITC* are significant if one is to understand the Supreme Court's reasoning in the present case. In *ITC*, the claimant companies had paid VAT to managers for their provision of investment management services. The managers subsequently paid the VAT to HMRC. However, under EU law, the charging of the VAT in such cases was unlawful. The managers had a claim to recover the VAT paid under s. 80 of the Value Added Tax Act 1994. However, claimant companies were dissatisfied with this for two reasons. First, the managers could only recover VAT paid within the three-year limitation period under s. 80. Secondly, the managers were entitled to, and they had, deducted VAT they paid to their own third-party suppliers from the VAT paid to HMRC. 'On the notional figures used by way of illustration in the case, the managers had deducted £25 from £100 VAT paid to them by the claimants and had therefore paid to HMRC, and recovered from HMRC, and passed on to the claimants, only £75'.[35] Proceeding under these notional figures, claimant companies brought a claim for the £25 paid during the three-year period. Additionally, claimant companies brought a claim for the sum paid that was time-barred under s. 80 of VATA. The Supreme Court held that HMRC had not been enriched by the notional £25. Further, the Court held that the VAT received was not enrichment at the expense of the claimant companies. The claimant had not transferred any value directly to HMRC, rather, the managers had. The Court held that 'at the expense of' required direct conferral of benefit, rather than a causality between

[34] [2017] UKSC 29, [2017] 2 W.L.R. 1200.
[35] Burrows, 'Narrowing the scope of unjust enrichment' (2017) 133 LQR 537.

defendant's gain and claimant's loss. The claimant's case failed on this ground.

The Supreme Court in *Prudential* applied this direct conferral requirement to the facts of the case and found it not to have been satisfied. The decision in *Sempra*, which held that in addition to the sum paid, the defendant was also enriched by the use of the money, was held to have 'questionable features'.[36] The Supreme Court's reasoning in criticising *Sempra* is worth quoting in full:

> 'If on 1 April the claimant mistakenly pays the defendant £1,000, with the result that the defendant is on that date obliged to repay the claimant £1,000, the defendant's repayment of £1,000 on that date will effect complete restitution. Restitution of the amount mistakenly paid in itself restores to the claimant the opportunity to use the money: there is no additional amount due in restitution. That is because there has been only one direct transfer of value, namely the payment of the £1,000. The opportunity to use the money mistakenly paid can arise as a consequence of that transfer, but a causal link is not sufficient to constitute a further, independent, transfer of value. *Contrary to the analysis of Lord Nicholls in Sempra Metals [2008] AC 561, para 102, the recipient's possession of the money mistakenly paid to him, and his consequent opportunity to use it, is not a distinct and additional transfer of value*.'[37]

Unfortunately, the reasoning here does not stand up to scrutiny. The Court is correct in holding that on 1 April, where the defendant pays the claimant £1,000, this repayment in itself will amount to complete restitution. However, the fact that the £1,000 of itself restores the claimant with the opportunity to use money is not due to the fact that the use value of money is

---

[36] *Prudential* (n 1) [71].
[37] ibid (emphasis added).

incapable of enriching the defendant as a distinct transfer of value. Rather, in this instance, the repayment is immediate, such that defendant has simply not been benefitted by the use value of the £1,000. To put it in a different way, because there has been no passage of time, there has been no opportunity to use the money, and so there is no use value. A conclusion that full repayment of the sum itself will always be the full measure of restitution does not follow from this example. Yet, this is precisely what the Court goes on to assert, transposing the first scenario of immediate repayment into the next scenario:

> 'The position is essentially the same if the £1,000 is repaid not on 1 April but on 1 May. There has been no transfer of value subsequent to 1 April, when the mistaken payment was made. The only transfer of value needing to be reversed remains the payment of the £1,000.'[38]

The position clearly is not the same in this scenario. A key distinction between the first scenario and the second scenario is the passage of time. In the first scenario, no time had passed between the claimant's payment and the defendant's repayment. As such, transfer of value, in the form of time value of money, could not have arisen. In the second scenario, this is not the case: a month has elapsed since the defendant received the money and, in that time, the defendant has been enriched by the opportunity to use the sum received. Not only is D enriched, but he is also enriched *at the expense of* C, who has lost the use of his money during this time, and consequently lost out on the revenue he could have gained with this money. According to the reasoning of the court in *Benedetti*,[39] this lost value should be calculated by reference to the price which a reasonable person in the

[38] ibid [72].
[39] *Benedetti* (n 32).

defendant's position would have had to pay to borrow an equivalent sum of £1000 for a term equivalent to the period of D's retention.

In other words, the first scenario where no time elapses and the second scenario, where a month elapses, are similar because in both cases, the defendant received two things: the face value of the money and the use value of the money. The fact that no time elapses in the first scenario does not change the fact that the defendant received two things – it alters the valuation of the use value of the money received, as this valuation is contingent on time. As such, it is incorrect to state that because in the first scenario, the total enrichment is the face value of money, it follows that in the second scenario, this will also be the case.

In the alternative, it is also possible to see this part of the judgment as holding that there has been a transfer of value in the form of the use value of money, but the law of unjust enrichment does not deem it to be a relevant transfer. In this case, we need to know what qualifies as a relevant transfer. Furthermore, if this benefit does not come from the claimant, then the question arises, where does the benefit come from? The judgment seems to suggest that the opportunity to use money arises from the defendant's failure to repay the debt.[40] If this is the case, then we face the problem of explaining how this accrual occurs.[41] It is submitted that the better view is to look at the benefit obtained by the defendant: it seems contrary to logic and common sense to suggest that benefit obtained from non-repayment of a debt (where non-repayment necessarily disadvantages a creditor) is not at the creditor's expense. It is when we look at it from this perspective that it becomes evident that the defendant's benefit would not have occurred but for the claimant's deprivation:

---

[40] *Prudential* (n 1) [74].

[41] Mitchell, 'End of the Road for Overpaid Litigation' [2018] The UK Supreme Court Yearbook (forthcoming).

namely, the face value and the use value of the money. It is submitted that the focus on transfer here merely leads to confusion and led the Supreme Court to an untenable conclusion.

As a final point, although the court extends the reasoning of *ITC* to the present case, specifically the directness requirement, it is important to note that on the facts, *ITC* was concerned with a wholly different scenario from the one in *Prudential*.[42] The issue in *ITC* was whether the claimant could bring a claim against HMRC, which was an indirect recipient (because claimant companies had paid the money to the managers, who then paid the tax to HMRC). *ITC* was therefore concerned with claims in unjust enrichment being brought against remote recipients. This is different from the present case, because the direct conferral requirement stemmed from the desire to restrict recovery in cases where the defendant was enriched by a party which was itself enriched by the claimant. *Prudential*, on the other hand, was concerned with the recovery of the opportunity to use money. It seems odd that the directness requirement in a multi-party case, notwithstanding the lack of clarity of the directness requirement itself, is being used to reach and justify a finding in *Prudential*, where the latter case occurs under a wholly different factual matrix. Here, as between two parties, the term 'directness' cannot mean the same thing as it did in *ITC*, as there simply were no indirect recipients of money.

5. Conclusion

The Supreme Court in *Prudential* gave five reasons for departing from the decision in *Sempra*, in so far as it allowed the recovery of

[42] Ibid.

compound interest for claims in unjust enrichment for the loss of opportunity to use money. Of these five, two are tenable. The Court was correct to have regard to the fact that an award of simple interest is now sufficient compensation under EU law, where the claimant has lost the opportunity to the use of money that has been paid as tax under legislation which is contrary to EU Law. The Court was also correct to note that the decision in *Sempra* was open to criticism as it failed to acknowledge the inconsistency of its decision with the existent legislative framework.

The next three reasons are highly problematic. The Court is incorrect in suggesting that legislation is incapable of addressing the problem of overpaid tax recovery litigation: all that is required is the inclusion of a transition period. Further, in taking account of the impact of such litigation on public finances, the court engages in result-led jurisprudence which narrows the law of unjust enrichment and will apply across the board in non-HMRC cases, possibly leading to undesirable results. Finally, the most untenable aspect of the judgment is the modification of the 'at the expense of' element of unjust enrichment, where the Court's reasoning is simply faulty. In light of this assessment, it is submitted that despite the accuracy of the first two considerations, the weakness of the latter three arguments renders the Court's decision in *Prudential* untenable.

PUBLIC LAW ARTICLES

A Causal Model for the Extraterritorial Application of Human Rights Treaties

Abe Chauhan*

Abstract—This article explores the existing models of extraterritorial application of human rights treaties. It concludes that these are inadequate as they allow for situations where no jurisdictional link is found between a State party and a victim of a human rights violation even where the State party has the *ability* to prevent this violation. This article therefore proposes that the existing models be replaced with the legal test for determining such ability: a causal model, applying principles of foreseeability and remoteness, which determines whether the State party has control over an act or omission causing an extraterritorial rights violation.

1. Introduction

---

* University College London. I am grateful to Professor Jeff King and Professor Christian Walter, as well as the OUULJ editorial team, for their comments on earlier drafts. All errors remain my own.

Human rights are founded on the principle of universality. This is apparent in the preamble of the United Nations Declaration of Human Rights which considers 'recognition […] of the equal and inalienable rights of all members of the human family [to be] the foundation of freedom, justice and peace in the world'. A truly universal scheme of human rights protection envisages that all individuals are empowered to enforce these rights against any State party.[1] However, human rights treaties are drafted by States which seek to limit the universal protection of rights where this would incur their own international responsibility.[2] As a result, the obligations contained under these treaties are limited not only in content but also in the pool of rights-holders to whom they are

[1] International law is founded on the principle of State sovereignty ('The principle…of supreme authority within a territory': Samantha Besson, 'Sovereignty', *Max Planck Encyclopedia of Public International Law*). It is therefore only through the consent of individual States that schemes of human rights protection emerge on an international level. Thus, there is no intrinsic universal scheme of protection; this is established by treaties negotiated on a multilateral basis.

[2] For example, the US introduced amendments during the drafting of the ICCPR to ensure that the US' international responsibility would not extend to its overseas activities. Eleanor Roosevelt, the then US representative and Chairman of the UN Commission on Human Rights, made clear that the US would not assume 'an obligation to ensure the rights recognized in [the Treaty] to citizens of countries under United States occupation': Summary Record of the Hundred and Ninety-Third Meeting, UN ESCOR Human Rights Committee (1950) UN Doc E/CN.4/SR.193, [14] (Mrs Roosevelt). The US has since maintained this strictly territorial interpretation of the application of the ICCPR: Second and Third Periodic Report of the United States of America to the UN Committee on Human Rights Concerning the International Covenant on Civil and Political Rights (21 October 2005) Annex 1, available at <https://www.state.gov/j/drl/rls/55504.htm#annex1> accessed 25 April 2019.

owed; the universal scheme of protection envisaged in the Declaration does not exist.

In determining this pool of rights-holders through the extraterritorial application of human rights treaties (Part 2), a balance must be struck between protecting of the rights of all individuals and ensuring that obligations imposed on State parties can be effectively achieved (Part 3). The current models for determining extraterritorial application only roughly correlate to this notion of effectiveness and create a *prima facie* presumption that those outside of a State party's territory do not enjoy rights against it (Part 4). The presumption must be reversed with the adoption of a causal model of extraterritorial application (Part 5). Causation constitutes a legal test for determining whether a State party is able to effectively secure rights under human rights treaties and thus is most closely aligned with the guiding principles of human rights protection, namely universality and effectiveness. The causal model can be introduced through customary principles of treaty interpretation and should apply to both negative and positive obligations.[3] State parties will retain the possibility of being exculpated, but only when it can be shown that they were unable to effectively secure the relevant rights, because of an insufficiently direct causal chain or because they acted reasonably in preventing human rights harm, thus discharging their obligations of due diligence. It is only in such cases that a State party has acted in accordance with the object and purpose of the relevant human rights treaty and where deviation from the universality of human rights is permissible.

[3] These are obligations to respect and protect human rights, respectively. On this, see Dinah Shelton and Ariel Gold, 'Positive and Negative Obligations' in Dinah Shelton (ed), *The Oxford Handbook of International Human Rights Law* (OUP 2013) 562–584.

2. The exercise of extraterritorial jurisdiction

A human rights treaty is applied extraterritorially where, at the moment of a violation by State A of a right protected under the treaty, the right-holder is located outside State A's territory. The extraterritorial element, therefore, is the location of the affected individual at the point of the act or omission by State A, rather than the location of State A's act or omission.[4] While many extraterritorial situations also involve extraterritorial State acts, this is not required for the issue of extraterritoriality to arise. For example, in *Montero v Uruguay*,[5] the relevant rights violation was a decision not to issue a passport to the applicant; this decision was taken within Uruguay's territory with extraterritorial effect on the applicant, who was situated in Germany.

There is no default rule on the extraterritorial application of human rights treaties in international law;[6] some treaties expressly require State A to ensure rights to those in its 'jurisdiction',[7] while others have a notion of 'jurisdiction'

---

[4] Marko Milanovic, *The Extraterritorial Application of Human Rights Treaties* (OUP 2011) 8.

[5] Com No 106/1981 (HRC, 31 March 1983), UN Doc CCPR/C/18/D/106/1981.

[6] Milanovic (n 4) 9–11. Some treaty obligations are applied universally, for example, the obligation to prevent genocide under the Convention on the Prevention and Punishment of the Crime of Genocide: *Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosnia and Herzegovina v Serbia and Montenegro)* [2007] ICJ Rep 43. This article will focus on human rights obligations which are limited in application by the notion of 'jurisdiction'.

[7] For example, Article 1 ECHR, Article 1 (1) ACHR and Article 2 (1) ICCPR.

introduced through subsequent interpretation and application.[8] This is distinct from jurisdiction in general international law, meaning the permission to exercise legal authority, and refers instead to factual control.[9] Whether particular obligations under a human rights treaty can be applied extraterritorially is a question of interpretation of the relevant treaty,[10] and must therefore be determined with reference to the customary rules on interpretation codified in Article 31 of the Vienna Convention on the Law of Treaties ('VCLT').[11] Jurisdiction must be interpreted in line with the term's ordinary meaning and the object and purpose of the relevant treaty – Article 31(1) VCLT – as well as with relevant rules of international law – Article 31 (3) (c) VCLT.

3. The principles of universality and effectiveness

In determining jurisdiction under human rights treaties, international human rights courts and treaty bodies must have regard to the interlinking principles of universality and

---

[8] For example, the ICESCR (*Legal Consequences of the Construction of a Wall* (Advisory Opinion) [2004] ICJ Rep 136, [112]) and CEDAW (CEDAW, 'General Recommendation No. 28 on the core obligations of States parties under article 2 of the Convention on the Elimination of All Forms of Discrimination against Women' (16 December 2010) [12], CEDAW/C/GC/28).

[9] See discussion in Part 5.A.

[10] As the ECtHR noted in *Banković v Belgium*, App No 52207/99 (ECtHR, 12 December 2001) [55].

[11] See *Territorial Dispute (Libya v Chad)*, (Merits) [1994] ICJ Rep 6 1045 [41], where the ICJ confirmed the customary status of Articles 31 and 32 VCLT.

effectiveness. The universality of human rights[12] is a foundational principle of the International Bill of Rights[13] and its fundamental nature requires that it be deviated from only in line with the object and purpose of human rights treaties, namely the *effective* protection of human rights.[14]

State A is able to effectively secure the relevant rights when it has some degree of power or influence over the relevant actors involved in the rights violation. Human rights are dual in nature, requiring observance of negative and positive obligations,[15] and therefore State parties are obligated to prevent the harmful actions of both its agents[16] and, generally to a lesser extent, of private parties. This must be effective in the sense that

---

[12] There is continuing opposition to the notion that human rights are universal. It is argued that they are instead a Western ideal which cannot be transposed onto all cultures: on this, see Louis Henkin, 'The Universality of the Concept of Human Rights' (1989) 506 The Annals of the American Academy of Political and Social Science 10. While this may be true as to the *content* of the right, this essay proceeds on the basis that the *existence* of universal rights is a fundamental principle of the international legal order. In any event, the causal model of extraterritorial application would not impose rights standards on all countries, but merely provide redress to those in other territories causally affected by a State party's actions: Chimene I Keitner, 'Rights Beyond Borders' (2011) 36 Yale Journal of International Law 55, 66–68.

[13] Universality is reflected in a great number of international human rights instruments including the UNDHR (see Part I). The International Bill of Rights consists of the UNDHR and the twin covenants (ICCPR and ICESCR): UN General Assembly Resolution 217 (III), 10 December 2014.

[14] Milanovic (n 4) 56.

[15] HRC, 'General Comment No. 31: The Nature of the General Legal Obligation Imposed on States Parties to the Covenant' (26 May 2004) [6], UN Doc CCPR/C/21/Rev.1/Add.13. See generally Shelton and Gold (n 3) 562.

[16] In the sense of Article 4 of the Articles on State Responsibility.

it does not impose too great a burden on State parties,[17] i.e. obligations beyond what could ever be achieved in light of the insufficient proximity between the rights violation and State A. For example, where a private party in another territory, with no relevant link to State A, violates the rights of an individual also located in another territory, State A would be *unable* to anticipate the actions of the private party, introduce preemptory measures, or indeed carry out proper investigations once the violation had taken place. In line with the shared object and purpose of human rights treaties, it is only in these circumstances, where State A is completely *unable* to fulfil *any* of the various obligations under the treaty, in which jurisdiction should not be found.[18] Any model of extraterritorial application of human rights treaties must therefore be viewed in light of this principle of effectiveness.

Judge Angelika Nußberger notes that international courts and treaty bodies must 'be aware of [their] own capacit[ies]'.[19] As these institutions have seen a greatly increased workload over the last half-century, Judge Nußberger suggests that a restrictive interpretation of jurisdiction may be necessary to ensure the effectiveness of such institutions in maintaining international justice. However, the question of who holds human

---

[17] This can be contrasted against the notion of effectiveness as applied elsewhere in international law, for example, by the CJEU through the notion of *effet utile*, which requires that international legal norms are observed in order to protect the efficacy of the legal order: '*effet utile*' *Encyclopaedic Dictionary of International Law* (3rd edn, 2009).

[18] Where State A's ability to protect against rights violations is minimal, it would still be under an obligation to cooperate, which constitutes a legal duty under Articles 55 and 56 UN Charter (see discussion at Part V.1).

[19] Angelika Nußberger, 'The Concept of 'Jurisdiction' in the Jurisprudence of the European Court of Human Rights' (2012) 65 Current Legal Problems 241, 253.

rights under a treaty is fundamental and the application of thresholds which, to a large extent, discount rights violations abroad undermines the integrity and coherence of international human rights law. Selectivity is necessary, but this must be done on a reasoned basis, by focusing on exemplary cases and ensuring a robust admissibility stage of proceedings.[20]

4. The failings of the spatial and personal models of extraterritorial application

In practice, international courts and treaty bodies have interpreted the notion of 'jurisdiction' under human rights treaties restrictively.[21] Jurisdiction is largely established using the spatial and personal models of extraterritorial application; these are engaged where the area in which the affected individual is situated or where the individual is themselves under State A's effective control.[22] These models, though, do not extend to all situations where State A is *able* to secure the rights of relevant individuals and consequently these interpretations of jurisdiction do not align with the principles of universality and effectiveness.

A. The spatial model

Article 1 of the European Convention on Human Rights ('ECHR') obliges States parties to 'secure to everyone within their jurisdiction the rights and freedoms defined in […] this

[20] As Judge Nußberger acknowledges: ibid 253.

[21] 'Jurisdiction' has been consistently interpreted across the ECHR, ICCPR, ACHR, ICESCR, CEDAW, and others as pertaining to control.

[22] 'Control' is rarely defined by international courts and treaty bodies and the threshold varies between the two models of extraterritorial application, as is discussed below in Parts IV.1 and IV.2.

Convention'. The standard for the spatial model of extraterritorial application was set by the European Court of Human Rights ('ECtHR') in *Loizidou v Turkey*,[23] concerning the Turkish occupation of Northern Cyprus, and requires that State A 'exercise[] effective control of [the relevant] area outside its national territory'.[24] State A has effective control where it exercises executive powers over the relevant area,[25] such that it generally directs the local authorities.[26]

The spatial model is restrictive as it requires control over an area in order for human rights obligations to be engaged. Invasion constitutes only one example of the numerous types of extraterritorial activity – military, political, economic – which are carried out by State parties abroad and which affect human rights. There are many other types of activity, though, which do not bring the relevant territory under State parties' effective control. For example, where an Iraqi civilian is taken into custody and killed by UK forces operating in Iraq,[27] it cannot sensibly be argued that the precise area in which the human rights violations took place was under the UK's effective control. The UK, though, is *able* to secure the rights of the relevant individuals as it exerts complete control over them. The spatial model therefore excludes extraterritorial situations where State A remains able to

---

[23] (Preliminary Objections) App No 40/1993/435/514 (ECtHR, 23 March 1995) [62].
[24] ibid [52].
[25] *Al-Skeini v United Kingdom*, App No 55721/07 (ECtHR, 7 July 2011) [131].
[26] *Issa v Turkey*, App No 31821/96 (ECtHR, 16 November 2004) [70].
[27] See *Al-Skeini v United Kingdom* (n 25). Only Mousa, the son of the sixth claimant, was detained under British custody, whereas the other claimants' relatives were killed on the streets of Basra while British troops were on patrol: Karen da Costa, *The Extraterritorial Application of Selected Human Rights Treaties* (Martinus Nijhoff 2013) 223.

secure the relevant right because it has control over people rather than territory.

B. The personal model

International courts and treaty bodies have avoided rigid application of the spatial model by also focusing on the direct relationship between State A and the affected individual.[28] In *Öcalan v Turkey*[29] the applicant claimed that Turkish authorities operating overseas had abducted him. The ECtHR looked beyond notions of *territorial* control, which limited the normative efficacy of the spatial model, and instead applied the *personal* model of jurisdiction.[30] The personal model arises where the individual whose right is damaged is themselves under State A's 'authority and control',[31] for example, where they have been captured by State A's agents[32] or where State A enjoys sovereign rights over them because they are its national.[33] Returning to the example of the Iraqi civilian, in *Al Skeini v United Kingdom*[34] the ECtHR found that, independent of control over the relevant area in Iraq, the detainee had been 'taken into the custody of [the UK]'[35] and was consequently under the UK's complete control and, therefore, under its jurisdiction.[36] The personal model operates alongside the spatial model and had been used in earlier

[28] *López Burgos v Uruguay*, Com No 52/1979 (HRC, 29 July 1981) [12.2], UN Doc CCPR/C/13/D/52/1979.
[29] (Admissibility Decision) App No 46221/99 (ECtHR, 14 December 2000).
[30] ibid [93].
[31] *Issa v Turkey* (n 26) [71]; da Costa (n 27) 162.
[32] For example, *López Burgos v* Uruguay (n 28).
[33] For example, *Montero v Uruguay* (n 5).
[34] *Al-Skeini v United Kingdom* (n 25).
[35] ibid [136].
[36] ibid [149] – [150].

decisions of the European Commission of Human Rights[37] as well as in communications of the Human Rights Committee ('HRC'),[38] the treaty body to the International Covenant on Civil and Political Rights ('ICCPR').

However, the personal model excludes extraterritorial situations where State A has insufficient proximity to the individual to control them but is nonetheless able to secure their rights. For example, global surveillance and modern armed conflict are often controlled centrally within State A's territory with effects abroad. The UK's surveillance regime involves the collection of data belonging to those situated outside the UK's territory and has impacts on their right to privacy (Article 8 ECHR);[39] they are not, however, controlled by the UK. Similarly, the US engages in drone attacks abroad which endanger civilian lives contrary to the right to life (Article 6 ICCPR);[40] however, the targets are not controlled by the State and thus the personal model is not engaged.[41] The extraterritorial element in such

[37] For example, *X v United Kingdom*, (Admissibility Decision) App No 7547/76 (ECommHR, 15 December 1977) [74].
[38] For example, in *López Burgos v Uruguay* (n 28) [12.2].
[39] See *Big Brother Watch and Others v United Kingdom*, App Nos 58170/13, 62322/14 and 24960/15 (ECtHR, 13 September 2018).
[40] See HRC, 'Concluding Observation on the United States of American' (23 April 2014) [9], UN Doc CCPR/C/USA/CO/4.
[41] Interestingly, in a recent case the German Higher Administrative Court for North Rhine-Westphalia was asked whether such drone strikes would offend the German Basic Law as the US Air Base in Ramstein, Germany plays an essential role in these operations. The German Court emphasised that there is no limitation to the extraterritorial effect of the right to life under the Basic Law: OVG NRW, Urteil vom 19.3.2019 – 4 A 1361/15 <https://www.justiz.nrw.de/JM/Presse/presse_weitere/PresseOVG/19_03_2019_1/190319a_Anlage.pdf> accessed 25 April 2019: 'Die Schutzpflicht des Staates aus Art. 2 Abs. 2 GG besteht bei Gefahren für

situations, merely the location of the affected individual, in no way impacts State A's *ability* to prevent the rights violations. The UK is able to respect Article 8 ECHR by prohibiting the storage of certain data and the US can effectively secure the right to life by adopting suitable measures to ensure that the prophylactic obligations under Article 6 ICCPR are met. Further, the personal model incentivises State A to commit human rights violations against an individual before performing proper checks, as such checks might require bringing individuals under State A's control and, therefore, jurisdiction.[42] Thus, the personal model also fails to balance effectiveness against the universality of rights as it derogates from the latter in situations where the former is not impaired.

5. A causal model of extraterritorial application

There are two problems inherent in the above models of extraterritorial application of human rights treaties. Firstly,

---

das Grundrecht auf Leben auch bei Auslandssachverhalten, sofern ein hinreichend enger Bezug zum deutschen Staat besteht.' ('The State's duty to protect under Art. 2 II BL also arises in cases of threats to the basic right to life abroad, so long as there is a sufficiently close relationship with the German State.'). See Leander Beinlich, 'Germany and its Involvement in the US Drone Programme before German Administrative Courts' (EJIL: *Talk!*, 8 April 2019) <https://www.ejiltalk.org/germany-and-its-involvement-in-the-us-drone-programme-before-german-administrative-courts/> accessed 25 April 2019.

[42] Noam Lubell, *Extraterritorial Use of Force Against Non-State Actors* (OUP 2010) 224; *Al-Skeini v United Kingdom*, (Concurring Opinion of Judge Bonello) App No 55721/07 (ECtHR, 7 July 2011) [15].

extraterritorial application is necessary to ensure that State A cannot perpetrate violations of treaty rights in other territories which it could not perpetrate in its own.[43] However, the existence of tests for jurisdiction, in addition to what an applicant must already prove, means that extraterritorial application is treated as an 'exception[]'[44] to be justified. This creates a *prima facie* presumption that, where the affected individual is situated outside of State A, they do not hold the relevant rights.[45] This has given rise to conflicting case law; courts have sought equitable outcomes in specific cases by formulating contradictory tests, resulting in a total loss of coherence.[46]

Secondly, although the spatial and personal models are successfully applied in situations where State A is able to secure rights, they do not constitute legal tests for determining such ability. They consequently exclude situations where State A is able to secure relevant rights and thus do not align with the principles of universality and effectiveness.[47] State A can, of course, guarantee the rights of those situated in an area it has annexed, or who have been captured by its agents. However, the decisive factor indicating whether State A is *able* to prevent a rights violation is not whether it has general control over a person, but

---

[43] As acknowledged by the HRC in *López Burgos v Uruguay* (n 28) [12.3] and the ECtHR in *Isaak v Turkey*, (Admissibility Decision) App No 44587/98 (ECtHR, 28 September 2006) 20.

[44] The 'exceptional' nature of extraterritorial application of the ECHR is emphasised often by the ECtHR, for example, in: *Banković v Belgium* (n 10) [67]; *Manitaras v Turkey*, (Admissibility Decision) App No 54591/00 (ECtHR, 3 June 2008) [26]; and *Al-Skeini v United Kingdom* (n 25) [131].

[45] Theodor Meron, 'Extraterritoriality of Human Rights Treaties' (1995) 89 American Journal of International Law 78, 81.

[46] *Al-Skeini v United Kingdom*, (Concurring Opinion of Judge Bonello) (n 42) [5].

[47] See Part III.

whether it controls the relevant act or omission which *causes* the violation.[48] It is in such circumstances in which State A has the power to prevent the human rights infringement. The spatial and personal models must therefore be replaced with a legal test for determining State A's *ability* to secure rights: a causal model which determines whether State A has control over an act or omission which causes an extraterritorial rights violation.

A. Introducing the causal model through principles of interpretation

A causal model of extraterritorial application was expressly rejected by the ECtHR in *Banković*.[49] This case concerned airstrikes by NATO forces on the former Federal Republic of Yugoslavia. A missile had hit a radio and television broadcasting centre, killing sixteen people and injuring a further sixteen. The ECtHR rejected what it considered a 'cause-and-effect' notion of jurisdiction,[50] where an individual who was causally affected by an act or omission emanating from a State party is brought under its jurisdiction. It considered this model to be contrary to the ordinary meaning of the term 'within their jurisdiction' in Article 1 ECHR as it rendered those words 'superfluous and devoid of any purpose'.[51] It further noted that this model conflates jurisdiction with the question of whether the applicant is a victim of a rights violation.[52]

---

[48] *Al-Skeini v United Kingdom*, (Concurring Opinion of Judge Bonello) (n 42) [11].
[49] *Banković v Belgium* (n 10) [75].
[50] ibid [75].
[51] ibid [75].
[52] ibid [75].

However, the ECtHR failed to distinguish between different notions of 'jurisdiction' in international law and their separate ordinary meanings in the sense of Article 31(1) VCLT. Jurisdiction under general international law refers to the permission to exercise legal authority in a given situation.[53] However, as Dr Ralph Wilde notes, the suggestion that human rights are owed only to those over whom States *lawfully* exercise power would be 'perverse'.[54] State party jurisdiction under human rights treaties is therefore a separate doctrine which refers to *de facto* rather than *de jure* jurisdiction; we look for control factually resembling the sovereign powers which a State enjoys over its territory and nationals. This was confirmed by the HRC, which has found that jurisdiction 'applies to those within the power or effective control of the forces of a State party acting outside its territory, regardless of the circumstances in which such power or effective control was obtained'.[55]

State party jurisdiction therefore need not be tied to notions of territory and arises where there exists factual control over the relevant rights violation.[56] This interpretation is supported by the drafting of jurisdiction clauses under human rights treaties, *inter alia*, Article 1 ECHR, Article 2(1) ICCPR and Article 1(1) American Convention on Human Rights ('ACHR'). These require each State party to secure the rights contained within the respective treaty to all those under its jurisdiction; such

[53] Cedric Ryngaert, *Jurisdiction in International Law* (OUP 2008) 5–10.
[54] Ralph Wilde, 'Triggering State Obligations Extraterritoriality: The Spatial Test in Certain Human Rights Treaties' (2007) 40 Israel Law Review 503, 513.
[55] General Comment 31 (n 15) [10]. See also: Wilde, ibid 513; Milanovic (n 4) 39. Indeed, this is apparent from the distinction between 'territory' and 'jurisdiction' under Article 2 (1) ICCPR.
[56] *Al-Skeini v United Kingdom*, Concurring Opinion of Judge Bonello (n 42) [12].

jurisdiction is therefore prefaced on the State party's *ability* to secure the rights.

In any event, there is no hierarchy among the tools of interpretation in Article 31 VCLT, and it would be justifiable to favour an interpretation of 'jurisdiction', in line with the treaty's object and purpose, which relies on an alternative ordinary meaning of the term.[57] Such interpretations have already been adopted in respect of numerous jurisdiction clauses. For example, the jurisdiction clause under Article 2(1) ICCPR requires each State party to guarantee the rights of those under its 'territory and jurisdiction'. This has been applied by the International Court of Justice[58] and the HRC[59] as meaning territory *or* jurisdiction. Similarly, Article 2(1) of the Convention against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment refers to 'acts of torture in any territory under [a State party's] jurisdiction' but is applied extraterritorially under the personal model.[60]

Such an interpretation is further supported by the duty to 'take joint and separate action in cooperation' under Articles 55 and 56 UN Charter,[61] a relevant rule of international law which constitutes an interpretative tool under Article 31(3)(c) VCLT. In many situations, it is *only* the State party in the territory in which

---

[57] Richard Gardiner, *Treaty Interpretation* (2nd end, OUP 2013) 290; ILC Report on the work of its Sixty-fifth session (2013) General Assembly Official Records Sixty-eighth Session, Supplement No 10, Chapter 4, 21 [4], UN Doc A/68/10.

[58] *The Wall* (n 8) [111].

[59] General Comment 31 (n 15) [10].

[60] CAT, 'General Comment No. 2: Implementation of Article 2 by States parties' (24 January 2018) [16], UN Doc CAT/C/GC/2.

[61] For a discussion on the legal basis of the obligation to cooperate in the UN Charter, see Tahmina Karimova, *Human Rights and Development in International Law* (Routledge 2016) 126–132.

the act or omission occurs which is able to prevent the extraterritorial harm. Outright rejection of the causal model therefore creates a vacuum in rights protection.[62] This runs contrary to the duty to cooperate, a fundamental principle of international law,[63] as it removes the ability of the affected State party, State B, to prevent the violation of rights in its territory. A causal model thus ensures the greatest level of cooperation to fulfil obligations under human rights treaties.

In respect of the ECtHR's second objection – that the causal model conflates jurisdiction with the requirement that the applicant is a victim of a rights violation – the test applied to determine whether or not the applicant is a victim of a rights violation is a preliminary survey which applies only a basic notion of causation.[64] The test applied under the causal model is considerably more rigorous.[65]

B. Case law applying a causal model

[62] It creates territorial non-application of the rights under human rights treaties: Kjetil M Larsen, '"Territorial Non-Application" of the European Convention Human Rights' (2009) 78 Nordic Journal of International Law 73.

[63] The duty to cooperate 'revitalizes' the principle of sovereign equality, which holds a central role in the international legal order: Georges Abi-Saab, 'Wither the International Community?' (1998) 9 European Journal of International Law 248, 261. Cooperation is essential in the field of international human rights as without it the realisation of rights would 'remain an unfulfilled aspiration': CESCR, 'General Comment No. 3: The Nature of States Parties' Obligations (Art. 2, Para. 1, of the Covenant)' (14 December 1990) [14], UN Doc E/1991/23.

[64] For example, in ECtHR jurisprudence an applicant must show that they were 'directly affected': *Tănase v Moldova*, App No 7/08 (ECtHR, 27 April 2010) [96].

[65] See Part V.3.

The ECtHR encountered clear normative difficulties following its rejection of the causal model in *Banković*,[66] as seen in *Andreou v Turkey*.[67] This case concerned the cross-border shooting of demonstrators situated on Greek-Cypriot territory by Turkish forces situated on Turkish-Cypriot territory. Turkey sought to deny responsibility by claiming that the affected individuals did not fall under its jurisdiction as they were neither situated in an area under its control (the spatial model) nor were they themselves directly under the control of Turkish authorities (the personal model). This analysis highlights the two difficulties with rejection of a causal model highlighted above. Firstly, it allows for States parties to deny jurisdiction because of damage where the extraterritorial element, the location of the affected individuals, has no impact on State A's *ability* to prevent the rights violation. The mere incidence of a cross-border, rather than intrastate, shooting does not limit Turkey's ability to prevent the shots being fired as this still took place on Turkish-controlled territory. Secondly, in this situation it was *only* Turkey which was able to prevent the harm; the Turkish forces operating on Turkish-Cypriot territory could be controlled only by Turkey. This denies Greece the ability to guarantee the rights of the applicant, contrary to the duty to cooperate.

In light of these difficulties, the ECtHR departed, although only implicitly, from its position in *Banković* by finding that:

> *even though the applicant had sustained her injuries in territory over which Turkey exercised no control*, the opening of fire on the crowd from close range, which was the *direct and immediate cause* of those injuries, had been such that the applicant

[66] *Banković v Belgium* (n 10).
[67] App No 16094/90 (ECtHR, 27 October 2009).

> should be regarded as 'within [the] jurisdiction' of Turkey within the meaning of Article 1 of the Convention.[68]

A causal model has been applied in several other admissibility decisions[69] and in the judgement of *Stephens v Malta (No. 1)*,[70] concerning the detention of an individual by Spanish authorities under the instructions of Malta. In finding jurisdiction, the ECtHR here emphasised that:

> it cannot be overlooked that the applicant's deprivation of liberty had its *sole origin* in the *measures taken exclusively* by the Maltese authorities […] By *setting in motion* a request for the applicant's detention pending extradition, the responsibility lay with Malta […].[71]

The Strasbourg Court did not assert that the actions of the Spanish authorities are directly attributable to Malta; these actions were instead framed as the causal consequence of Maltese actions, attributable to the respondent.

An explicit application of the causal model can be found in Advisory Opinion 23/17 of the Inter-American Court of Human Rights ('IACtHR').[72] It was invited to consider situations

---

[68] ibid [25]. Discussed by Michael Duttwiler, 'Authority, Control and Jurisdiction in the Extraterritorial Application of the ECHR' (2017) 30 Netherlands Quarterly on Human Rights 137, 148 (emphasis added).

[69] In two further admissibility decisions (*Pad and Others v Turkey*, (Admissibility Decision) App No 60167/00 (ECtHR, 28 June 2007) [54] and *Kovačić and Others v Slovenia*, (Admissibility Decision) App No 45653/99 (ECtHR, 3 June 2008) 52) the ECtHR has implicitly found clear causal harm to suffice for establishing jurisdiction.

[70] App No 11956/07 (ECtHR, 21 April 2009).

[71] ibid [51]–[52] (emphasis added).

[72] *Environment and Human Rights*, Advisory Opinion OC-23/17, IACtHR Series A No 23 (2018).

of transboundary environmental harm[73] with simultaneous *human rights* harm.[74] The IACtHR took an expansive approach to interpreting 'jurisdiction' under Article 1(1) ACHR, finding that:

> Under the American Convention, when transboundary harm which affects Convention rights occurs, it is understood that the persons whose rights have been injured are under the jurisdiction of the State of origin [of the transboundary harm] if a relationship of causation exists between the act that occurred in its territory and the infringement of the human rights of persons outside its territory.[75]

This is a welcome step towards an approach which achieves a principled interpretation of jurisdiction in line with the notions of universality and effectiveness, particularly as it views positive obligations as applicable under a causal model. However, limiting the model to cases of transboundary harm is as arbitrary as finding the requirement of jurisdiction only to be met in

---

[73] This occurs when State A breaches its obligation to prevent activities taking place on its territory which cause significant environmental damage to that of State B: Article 3 Draft articles on Prevention of Transboundary Harm from Hazardous Activities; Philippe Sands, *Principles of International Environmental Law* (2nd edn, CUP 2012) 317f.

[74] Such a situation was the subject of *Aerial Herbicide Spraying*. Colombia was engaged in toxic herbicide spraying over its territory in order to combat the cultivation of coca, however this had numerous environmental and human rights impacts in Ecuador: see Memorial of Ecuador, 28 April 2009, [9.9], and Counter-Memorial of Colombia, 29 March 2010, [9.10]–[9.15], available at <https://www.icj-cij.org/en/case/138/written-proceedings> accessed 25 April 2019. This dispute was settled: *Aerial Herbicide Spraying (Ecuador v Colombia)*, Order of 13 September 2013, ICJ Reports 2013, 278.

[75] Advisory Opinion OC-23/17 (n 72) [101].

situations of military occupation under the spatial model[76] or passport confiscation under the personal model.[77]

Although these decisions are non-binding[78] and do not amount to outright endorsements of a universally-applicable causal model of jurisdiction, they highlight the normative pull of an interpretation of jurisdiction relying on principles of causation.

C. The relevant causal test

Principles used to determine causation in international law are closely tied with State A's ability to guarantee human rights. The causal model does *not* amount to the universal application of human rights treaties as it includes principles which will exculpate State parties where they are, despite a factual causal link between their acts and omissions and human rights violations, unable to secure the relevant rights. This is evident when looking to the two types of obligations owed under human rights treaties, negative and positive.

I. Negative obligations

Causation is assessed in two stages: first, an international court looks for cause-in-fact; second, it assesses scope of

[76] For example, *Loizidou v Turkey* (n 23).
[77] For example, *Montero v Uruguay* (n 5).
[78] There is no *stare decisis* system in the jurisprudence of international courts (see generally Gilbert Guillaume, 'The Use of Precedent by International Judges and Arbitrators' (2011) 2 Journal of International Dispute Settlement 5). Decisions, though, constitute subsidiary means of interpretation under Article 38 (1) (d) ICJ Statute. Further, international courts will generally not deviate from their previous decisions on the same matter for the sake of legal certainty: *Goodwin v United Kingdom*, App No 28957/95 (ECtHR, 11 July 2002) [74].

responsibility.[79] Cause-in-fact requires State A's act or omission to be a 'but for' cause of the human rights damage.[80] Where there are multiple factual causes including State A's act or omission, State A will only be deemed to have caused the wrongful harm if its contribution was a 'principal cause';[81] this is in line with the principle of effectiveness as, where State A is only an ancillary cause of the harm, they lack the ability to prevent it and therefore jurisdiction would not be established. A further limitation to liability when establishing factual causation is that the burden of proof for establishing a causal link rests with the applicant.[82]

The scope of responsibility stage of the causation analysis further limits the liability of State A for the harm it factually caused. The principle of foreseeability requires that the wrongful harm which eventualises would have been foreseen by a person of normal prudence when the relevant act or omission was carried

---

[79] Markus Kellner and Isabelle C Durant, 'Causation' in Atilla Fenyves *et al* (eds), *Tort Law in the Jurisprudence of the European Court of Human Rights* (De Grutyer 2012) 450; Ilias Plakokefalos, 'Causation in the Law of State Responsibility and the Problem of Overdetermination: In Search of Clarity' (2015) 26 European Journal of International Law 471, 476–478; Dávid Pusztai, 'Causation in the Law of State Responsibility' (PhD thesis, University of Cambridge 2017) 119–129, 150–151.

[80] Kellner and Durant, ibid 457–460; *Case of YATAM v Nicaragua*, IACtHR Series C No 127 (2005) [239]–[245].

[81] *Campbell and Cosans v United Kingdom*, App Nos 7511/76 and 7743/76 (ECtHR, 25 February 1982) [26]; *Khodorkovsky and Lebedev v Russia*, App Nos 11082/06 and 13772/05 (ECtHR, 25 July 2013) [940]; Pusztai (n 79) 189–192.

[82] This is true of any legal submission to be substantiated by fact: see, for example, Caroline E Foster, 'Burden of Proof in International Courts and Tribunals' (2010) 29 Australian Yearbook of International Law 27, 41–42; *Case Concerning the Land and Maritime Boundary Between Cameroon and Nigeria (Cameroon v Nigeria)*, (Preliminary Objections) [1998] ICJ Rep 275, 319.

out.[83] This standard will vary dependent on the information available to State A at the relevant time.[84] In *Lemire v Ukraine*[85] the tribunal emphasised that the victim must show that each intermediary link in the causal chain can be traced back to the initial act or omission.[86] Foreseeability is therefore premised on State A's awareness of the likely harmful effects and therefore its *ability* to avoid rights violations.[87]

Where there is insufficient closeness between State A's act or omission and the rights violation, the principle of remoteness operates to exculpate States.[88] The tribunal in *Micula v Romania*[89] formulated a remoteness test for cases of multiple successive causes where subsequent causes which disrupt the chain of causation may constitute intervening acts which render the resulting damage too remote.[90] The notion of remoteness generally affords international courts discretion to exculpate States on a number of grounds, and has been utilised in the jurisprudence of the ECtHR.[91] Remoteness is closely tied with the

---

[83] Marta Infantino and Eleni Zervogianni, 'Summary and Survey of the Results' in Marta Infantino and Eleni Zervogianni (eds), *Causation in European Tort Law* (CUP 2017) 604.
[84] *Bosnia Genocide* (n 6) [430].
[85] ICSID Case No ARB/06/18, Award of 28 March 2011.
[86] ibid [164]–[167].
[87] Infantino and Zervogianni (n 83) 604.
[88] ibid 606–607.
[89] *Ioan Micula, Viorel Micula, SC European Food SA, SC Starmill SRL and SC Multipack SRL v Romania*, ICSID Case No ARB/05/20, Award of 11 December 2013.
[90] ibid [926].
[91] Gerhard Dannemann, *Schadesnersatz bei Verletzung der Europäischen Menschenrechtskonvention* (Heymann 1994) 146–147; Kellner and Durant (n 79) 449, 467–471. See, for example, *Vereinigung Bildender Künstler v Austria*, App No 68354/01 (ECtHR, 25 January 2007) [44]. The ECtHR also refers to the requirements of a clear, direct or sufficient causal link when it chooses to reject an applicant's claim because, for example, the

principle of effectiveness; it is based on the notion that State A cannot control all subsequent causal contributions to its acts or omissions and thus lacks the *ability* to prevent the harm.

These principles greatly limit State A's liability for extraterritorial human rights harm. For example, in the *Aerial Herbicide Spraying Case*,[92] Colombia was engaged in toxic herbicide spraying over its territory in order to combat the cultivation of coca. This had numerous human rights impacts as the herbicides travelled over the border into Ecuador. Under the causal model, if the affected individuals could show that there is but-for causation, Colombia would be *prima facie* responsible. Colombia would be exculpated, though, where the spraying is only a negligible contribution to the general harmful exposure in Ecuador as its contribution would not be the principal cause. Further, if the aerial spraying has a completely unexpected health effect on the relevant individuals, the damage would be unforeseeable. If the aerial spraying only causes damage to health because it reacts with chemicals being sprayed by the affected individuals themselves, there would be an intervening act rendering the harm too remote.

The causal test has a number of advantages when determining the scope of State A's liability. Firstly, it favours case-

---

wrongful harm falls outside of the protective scope of the obligations violated: *Isaak v Turkey*, App No 44587/98 (ECtHR, 24 June 2008) [137]; Kellner and Durant, 468.

[92] (n 74). Such cases of transboundary environmental and human rights damage are increasing in frequency: Peter Beaumont, 'Rotten eggs: e-waste from Europe poisons Ghana's food chain' *The Guardian* (London, 24 April 2019) <https://www.theguardian.com/global-development/2019/apr/24/rotten-chicken-eggs-e-waste-from-europe-poisons-ghana-food-chain-agbogbloshie-accra> accessed 25 April 2019. These would, *prima facie*, give rise to jurisdiction under the causal model.

by-case assessment; it incorporates discretion by allowing courts to consider the culpability of States parties in each case. In particular, the notion of remoteness gives courts considerable breadth in exculpating State A for extraterritorial rights violations.

Secondly, a causal model would ensure greater legal certainty; in comparison to the inconsistently applied spatial and personal models, the question of whether an individual can be causally harmed by State A in a given instance is initially a question of fact rather than law. This may require a degree of cooperation among parties to a dispute to provide evidence related to the rights violations, however this is unlikely to impede the efficacy of international human rights courts and treaty bodies, which have been willing to find violations in cases where the respondent State withholds information.[93]

Thirdly, a causation-based analysis would avoid situations where extraterritorial harm to human rights is not subject to review as it would ensure that a claim could be brought against any State party to a human rights treaty which causally contributed to the human rights harm to an individual. The causal model thus removes vacuums in the protection of rights under human rights treaties.

II. Positive obligations

The causal model would also balance the principle of universality and the effective protection of treaty rights if applied to positive human rights obligations. Professor Marko Milanovic is wary of imposing an expansive model of extraterritorial application for positive obligations because, in such instances, the obligations on State A are so wide-ranging that it requires some degree of

[93] For example, *Timurtaş v Turkey*, App No 23531/94 (ECtHR, 13 June 2000).

control, over the area in which the affected individual is situated, in order to meet them.[94] His proposed model of extraterritorial application[95] would remove the requirement of jurisdiction when determining violations of negative obligations, as well as for procedural positive obligations which are required to make negative obligations effective,[96] while maintaining the spatial model of jurisdiction in applying substantive positive obligations extraterritorially.[97] Professor Milanovic considers this to strike the best balance between the principle of universality and imposing manageable burdens on States parties as it only deviates from a universal approach when applying obligations to protect rights, which he believes cannot be effectively secured in extraterritorial scenarios.

However, extraterritorial scenarios relate to the location of the *affected individual,* whereas positive due diligence obligations require State action in relation to *third parties*; where the former is located extraterritorially this does not render the latter beyond State A's control. This is apparent in the aerial spraying example, now with a private party engaged in aerial spraying activities in Colombia which cause human rights damage to individuals situated in neighbouring Ecuador.[98] Under Professor Milanovic's model, Colombia would come under no due diligence obligations to prevent the spraying activities as the affected individuals are

---

[94] Milanovic (n 4) 209–212.

[95] ibid 209–222.

[96] For example, the obligation to initiate an effective public investigation into any death where it appears that the negative obligation has been breached: Milanovic (n 4) 216. See Laurens Lavrysen, *Human Rights in a Positive State* (CUP 2016) 50–53.

[97] For example, due diligence obligations which are prophylactic in nature. See Lavrysen ibid.

[98] The same scenario was discussed by Professor Milanovic: Milanovic (n 4) 218.

not in an area under Colombia's effective control. However, the extraterritorial element in no way diminishes Columbia's ability to prevent the rights damage as the private party's activities emanate from Columbia's territory; Columbia remains *able* to take reasonable steps to prevent the rights violation by carrying out investigations, identifying those responsible and imposing appropriate punishment.[99] Such situations were acknowledged by the Committee on Economic, Social and Cultural Rights, the treaty body to the International Covenant on Economic, Social and Cultural Rights, which demands that each State party protect the rights of those situated outside its territory who are harmed by the activities of companies domiciled within its territory.[100]

The application of the causal model to positive obligations would impose burdens that can effectively be met as responsibility, as with negative obligations, would be limited by principles of causation. Jurisdiction in these instances would be limited by factual causation, foreseeability[101] and remoteness.

---

[99] *Velásquez v Honduras*, IACtHR Series C No 4 (1988) [174].

[100] CESCR, 'General Comment 23 on the Right to Just and Favourable Conditions of Work (article 7 ICESCR)' (7 April 2016) [70], UN Doc E/C.12/GC/23. See also the UN Guiding Principles on Business and Human Rights, UN General Assembly Resolution 17/4, 16 June 2011. In a recent case, the UK Supreme Court found that villagers in Zambia, where a UK-domiciled company operated a mine, could bring a group tort claim before English courts against the company in relation to alleged toxic emissions which caused human rights and environmental harm: *Vedanta Resources PLC and another v Lungowe and others* [2019] UKSC 20.

[101] The obligation to conduct due diligence only arises when State A knows or ought to have known of the rights violation: *Osman v United Kingdom*, App No 87/1997/871/1083 (ECtHR, 28 October 1998) [116]. See also *Mastromatteo v Italy*, App No 37703/97 (ECtHR, 24 October 2002) [68], [76]; Danwood M Chirwa, 'The Doctrine of State Responsibility as a Potential Means of Holding Private Actors

Further, the positive obligation of due diligence is one of conduct, rather than result;[102] State A is obligated to take *reasonable* steps to prevent human rights harm,[103] determined in line with State A's knowledge of the relevant violations and other factors relating generally to its capacity to intervene.[104] More generally, positive obligations 'must be interpreted in a way which does not impose an impossible or disproportionate burden on the authorities';[105] therefore positive obligations are already applied in accordance with the principle of effectiveness.

6. Conclusion

At a time when States exercise power in all parts of the world, notions of jurisdiction tied to control over an area or an individual are inadequate as they allow States to be exculpated for human rights violations where they remain able to secure these rights. International courts and treaty bodies have sought to remedy this by expanding the meaning of 'jurisdiction' under human rights treaties, however the existing models of extraterritorial application have been outpaced by the increased complexities of States parties' international operations. The only way to ensure that these operations are governed by international human rights

Accountable for Human Rights' (2004) 5 Melbourne Journal of International Law 1, 16–17.

[102] *Bosnia Genocide* (n 6) [430].

[103] *Velásquez v Honduras* (n 99) [174]; Shelton and Gold (n 3).

[104] As the ICJ found in relation to due diligence obligation to prevent genocide, such obligations depend on many factors including the geographical distance of the State from the relevant events and the strength of its political and other links to the relevant actors: *Bosnia Genocide* (n 6) [430].

[105] *Osman v United Kingdom* (n 101) [116].

law is to deem all individuals causally affected by a State party's act or omission to fall under its jurisdiction. A causal model, applied explicitly and in all cases, would use principles of foreseeability and remoteness and reverse the presumption that rights are not owed extraterritorially in favour of the affected individual; States would only be exculpated where they were unable to effectively secure the rights harmed, thus striking the correct balance between reasonable burdens on States parties and the principle of universality.

Human Rights, Deprivation of Liberty, and the Mental Capacity (Amendment) Bill 2018

Oliver Pateman*

Abstract—This article considers the new Mental Capacity (Amendment) Bill and how it interacts with the Human Rights Act, particularly Articles 5 and 6. These two articles protect the right to liberty and the right to a fair trial. The Bill introduces new legal procedures, which are commonly used in care homes and hospitals, to deprive people of their liberty. This article concludes that, in neglecting to consider the human rights implications, the Bill may fail to stand up to scrutiny in the courts, or be insufficiently resourced to give meaningful human rights protection to cared-for persons.

* Harris Manchester College, Oxford. I am sincerely grateful to the OUULJ editorial team for comments made on previous drafts. All errors remain my own.

1. Introduction

The Mental Capacity (Amendment) Bill 2018 is, at the time of writing, awaiting Royal Assent.[1] This will soon become the Mental Capacity (Amendment) Act 2019, but will be referred to henceforth as "the Bill". The decision to amend the current Mental Capacity Act 2005 ('MCA') has its origins in a post-legislative report commissioned in 2014.[2] The report, and the advice of the Law Commission, led to the proposal by the Government of a series of amendments to the MCA, particularly to replace the statutory Deprivation of Liberty Safeguards ('DoLS'). Assessments complying with these safeguards are required to authorise deprivations of liberty, and are commonly applied to older people with dementia, as well as those with autism and specific learning disabilities. The MCA and DoLS are important because they serve the interests of some of the most vulnerable people in our society, and are in place to ensure that deprivations of liberty for care and treatment are in the best interests of the cared-for person. However, the demand for DoLS applications has increased exponentially, and for this reason, the DoLS is changing, following an administrative overload and a growing consensus that the scheme is broken.

The legislation, brought onto the statute book following the Bill, will interact with the UK's international obligations under the European Convention on Human Rights, brought into English Law through the Human Rights Act 1998. The most relevant article of the Act here is Article 5, the right that secures

---

[1] Mental Capacity (Amendment) HL Bill (2017-19) 171.
[2] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: post-legislative scrutiny*, (HL 2013-14, 139 – I) para 1.

the liberty and security of the person from arbitrary detention.[3] Deprivations of liberty, characterised by constant supervision and an inability to leave, must currently be authorised by DoLS to be 'Article 5-compliant', under the Art 5(1)(e) exception that allows for the detention of people of 'unsound mind'.[4] In order to give effect to these rights for the practical benefit of a cared-for person, the MCA was drafted with Article 5 in mind. It also includes a 'best interests' principle that 'An act done, or decision made, under this Act for or on behalf of a person who lacks capacity must be done, or made, in his best interests'.[5] This was assured, on a practical level, in the DoLS by a mandatory 'best interests' assessment, which ensured that the cared-for person's wishes and feelings were central to the decision being made. A decision that is as close as possible to the one that the cared-for person would have made if they did not lack capacity will satisfy the requirement of Article 5 that detention should be as humane and compassionate as possible, and not arbitrary.[6] However, while aspects of the Bill represent admirable positive changes, the replacement to the DoLS, known as Liberty Protection Safeguards ('LPS'), will deprioritise the wishes of the cared-for person (the standard abbreviation is 'P') by removing the best interests assessment, leading to potentially arbitrary deprivations of liberty, as well as lead to conflicts of interest between the various people responsible for looking after them.[7]

---

[3] Human Rights Act 1998, sch 1, art 5.
[4] *ibid*
[5] Mental Capacity Act 2005, s 1(5)
[6] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: post-legislative scrutiny*, (HL 2013-14, 139 – I) summary.
[7] The Law Society, *Mental Capacity (Amendment) Bill 2018 House of Lords Committee Stage* (The Law Society 2018) 3.

The 2014 post-legislative report into the MCA also concluded that the Court of Protection ('CoP'), the court that arbitrates decisions relating to the DoLS, is over-worked and costs too much.[8] Established in 2007, the CoP deals with decisions regarding financial or welfare matters for those who are unable to so decide at the relevant time, and effectively sees to the implementation of the MCA.[9] The court is responsible for deciding when a person lacks mental capacity, for appointing Deputies and administering Lasting Powers of Attorney, and for ruling when and in what circumstances a person can be deprived of their liberty.[10] However, the CoP is ill-equipped on its current footing to continue to serve as the principal arbiter of mental capacity, raising Article 6 concerns about whether the court is able to hear cases within a 'reasonable' length of time, and whether claimants in the CoP who are not entitled to legal aid are able to access the court and have a fair hearing.[11] Worryingly, it is not clear how the proposed reform will address any of the objections outlined above.

Five years after the recommendations in the post-legislative report, the government is going to amend the MCA. Many of the most vulnerable people in our country will be affected by this legislative change. When proposing the Bill in its Second Reading to Parliament, Lord O'Shaugnessy, the government minister responsible, said that it was his aim that the

[8] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: post-legislative scrutiny*, (HL 2013-14, 139 – I) para 203.
[9] HM Courts and Tribunals Service, 'Court of Protection' (*Gov.uk*, 4 June 2015) <https://www.gov.uk/courts-tribunals/court-of-protection> accessed 27 September 2018.
[10] HM Courts and Tribunals Service, 'Court of Protection' (*Gov.uk*, 4 June 2015) <https://www.gov.uk/courts-tribunals/court-of-protection> accessed 27 September 2018.
[11] Human Rights Act 1998, sch 1, art 6.

system would afford the necessary protections for the 2 million people in our society who have impaired capacity.[12] However, as it stands, the Bill risks quietening the voices of these people by diluting the statutory requirement that any deprivation should be in their best interests, and, without major procedural changes, the CoP may be unable to function. This paper will explore some of the reasons why the old Mental Capacity Act was no longer fit for purpose as well as explain some of the problems with the proposed Bill, particularly with regard to its conflicts with Articles 5 and 6 of the Human Rights Act.[13]

2. Background: Inadequacy of the DoLS

The growth in DoLS applications follows the Supreme Court decision in *Cheshire West and Chester Council v P*.[14] The appeals in this case concerned the criteria for judging whether the living arrangements for mentally incapacitated persons (initialled 'P', 'MIG' and 'MEG' in this case) amounted to a deprivation of liberty. If so, the deprivation should have been authorised by DoLS and subject to (at least) annual, independent checks.[15] When drawing up the DoLS scheme in 2005, the government did not fully consider what arrangements or procedures for controlling behaviour might amount to a deprivation of liberty under Article 5. The resources allocated for reviewing DoLS

[12] HL Deb 16 July 2018, vol 792, col 1061.
[13] Human Rights Act 1998, sch 1.
[14] *Cheshire West and Chester Council v P and another* [2014] UKSC 19, [2014] AC 896.
[15] *ibid*

applications were, therefore, insufficient. As Baroness Hale (then Deputy President of the Supreme Court) remarked in the case:

> 'People who lack the capacity to make (or implement) their own decisions about where to live may justifiably be deprived of their liberty in their own best interests. They may well be a good deal happier and better looked after if they are. But that does not mean that they have not been deprived of their liberty. We should not confuse the question of the quality of the arrangements which have been made with the question of whether these arrangements constitute a deprivation of liberty.'[16]

Baroness Hale held in *Cheshire West* that Article 5 was, self-evidently, just as applicable to people who lack capacity as to those who have it, even though it does permit detention and deprivation of liberty under certain circumstances. She held that the interventions, even when made for what was believed to be P's benefit, did amount to a deprivation of liberty under Article 5, and therefore required a DoLS assessment. She stated that:

> 'It seems to me, what it means to be deprived of liberty must be the same for everyone, whether or not they have physical or mental disabilities. If it would be a deprivation of my liberty to be obliged to live in a particular place, subject to constant monitoring and control, only allowed out with close supervision, and unable to move away without permission even if such an opportunity became available, then it must also be a deprivation of the liberty of a disabled person. The fact that my living arrangements are comfortable, and indeed

[16] *ibid* [34].

> make my life as enjoyable as it could possibly be, should make no difference. A gilded cage is still a cage.'[17]

In their Lordships' view in *Cheshire West*, interventions such as restrictive clothing and the insertion of fingers into P's mouth to prevent him from eating his continence pads, were all adjudged to be deprivations of liberty under Article 5.[18] However, the two key elements, not being allowed to leave and being kept under continuous supervision, were the 'acid test' for determining whether there has been a deprivation of liberty that, without DoLS, would constitute an infringement of Article 5.[19] The commissioners' initial impact assessment commissioned in 2005 clearly did not consider that interventions such as those in *Cheshire West* would constitute a potential breach of Article 5. It suggested that the DoLS would only be used in a relatively small number of extreme cases; 21,000 at most, decreasing to 1,700 authorisations per year by 2015/16.[20] This is a stark contrast to the 195,840 applications that were in fact made in 2015/16 in England alone in the aftermath of the *Cheshire West* decision.[21] Indeed, Baroness Finlay of Llandaff, Chair of the Mental Capacity Forum, remarked in Parliament that there are over 108,000 applicants in backlog waiting to be processed under the DoLS, and that the 'current DoLS system has effectively fallen over'.[22] The need to apply DoLS in more situations than was initially envisaged has led to great expense and delay, which necessitates the development

---

[17] *ibid* [45]-[46].
[18] *ibid* [51].
[19] *Cheshire West and Chester Council v P and another* [2014] UKSC 19, [2014] AC 896 [54].
[20] Law Commission, *Mental Capacity and Deprivation of Liberty* (Law Com No 372, 2017) para 4.13 (9).
[21] *ibid*
[22] HL Deb 5 September 2018, vol 792, col 1823.

of new statutory safeguards for the deprivation of liberty. However, a failure to protect the rights of P under the new scheme will be incompatible with human rights and will condemn the government to repeating the same mistakes. Any new statutory scheme to authorise deprivations of liberty must offer effective as well as efficient protection of human rights, otherwise it risks being undermined by another *Cheshire West*-style case.

3. Introduction of the LPS

Liberty Protection Safeguards ('LPS') have been proposed to replace the DoLS, bringing several practical advantages. They will remove much of the unhelpful red tape surrounding DoLS, which the Law Commission found to be ignored in practice anyway. Through this, the LPS could save local authorities around £200 million per year[23] and remove the need for repeat assessments when moving P between different care settings, such as from their care home to an ambulance, then to a hospital.[24] However, the LPS also dilute the overall ethos of the existing legislation pertaining to deprivation of liberty – which is 'to put the patient at the centre'.[25]

---

[23] Department of Health and Social Care, 'New law introduced to protect vulnerable people in care' (Gov.uk, 3 July 2018) <https://www.gov.uk/government/news/new-law-introduced-to-protect-vulnerable-people-in-care> accessed 27 September 2018.
[24] Department of Health and Social Care, 'New law introduced to protect vulnerable people in care' (*Gov.uk*, 3 July 2018) <https://www.gov.uk/government/news/new-law-introduced-to-protect-vulnerable-people-in-care> accessed 27 September 2018.
[25] Parliamentary Office of Science and Technology, 'note #381 Mental Capacity and Healthcare' (researchbriefings.parliament, 1 June 2011)

A. Inadequacy of LPS: No 'best interests' assessment

The first problem with the LPS is that the Bill does not consider the best interests of P in a way that is consistent with Article 5's protection against arbitrary deprivation of liberty. Article 5 1(e) requires that:

> 1 Everyone has the right to liberty and security of person. No one shall be deprived of his liberty save in the following cases and *in accordance with a procedure prescribed by law*: […]
>
> > (e) the lawful detention of persons for the prevention of the spreading of infectious diseases, of persons of unsound mind, alcoholics or drug addicts or vagrants;[26]

A 'procedure prescribed by law' is therefore required for Article 5 rights to be upheld where there is a deprivation of liberty.[27] The three conditions to authorise a deprivation under paragraph 12 of the proposed LPS procedure are that:

> (a) the cared-for person lacks capacity to consent to the arrangements,
>
> (b) the cared-for person has a mental disorder, and
>
> (c) the arrangements are necessary to prevent harm to the cared-for person and proportionate in relation to the

<https://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-381> accessed 24 September 2018.

[26] Human Rights Act 1998, sch 1 art 5 (emphasis added).

[27] *ibid*

> likelihood and seriousness of harm to the cared-for person.[28]

When examining the conditions of Article 5, and the origins of the MCA, it is clear that there are substantive conditions, aside from the mere existence and deployment of a procedure, required to authorise a deprivation of liberty. The LPS procedure will authorise a deprivation of liberty where it is 'necessary' and 'proportionate'; however it is not clear that these two conditions alone are sufficient to protect against arbitrary deprivations of liberty.

In 2004, the European Court of Human Rights considered the need for a statutory process to authorise deprivations of liberty in *HL v UK*, also referred to as the '*Bournewood Case*'.[29] In *HL*, the disabled patient, who lacked the capacity to consent or object to any medical treatment, was informally admitted to a hospital and treated with a sedative.[30] The doctors found that *HL* had been compliant with the treatment and made no attempt to leave.[31] The question that the European Court had to answer was whether *HL* had been deprived of his liberty, in that he had been detained, but under no statutory procedure. Echoing the minority opinion of Lord Steyn in the preceding domestic litigation in the House of Lords, the Strasbourg Court held that:

> 'As a result of the lack of procedural regulation and limits, the hospital's health care professionals assumed full control of the liberty and treatment of a vulnerable

[28] Mental Capacity (Amendment) HL Bill (2017-19) 171 sch 1 para 12.
[29] *HL v UK* App no 45508/99 (ECtHR, 5 October 2004).
[30] *ibid* [11].
[31] *ibid* [13].

> incapacitated individual solely on the basis of their own clinical assessments completed as and when they considered fit: as Lord Steyn remarked, this left 'effective and unqualified control' in their hands. While the Court does not question the good faith of those professionals or that they acted in what they considered to be the applicant's best interests, the very purpose of procedural safeguards is to protect individuals against any "misjudgments" and professional lapses […] this absence of procedural safeguards fails to protect against arbitrary deprivations of liberty on grounds of necessity.'[32]

The court held that without a legal procedure, 'effective, unqualified control' was left in the hands of the clinician. This did not protect P from misjudgements or professional lapses and was therefore insufficient to satisfy the requirements of Article 5.[33] However, subsequent cases have indicated more substantive conditions for any potential 'legal procedure'. The current DoLS procedure was recently reviewed and endorsed in the case of *RB v UK*. In that case, the European Court found that a legal procedure to deprive people of their liberty must protect against arbitrariness by applying 'fair and proper procedures for deprivation of liberty'.[34] The Supreme Court in *Cheshire West* suggested yet more substantive conditions. At the conclusion of her speech, Baroness Hale suggested that:

---

[32] *ibid* [122]-[124]
[33] *ibid* [122]-[124].
[34] *RB v UK* App no 6406/15 (ECtHR, 12 September 2017) [22].

> 'Because of the extreme vulnerability of people like P, MIG and MEG, I believe that we should err on the side of caution in deciding what constitutes a deprivation of liberty in their case. They *need* a periodic independent check on whether the arrangements made for them are *in their best interests*. Such checks need not be as elaborate as those currently provided for in the Court of Protection or in the deprivation of liberty safeguards (which could in due course be simplified and extended to placements outside hospitals and care homes). Nor should we regard the need for such checks as in any way stigmatising of them or of their carers. Rather, they are a recognition of their equal dignity and status as human beings like the rest of us.'[35]

In Baroness Hale's view, the best interests of a patient must therefore be established as part of a fair and proper legal procedure, as she noted was the case in the DoLS legislation. Following this reasoning, the best interests of P goes with the assurance that the procedure is fair and proper, and this in turn ensures that the deprivation is, so far as possible, not 'arbitrary' under Article 5.

The LPS requires any proposed deprivation, usually issued to prevent harm, to be necessary and proportionate, and it could be argued that this will necessarily be in P's best interests. However, it is still concerning that, according to the LPS regime, no explicit consideration needs to be given to deciding what is in P's best interests. It is also inconsistent with the empowering ethos of *Cheshire West*. If it can be shown that a deprivation of liberty under the new LPS has been authorised contrary to the

[35] *Cheshire West and Chester Council v P and another* [2014] UKSC 19, [2014] AC 896 [57] (emphasis added).

best interests of P as a result of a failure to consider P's wishes and intentions (or indeed by any other means), this may constitute an infringement of Article 5.

B. Inadequacy of LPS: The role of care home managers

The second problem with the proposed legislation is that the cast of eligible characters to carry out the LPS is too large and under-qualified. This may lead to a misapplication of LPS standards and arbitrary detention under Article 5. The Bill effectively moves responsibility for carrying out lack of capacity assessments away from local authorities to the National Health Service ('NHS'), which is reasonable, but also to care home managers, which is less so. This shift was lauded as one of the great advantages of the Bill as proposed by Lord O'Shaughnessy at Second Reading, as it '[enables] a more streamlined and clearly accountable process in which the NHS has a clear role in helping to afford people their rights.'[36] However, though it is fairly common for medical professionals and social workers to carry out assessments of mental capacity, it is not the case that care home managers will be familiar with them, despite the government's insistence to the contrary.[37] The government has also made no provision for care home managers to receive any training in this regard, though Baroness Blackwood recently assured the House of Lords that there would be an impact assessment following Royal Assent.[38] This aspect of the Bill has attracted censure, that there will be insufficient independent scrutiny to ensure the competency of those proposing deprivations of liberty. As Lord Hunt of Kings Heath commented:

---

[36] HL Deb 16 July 2018, vol 792, col 1059.
[37] HL Deb 5 September 2018, vol 792, col 1819.
[38] HL Deb 24 April 2019, vol 797, col 624.

> 'The Court of Protection has a record of rejecting capacity assessments conducted by consultant psychiatrists with years of training in mental health and specifically in relation to the MCA. The Bill permits care home managers to assess capacity in this context. There is no way that will withstand scrutiny by the court, and there are likely to be even more cases in which assessments of incapacity are overturned as care home managers with little or no relevant training are required to carry out what can be a complex task.'[39]

Additionally, as Lord Touhig made clear at the Bill's Second Reading, there is no consistent standard at all in employing care home managers, demonstrable by comparing different job descriptions.[40] One particularly egregious example was an 'advertisement [that] sought candidates with proven home management experience, strong marketing, commercial and business acumen'.[41] Possessing knowledge of Care Quality Commission ('CQC') standards came in third after these commercial requirements. Care home managers currently have one principal responsibility; that they should run their care home efficiently. Efficient care homes are often full care homes. They certainly have many separate considerations apart from assessing P's capacity, such as organising staffing, maintaining the fabric of buildings, and organising employee training, not to mention developing their 'business acumen'. Their disparate responsibilities aside from the welfare of their residents demonstrate that care home managers are certainly not best placed to make or even sign off on assessments regarding the

[39] HL Deb 5 September 2018, vol 792, col 1830.
[40] ibid, col 1823.
[41] HL Deb 5 September 2018, vol 792, col 1823.

deprivation of their residents' liberty, further opening the possibility of arbitrary detention under the meaning of Article 5.

C. Inadequacy of LPS: Limited role of interested third parties in assessing capacity and wishes of P

The third problem with the proposed LPS is the limited role it gives to others who are outside the responsible body, but who are still heavily invested and involved in the care of P. It is a requirement in the LPS schedule that the cared-for person must actually lack capacity to consent to the arrangements.[42] These individuals are often most able to shed light on whether P actually lacks capacity or is just making an unwise decision. In order to fulfil s. 3 of the MCA, the so-called 'functional test', to deprive someone of their liberty, P must be unable to understand and retain information, weigh it in their mind, and communicate their decision.[43] In applying this functional test, the CoP must decide whether P has actually lost capacity. This is a guiding principle of the MCA, that, 'A person is not to be treated as unable to make a decision merely because he makes an unwise decision'.[44] The fine line between decisions that lack capacity and unwise decisions was discussed in *Heart of England NHS Foundation Trust v JB*,[45], where a schizophrenic patient refused treatment on her foot such that it eventually became detached from the rest of her body. In that case, Jackson J remarked:

[42] Mental Capacity (Amendment) HL Bill (2017-19) 171 sch 1 para 12.
[43] Mental Capacity Act 2005, s 3.
[44] ibid, s 1(4).
[45] *Heart of England NHS Foundation Trust v JB (by her litigation friend, the Official Solicitor)* [2014] EWHC 342 (COP), (2014) 137 BMLR 232.

> 'The temptation to base a judgement of a person's capacity upon whether they seem to have made a good or bad decision, and in particular upon whether they have accepted or rejected medical advice, is absolutely to be avoided. That would be to put the cart before the horse or, expressed another way, to allow the tail of welfare to wag the dog of capacity. Any tendency in this direction risks infringing the rights of that group of persons who, though vulnerable, are capable of making their own decisions.' [46]

The Bill acknowledges that the responsible body who will be issuing an authorisation for a deprivation of liberty, or the care home manager who seeks its approval may not have the complete picture about how P uses and weighs information and will therefore not always know if P lacks capacity. This is why it builds in a compulsory consultation stage at s. 20 of the LPS:[47]

> 20 (1) Consultation under this paragraph must be carried out—
>
> > (a) if the arrangements are care home arrangements and authorisation is being determined under paragraph 16, by the care home manager;
> >
> > (b) otherwise, by the responsible body.
>
> (2) The following must be consulted—
>
> > (a) the cared-for person,

[46] *ibid* [7].
[47] Mental Capacity (Amendment) HL Bill (2017-19) 171 sch 1 para 20.

> (b) anyone named by the cared-for person as someone to be consulted about arrangements of the kind in question,
>
> (c) anyone engaged in caring for the cared-for person or interested in the cared-for person's welfare,
>
> (d) any donee of a lasting power of attorney or an enduring power of attorney (within the meaning of Schedule 4) granted by the cared-for person,
>
> (e) any deputy appointed for the cared-for person by the court, and
>
> (f) any appropriate person and any independent mental capacity advocate concerned (see Part 5).
>
> (3) The main purpose of the consultation required by sub-paragraph (2) is to try to ascertain the cared-for person's wishes or feelings in relation to the arrangements.
>
> (4) If it is not practicable or appropriate to consult a particular person falling within sub-paragraph (2) the duty to consult that person does not apply[48]

Under para. 20(2), the responsible body or care home manager must ensure that a whole host of people concerned for P's welfare are consulted before any authorisation.[49] However, this need only be done where it is practicable and appropriate,

[48] ibid.
[49] Mental Capacity (Amendment) HL Bill (2017-19) 171 sch 1 para 20(2).

making any meaningful consultation into the best interests of P entirely contingent on its practicality and appropriateness in the view of the responsible body or care home manager.[50]

However, the evidence regarding the ability to weigh information provided by the close family members of P has had an impact on previous case outcomes. In *King's College Hospital NHS Foundation Trust v C*, the evidence of C's daughters was essential for the court to decide whether C lacked capacity. C refused medical treatment – kidney dialysis – following a suicide attempt with 60 Paracetamol and a bottle of Veuve Cliquot, on the grounds that she did not want to grow old and lose her beauty.[51] MacDonald J noted, 'C is, as all who know her, and C herself appears to agree, a person who seeks to live life entirely, and unapologetically on her own terms; that life revolving largely around her looks, men, material possessions and 'living the high life"[52] The evidence provided by C's daughters allowed MacDonald J to decide that C indeed had capacity to weigh up and process information about her condition, but had weighted it at naught, considering her fear of lifelong treatment and disability not to be evidence of a lack of capacity.

An LPS can be issued for deprivations of liberty intended only to be effective in the short-term, such as admissions to hospital for emergency treatment. However, the same mechanism is also used to authorise a deprivation of liberty valid for up to three years, and can therefore be used to determine the long-term living arrangements of P. Making the consultation contingent on practicality and appropriateness irrespective of the intended

[50] Ibid, para 20(4).
[51] *Kings College Hospital NHS Foundation Trust v C* [2015] EWCOP 80, [2016] COPLR 50 [14].
[52] *Ibid* [8].

length of the deprivation does not capture the difference between moderate and acute situations. In moderate situations, such as long-term living arrangements, it should be essential to consult, but in acute situations (such as emergency treatment in a hospital), a long consultation process with all interested parties may not be in the best interests of P. Furthermore, where extended consultation is impracticable, it is difficult to see how a doctor or social worker tasked with completing an LPS Assessment could possibly comment on Ps' personality, nor their system of values, when their contact with them is limited; maybe only briefly on a hospital ward or in a care home lounge. As Lord Hunt of Kings Heath mentioned in the Lords' debate, this is particularly relevant when considering the recruitment crisis in care homes: there is currently a staff turnover rate of 27% in care homes, and one in five has no registered manager in position, despite this being mandatory.[53]

In order for the Mental Capacity Bill to be consistent with the principles of *King's College Hospital NHS Foundation Trust v C*, and those explained in s.20(3), the consultation process must be conducted so as to elucidate the wishes and feelings of P to ensure that any deprivation of liberty is in their best interests. This is less likely to happen where consultation is limited to whatever is practicable and appropriate in the opinion of the responsible body. Baroness Finlay of Llandaff also explained that she was 'concerned that, if the lasting power of attorney or the court-appointed deputy are not really given appropriate prominence in our process, we could find that the careful choice of a trusted person becomes effectively downgraded in the system when we are trying to consider what is in P's best interests'.[54]

---

[53] HL Deb 5 September 2018, vol 792, col 1820.
[54] Ibid, col 1811.

It is arguable that a duty to consult interested parties will be very onerous for hospitals and care homes to carry out, and so a practicality requirement is needed if the Bill is to deliver on its promise to save money for the taxpayer. However, no great sums need be expended in tracking down erstwhile, long-lost relatives or those who have a peripheral interest in P's care, since these people will, in all likelihood, not be able to make any meaningful contribution to understanding whether or not P lacks capacity. Consultation should take place where a long-term, abiding provision is proposed if it is reasonably possible, not merely where practical. This nuance should be provided for in the Bill. Furthermore, the fact that the Bill does not require court-appointed deputies and attorneys to be consulted when considering long-term living arrangements, for instance, when these seem to be the most obvious people to consult, especially when considering the restriction of rights enjoyed under Article 5, suggests that the issuing of deprivations of liberty would be in contravention of those rights.

4. The Court of Protection

As the 2014 post-legislative report made clear, it is not just the MCA itself that is in need of reform, but the court that arbitrates on it as well. This brings into consideration additional concerns with regard to Article 6 and the right to a fair trial within a reasonable period of time.[55] Many people are unaware of the existence of the CoP, let alone the work it does or the power it has. However, almost everybody knows someone whose life is closely relevant to the law on which it adjudicates. Two of the

[55] Human Rights Act 1998, sch 1, art 6.

major problems associated with the CoP mentioned in the House of Lords Report were the ever-increasing workload of the court, and the costs for litigants.

A. Inadequacy of the Court of Protection: Increased Workload

The workload of the CoP has increased massively, and it is forecast to continue growing into the future. In the 2014 post-legislative report, District Judge Elizabeth Batten stated that the court's workload has increased by 25%, whilst staffing was reduced by 30%.[56] In 2008, the CoP handled fewer than 1000 health and welfare applications; but, in 2016, it handled over 4000 and in excess of 25,000 cases overall. The number of potential users of the CoP is also expected to increase due to the UK's ageing population - in 2016, the government projected that there will be 1.42 million more households with someone over 85 by 2037 – an increase of 161% over 25 years.[57] Similarly, in 2014, the Alzheimer's Society predicted that 1.142 million people would suffer from dementia by 2025[58] – the approximate population of Birmingham. There is already a shortage of resources and an inability to process cases that come to the CoP in a timely or efficient fashion. The current and projected demand for applications for deprivation of liberty with the current CoP may be irreconcilable with the Article 6 right that everyone is entitled to a fair hearing within a reasonable period of time.

---

[56] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: post-legislative scrutiny*, (HL 2013-14, 139 – I) para 205.
[57] Government Office for Science, *Future of an Ageing Population* (Foresight 2016) p 9.
[58] Alzheimer's Society, *Dementia UK Report* (Alzheimer's Society, 2014).

B. Inadequacy of the Court of Protection: High Cost for Litigants

The House of Lords also identified the cost for litigants at the CoP as a serious cause for concern. The Court oversees the appointment and dismissal of all Deputies, and grants and rescinds Lasting Powers of Attorney ('LPA'). On 4th July 2018, the Parliamentary Under-Secretary of State for Justice, Lucy Frazer QC MP, admitted this fact in a statement to the House of Commons and announced that the cost of certain proceedings in the CoP would be reduced.[59] Additionally, applicants for an LPA or Deputyship can apply for a fee exemption or to have the fees reduced. However, if the applicant (in welfare cases), or P (in finance cases) has savings of over £3,000, it is unlikely that the Office of the Public Guardian will be able to help with the fee.[60] The decrease has been insufficient; applying for an LPA to become a deputy can still be an extremely expensive process. For example, the starting cost is £82 to register an LPA with the Office of the Public Guardian, which grants minimal, basic control over either finances or welfare.[61] If P's attorney dies, then the LPA dies with them, and the whole process must begin again.[62] To become a Deputy, there is a £385 application fee for each type of deputyship (finance and welfare), with an additional £500 fee in the event of a hearing.[63] There is an annual fee of up

[59] HC Deb 4 July 2018, cWS.
[60] Office of the Public Guardian, 'Make, register or end a lasting power of attorney' (*Gov.uk*, 17 June 2018) < https://www.gov.uk/power-of-attorney/register> accessed 3 May 2019.
[61] Ibid.
[62] Ibid.
[63] Office of the Public Guardian, 'Deputies: make decisions for someone who lacks capacity' (Gov.uk, 17 June 2018) <https://www.gov.uk/become-deputy/fees> accessed 27 September 2018.

to £320 for supervision (where assets are over £21,000), and a £100 flat assessment fee if you are a new deputy.[64] However, most significantly, for either the LPA or the Deputyship application, there is no guarantee of success after payment of the fee, and no prospect of reimbursement (unless P dies within five days of the CoP receiving the application).[65]

	Finance **or** Welfare	Finance **and** Welfare
Lasting Power of Attorney	£41-82/person	£82-164/person
Deputyship	£0-385/person +£160-320 annual supervision fee +£50-100 to be appointed new deputy +£5 official copy of documents (+£500 for a hearing) = £215-1310/person/in first year	£0-770/person +£160-320 annual supervision fee +£50-100 to be appointed new deputy +£5 official copy of documents +£500 for a hearing = £215-1695/person/in first year

Reducing these costs, or introducing partial refunds (perhaps deducting administration costs) for unsuccessful or ineffective applications, would stop discouraging people from planning ahead and would, therefore, lead to less litigation in the CoP. In the long term, this would also free up the court's time to deal with the most serious matters. The current prospect of

[64] Ibid.
[65] Ibid.

paying over a thousand pounds for a family member to become a court-appointed deputy with no guarantees of success is daunting to say the least, but contested litigation is obviously far more costly. A report by the University of Cardiff stated that the cost of proceedings 'is a major barrier to accessing justice and is likely to have a significant chilling effect on bringing disputes and serious issues before the CoP'.[66] As an indication of the average amount spent in litigation, legal aid certificates cost around £7,288 in deprivation of liberty cases that require the assistance of the CoP.[67] The prohibitive costs to individuals involved with acquiring an LPA or becoming a court-appointed deputy encourage prospective representatives for the relevant person to postpone planning for a loss of capacity for a rainy day, but if the cost could be reduced at the outset, there could be significant savings for the public purse in the long run.

Equally, if a person lacks litigation capacity outright, it is self-evident that they will require support to access the legal system and be furnished with a suitable advocate. However, the House of Lords noted in their 2014 report particular concerns that individuals have been unable to have their cases heard in the CoP for lack of legal representation, even by the Official Solicitor.[68] This is because the Legal Aid, Sentencing and Punishment of Offenders Act 2012 excluded non-means tested eligibility for legal aid to contest cases where a deprivation of liberty has been ordered by the CoP, or cases where a deprivation

[66] Lucy Series, Phil Fennell, Julie Doughty and Adam Mercer, 'Welfare Cases in the Court of Protection: A Statistical Overview' (Cardiff Law School research project, Cardiff, September 2017) p 4.
[67] Ibid.
[68] Select Committee on the Mental Capacity Act 2005, *Mental Capacity Act 2005: post-legislative scrutiny*, (HL 2013-14, 139 – I) para 71.

is alleged, but disputed.[69] As a result, the Official Solicitor has had to turn away cases for want of funding. The House of Lords remarked that, 'a group of solicitors and barristers reported that this had resulted in cases where no court proceedings could take place, because there was no-one willing or able to act as litigation friend, and they argued for the Official Solicitor's office to be resourced so that he is genuinely a litigation friend of last resort who can act regardless of P's resources, as he does in medical treatment cases'.[70] Granting legal aid as a matter of course, or at least prioritising the provision of legal aid in loss of capacity claims, would limit the noted chilling effect and would ensure that cases heard at the CoP are not prohibitively expensive for litigants.[71]

It is true that legal aid budgets have been cut significantly and are already under extreme pressure. Allocating an increased proportion of scarce resources to deprivation of liberty cases may cause strain elsewhere in the system. However, if somewhere in the region of £200 million will be saved in replacing the DoLS system with the LPS, perhaps some of this should be pledged to give meaningful human rights protection to those who lack capacity.[72] In this way, people who lack litigation capacity would not fall through the cracks, and deprivation of liberty cases can be heard with all parties fully prepared and similarly briefed. Both of these beneficial consequences would improve the security of Article 6 rights. It is obvious that a person who is alleged to lack

[69] Ibid, para 247.
[70] Ibid, para 248.
[71] Ibid, para 72.
[72] Department of Health and Social Care, 'New law introduced to protect vulnerable people in care' (*Gov.uk*, 3 July 2018) <https://www.gov.uk/government/news/new-law-introduced-to-protect-vulnerable-people-in-care> accessed 27 September 2018.

capacity even simply to look after him or herself, let alone defend him or herself should *prima facie* be entitled to decent legal representation.

5. Conclusion

Overall, it is clear that the Mental Capacity (Amendment) Bill is open to a number of different challenges on the basis of the UK's obligations under the Human Rights Act. This is particularly the case when considering the LPS alongside Article 5 and Article 8. The new LPS are in danger of falling into the same trap as DoLS; the removal of the best interests assessment and the deprioritisation of P's wishes in particular, as well as the weakening of the requirements for consultation before authorising the deprivation, show that the statute fails to protect these human rights adequately. When we consider the work of the Office of the Public Guardian and the CoP, as well as the costs involved for litigants, there is also cause for serious concern with regards to Article 6. This is especially true when faced with the prospect that lacking litigation capacity is not considered to be a 'special case' in an application for legal aid when the CoP asks that they be represented by the Official Solicitor.

Ultimately, restricting the liberty of citizens is among one of the most sensitively balanced exercises that the law can undertake. Because of this, the laws that authorise deprivations of liberty should be nuanced and precise, and the consequences of getting the balance wrong can have significant consequences for human rights. As it stands, the Mental Capacity (Amendment) Bill is likely to receive Royal Assent some time very soon. However, the Bill as it is currently does not provide adequate protection. By

not considering the human rights implications from the very outset and taking them as the starting point for drafting a new Bill, and by giving too much weight to questions of practicality, it may well be that the new LPS will also be put to the test in the courts, just as the DoLS was in *Cheshire West*. Any system for the deprivation of liberty must, of course, work in practice. However, restoring the best interests assessment for P before any authorisation of a deprivation of liberty, and reflecting the nuanced situations of where consultation ought to be required would be a great improvement towards protecting P's human rights. The work of the Court of Protection must also be improved, and the government should make sure that legal aid is provided where it is needed. The cost of appointing deputies and LPAs through the Office of the Public Guardian ought to be reduced such that the interests of P are not put off but can be resolved non-contentiously and without unreasonable expense while P retains capacity. This would go some way towards ensuring that the work of the Court of Protection is reserved for the most finely-balanced and difficult cases, freeing up resources for those in greatest need. With these changes, we might get some way closer to a fair CoP.

Recognizing Human Irrationality: Consumer Behaviour and Assessing Commercial Strategies in Digital Markets under Art. 102 TFEU

Tatiana Podstolna*

Abstract—Maintaining effective competition in digital markets is one of the most serious challenges facing EU competition law. It is crucial for the law to strike the right balance between protecting consumer welfare and encouraging innovation. This article evaluates the changes to the application of Art 102 TFEU to instances of tying and bundling, and refusals to supply, in digital markets affected by network effects. It will argue that the CJEU's and the Commission's detailed focus on the actual impact of putative infringements on the irrational consumer has resulted in decisions appropriately protective of the competitive process.

1. Introduction

Analysis of consumer [1] behaviour is crucial at every stage of

* Worcester College, Oxford. I am grateful to the OUULJ editorial team for comments made on earlier drafts. All errors remain my own.

[1] In accordance with the standard use of the term by the European Commission in the context of competition law, the concept of 'consumers' will for the purposes of this article encompass all direct or indirect users of the products affected by the conduct, including intermediate producers that use the products as an input, as well as

assessing the impact that an undertaking's [2] unilateral conduct has on competition within the internal market. Pursuant to Article 102 TFEU (formerly Article 82 EC), any abuse by an undertaking of a dominant position within the internal market shall be prohibited as incompatible with the Treaties insofar as it may affect trade between Member States. Consumer behaviour is necessary for defining the relevant market from the point of view of demand substitutability [3] which, in turn, helps to establish whether the undertaking enjoys a dominant position in said market. [4] Consumer welfare is taken into account during the evaluation of the object or effect of the undertaking's practices. [5] Considerations of efficiencies favourable to consumers may serve as an objective justification for conduct which would have otherwise been abusive. [6]

---

distributors and final consumers both of the immediate product and of products provided by intermediate producers. Where intermediate users are actual or potential competitors of the dominant undertaking, the assessment focuses on the effects of the conduct on users further downstream. C.f. Commission Communication, '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 45/02.

[2] The concept of an undertaking encompasses every entity engaged in an economic activity, regardless of the legal status of the entity and the way in which it is financed. C.f. Joined Cases C-159/91 and C-160/91 *Poucet et Pistre* [1993] ECR I-637, para 17.

[3] Commission Communication 'Commission Notice on the definition of relevant market or the purposes of Community competition law' [1997] OJ C 372/03, paras 15-19 explain that demand substitution entails a determination of the range of products which are viewed by the consumers as substitutes – i.e. products to which she would switch in case the price of the original product underwent a small but permanent increase.

[4] Commission Communication (n 1) para 10.

[5] ibid para 5.

[6] ibid para 30.

The extensive analyses of consumer behaviour conducted by the European Commission during its investigations of Microsoft and, more recently, Google, have attracted a lot of attention from the legal profession and the technology industries alike. For example, concerns were raised about the appropriateness of interference with fast-paced, highly innovative markets by competition authorities. [7] Commentators have highlighted that the Commission and – in case of Microsoft – the CJEU, have chosen to adopt a much more interventionist approach than competition authorities in the US. [8]

Of course, simply pointing out that one approach is more interventionist than the other is not particularly helpful. Much more interesting is the question whether the EU's approach is sound – i.e., whether it corresponds to, and appropriately interacts with, the commercial realities.

The present article will seek to analyse the changes introduced to the interpretation of Article 102 TFEU by the recent digital markets case law. [9] Part 2 of this article comprises an overview of the characteristics of digital markets which have been pertinent to competition law enforcement. It will be suggested that as a result of network effects influencing these markets, they are particularly susceptible to harm by tying and bundling, and refusal to supply. Part 3 and Part 4 will examine the influence of concerns about consumer welfare and the phenomenon of consumer bias on the application of Art 102 TFEU to instances of tying and bundling, and refusal to supply. It will be suggested that the changes are justified insofar as they

[7] David S. Evans, Atilano Jorge Padilla, Michele Polo, 'Tying in Platform Software: Reasons for a Rule-of-Reason Standard in European Competition Law' [2002] World Competition Law and Economics Review 509, 510.
[8] Ioannis Kokkoris, 'The Google Case in the EU: Is There a Case?' (2017) 62 The Antitrust Bulletin 313, 314.
[9] Inter alia Case T-201/04 *Microsoft Corp. v Commission* [2007] 5 CMLR 11; Case AT.39740 *Google Search (Shopping)*; Case AT.40099 *Google Android*.

protect consumers' choice and the competitive process in the long-term.

2. Stimulus for change: particularities of the digital markets

In order to ascertain whether an undertaking has abused its dominant position, it is necessary to assess the characteristics of the market in which the undertaking operates. Thus, application of Art 102 TFEU will necessarily entail assessing the number of competitors, the existence of entry barriers and the distinct commercial practices typical on that market. [10] However, in spite of the Article's broad language, which has allowed for flexible application to different markets, its application to undertakings operating in digital markets has proven to be a challenge. [11]

The Commission has expressed concern over consumers' tendency to disproportionately prioritise easily visible and accessible products over products of higher or equal quality [12] in the context of digital markets. This phenomenon is just one of many manifestations of consumer irrationality – i.e. behaviour on the market which does not maximize their economic welfare. [13] The essential question that needs to be answered is this: when does profiting off consumer's irrationality, which is a typical corporate strategy, cross the line between legal marketing techniques and the realm of anti-competitive practices? Before we

[10] Commission Communication (n 1).
[11] Stefan Holzweber, 'Tying and bundling in the digital era' (2018) 14 ECJ 342, 343.
[12] Case AT.39740 *Google Search (Shopping)* 124.
[13] Nicholas Petit and Norman Neyrinck 'Behavioral Economics and Abuse of Dominance: A Proposed Alternative Reading of the Article 102 TFEU Case-Law' [2010] Global Competition Law Centre Working Paper Series 203.

go on to assess where the Commission and the CJEU have seen it fit to draw this line, it may be helpful to consider the typical consumer-affecting characteristics of digital markets, so as to gain a preliminary understanding of why these markets have become an area of special concern.

Digital technology has quickly become an important staple in the life of an average European. For instance, in 2018 over 82% of the UK population owned a smartphone, [14] 89% used the Google Search engine, 79% had a Facebook account, 79% used YouTube, and 41% used Instagram. [15] Search engines and social media are a good example of digital software which involves and encourages data sharing with third parties. Indeed, the entire point of such software is that the more users contribute data to the software, the more attractive services the software can provide. Yet, the implications of data sharing are lost on many consumers. [16] Many more may feel uneasy about the extent of data sharing but may not understand how to monitor or prevent it. [17]

Furthermore, the diffusion of digital technology and a 'culture' of data sharing has led to the development of more pervasive and aggressive advertisement. While in the past a person could expect their home to provide some reprieve from commercial pressures, today most people carry in their smartphone a portable collection of advertisements everywhere they go.

---

[14] 'Global Mobile Market Report' (*Newzoo*, 2018) <https://newzoo.com/insights/rankings/top-50-countries-by-smartphone-penetration-and-users/> accessed 23 March 2019.
[15] 'Social Media Demographics' *(We Are Flint Consultancy*, 2018) <https://weareflint.co.uk/main-findings-social-media-demographics-uk-usa-2018> accessed 23 March 2019.
[16] 'Attitudes towards data sharing' (*Open Data Institute*, 2017) <https://theodi.org/article/odi-survey-reveals-british-consumer-attitudes-to-sharing-personal-data/> accessed 23 March 2019.
[17] ibid.

The European Union has reacted to these market realities with distinct instruments including an increasingly complex and demanding legislation on data protection [18] and consumer protection. [19] Competition law is therefore just one of the methods of ensuring consumer welfare, yet its role cannot be dismissed. The Director-General of Competition's predominant concern has been the risk that abuse of dominance by the few successful undertakings will paralyse innovation in the long-term, making new entry into the market exceedingly difficult. [20] Digital markets have been flagged up by both academic commentators [21] and the Commission [22] as strongly influenced by network effects, whereby the benefit of the offered product or service increases with the number of users and the amount of data about users. [23] For instance, the value that Facebook can offer to its users increases as more people create accounts.

---

[18] Council Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

[19] Council Directive (EU) 2011/83 of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1994/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance [2011] OJ L304/64.

[20] COM (2018) 482, Report on Competition Policy 2017 from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 4

[21] Maurits Dolman and Thomas Graf, 'Analysis of Tying Under Article 82 EC: The European Commission's Microsoft Decision in Perspective' [2004] World Competition Law and Economics Review 234.

[22] *Google Search* (n 9) 62.

[23] ibid.

Edelman's examination of different practices of Google shows us that: 1) network effects are a powerful incentive for undertakings to engage in more aggressive marketing schemes; and 2) in the case of dominant undertakings, this may sometimes result in them disregarding their 'special responsibility' not to abuse their dominant position by engaging in self-preferencing conduct. [24] While trying to make their products visible is the goal of any undertaking, Dolman and Graf rightly remind us that increasing visibility will be particularly profitable in markets affected by network effects, since any product sold will have an impact beyond the immediate customer. [25] Whereas non-dominant undertakings and/or undertakings operating on only one market will have no choice but to increase the visibility by way of ordinary advertising, undertakings which operate in several markets and are dominant in at least one of them may therefore feel particularly tempted to harm competition by leveraging market power and/or self-preferencing.

The leveraging theory of harm is associated with tying and bundling, [26] whereby an undertaking uses its dominant position in market A to strengthen its position in market B. [27]

---

[24] Benjamin Edelman 'Does Google Leverage Market Power Through Tying and Bundling?' (2015) 11 Journal of Competition Law and Economics 365.

[25] Dolman (n 21).

[26] Jurian Langer, *Tying and Bundling as a Leveraging Concern under EC Competition Law* (Kluwer Law International 2007) 19.

[27] For a period of time it remained deeply controversial whether tying and bundling is at all capable of harming competition: e.g., according to the Chicago school's one-monopoly-profit theorem, a dominant undertaking could never earn more than a single monopolistic price because it can already charge it on the market in which it is dominant, and would therefore never have the incentive to engage in market power leveraging. However, subsequent commentators have noted that the Chicago school vision was too simplistic, limited to short-term effects of leveraging and only operable within the confines of an artificial, strictly defined market – and, crucially, failing to take note of network effects. Should the reader be interested in reading more

Self-preferencing is a dominant feature in cases of refusal to supply, whereby the dominant undertaking promotes its own downstream product at the expense of downstream products marketed by competitors. [28]

The following parts 3 and 4 will examine the changes that innovation and the corresponding changes in consumer behaviour have brought to the application of Art 102 TFEU as regards tying and bundling and refusal to supply, respectively.

3. *Tying and Bundling*

Tying and bundling may be anti-competitive when it falls within the ambit of Art 102(d) TFEU which prohibits the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts. Instances of tying and bundling encompass:

1. pure bundling, where goods A and B are sold together fixedly, and neither can be purchased without the other;
2. mixed bundling, where goods A and B remain available as standalones, but the bundle is offered at a discount;
3. tying, where the offered goods are either just A or the bundle A-B (the tying and the tied product). [29]

A special form of bundling particularly common in digital markets is technical bundling, where two goods are incorporated

---

widely on the subject of economic theories of harm to competition, see e.g. Richard Posner, *Antitrust Law: an Economic Perspective* (University of Chicago Press 1976); Robert H. Bork, *The Antitrust Paradox — A Policy at War with Itself* (Basic Books 1978); Barry Nalebuff, *Bundling, Tying and Portfolio Effects – Report for the UK Department of Trade and Industry* (2003) for contrasting views.

[28] Case C-7/97 *Oscar Bronner* [1998] ECR I-7791, para 32 and 37.

[29] Langer (n 26) 4-5.

into one another in a way which makes it impossible for an ordinary consumer to either dismantle or to assemble in the first place. [30] Somewhat confusingly, the term 'technical bundling' has been used by the CJEU and the Commission regardless of whether the conduct in question consists of pure or mixed bundling, or a tie. The terms address two different issues. The terms 'pure bundling', 'mixed bundling' and 'tie' refer to the combinations of goods *offered for sale*. In comparison, 'technical bundling' refers to the manner in which the goods are *integrated* at the production stage (as opposed to merely packaged together).

A common non-digital instance of 'healthy' technical bundling is the assembly of car parts to make up a single car. In that particular case, the specialized market for car parts will continue to exist, for instance, for the purposes of car repairs. However, the market for cars will exist as a distinct one, since the car itself is valuable for the consumer precisely because it has been assembled and thus offers many functions that separate car parts do not. Indeed, without the existence of a car as a distinct product, the market for car parts would be non-existent – no one would have the incentive to participate in it except with the vision of eventually starting to produce cars.

The precise test for assessing whether technical bundling infringes Art 102 TFEU was articulated by the Commission in *Microsoft:*

1. the tying and tied goods must be separate and;
2. the undertaking concerned must be dominant in the tying product market and;
3. the undertaking concerned must not give customers the choice of obtaining the tying product without the tied product and
4. the tying arrangements must foreclose competition. [31]

[30] ibid.

[31] Microsoft (Case COMP/C-3/37.792) Commission Decision [2007] OJ L32/23*,* recital 794.

This article will further focus on the criteria of 'separateness' (1) and 'customers' choice' (3) insofar as these have proved especially controversial and problematic in the context of digital markets.

A. Separate products

As tying and bundling requires the existence of (at least) two separate products, the question arises how to tell apart two tied products from what has always been or has become a single product. Without separate products, the conduct falls outside Art 102(d) which requires that a *supplementary obligation* be undertaken aside from the one that the consumer primarily wanted to undertake. Defining the products is also the first step towards establishing whether the second *Microsoft* criterion of 'dominance in the tying market' is fulfilled, since dominance can only be assessed by reference to a particular product. Separateness of products can therefore be considered a gateway into abuse by tying and bundling.

This issue of products' differentiation is not particular to digital markets, but is especially pertinent to them, as innovation will often consist of combining several products into a new one to increase the value it brings to the consumer. [32] As Yu notes, technical bundling will often occur during product design and can be considered a product design method. [33] Other than design features, it will often offer consumers entirely new functionalities. For instance, at the beginning of the 2000s, a mobile phone had two main functions: making phone calls and sending text messages. In comparison, today a leading mobile phone manufacturer could not possibly issue a smartphone without, at the very least, Internet access, a camera, a music player, an app-

[32] David A. Heiner, 'Assessing Tying Claims in the Context of Software Integration: A Suggested Framework for Applying the Rule of Reason Analysis' (2005) 72 U. Chi. L. Rev. 123, 127.
[33] Qiang Yu, 'Technically tying applications to a dominant platform in the software market and competition law' (2015) 36 ECLR 160, 162.

store and several pre-installed applications like an Internet browser and a range of social media applications.

Pre-*Microsoft*, the Court focused, inter alia, on the lack of functional interchangeability between the different products sold by the undertaking as the test for separateness. [34] However, as Langer noted, such an approach would mean that complementary products could never constitute an integrated product, since they are by definition intended to be used alongside each other rather than interchangeably. The 'lack of functional interchangeability' test was therefore liable to generate false negatives. [35]

In *Microsoft* the test for distinctness of products was reassessed. The case was concerned with Microsoft's practice of incorporating its Windows Media Player into the Windows Operating System ('OS'). From a technical standpoint it could reasonably be argued that the two products were merged into one. However, the Commission determined that the relevant question to ask is whether there exists consumer demand for the tied product (i.e. the media player) on a stand-alone basis. In addition, it took into account the existence of undertakings specialized in the manufacture of the tied product without the tying product and the dominant company's own conduct in regard to the product(s) in question. [36] Its method was subsequently accepted on appeal by the Court of First Instance ('CFI'). [37]

Dolman and Graf pointed out that the criterion of separate products functions as a proxy for the efficiencies defence: the tie will be efficient from the consumer's perspective and therefore preferred over separate products if it saves them substantial costs. [38] As a consequence of its efficiency, the demand for the tied product as a standalone will disappear and the bundle will 'transition' into a single product. In case of WMP

[34] Case T-30/89 *Hilti* [1991] ECR-II-1439, para 18.
[35] Langer (n 26) 146.
[36] *Microsoft* (n 31) recitals 802-804.
[37] *Microsoft* (n 9) para 916-932.
[38] Dolman (n 21) 230.

and Windows OS, this did not happen: Microsoft was unable to prove that any substantial efficiencies were generated by the tying itself, the CFI having taken a sceptical view on the argument that efficiencies arising from standardisation unilaterally imposed by a dominant undertaking should be taken into account. [39]

Mariotti and Holzweber have highlighted possible concerns with the consumer demand test.

Firstly, consumer demand changes, which makes it crucial to pinpoint a time at which we want to assess it. [40] Secondly, the demand test does not resolve the dilemma concerning bundles which enjoy demand independently of the integral products, and thus may allegedly produce false-positives. [41] Thirdly, the demand test may lead to the counter-intuitive results where there would be no demand for the tying product without the tied product. [42] It will be proposed that neither of these critiques is sufficient to abandon the consumer demand test.

The first concern is not as urgent as it may first appear. The threshold for separateness of products is very low: demand for the tied product as a stand-alone might be dwindling; and it is not required that competitors who offer the tied product exclusively as a stand-alone are *effective* competitors, merely that the number of consumers who accept such offers is 'not insignificant'. [43] Fluctuations of demand are thus unlikely to be a significant obstacle to applying the test. The only remaining problem is the scenario where the competition has already been eliminated and there is no longer any demand for the tied product as a standalone product, which might necessitate a historical analysis on the part of the Commission. However, seeing that one of the Commission's enforcement priorities is prevention of

---

[39] *Microsoft* (n 9) para 1152.
[40] Renato Mariotti 'Rethinking Software Tying', (2000) 17 Yale J. on Reg. 367, 377–379.
[41] Holzweber (n 11) 357.
[42] ibid, 17.
[43] *Microsoft* (n 9) para 932.

foreclosure, rather than reaction to historical foreclosures, the test is well in line with the Commission's actual practice. [44]

As to the second concern, Holzweber gives an example of a painkiller which, when combined with a soporific, results in an antidepressant, which is a drug that operates in an entirely distinct market. [45] Notably, Holzweber's argument is that such tying should not be considered harmful because no changes in market power of the tying and the tied product occur. However, it is respectfully suggested that Holzweber's objection would only be relevant in jurisdictions where tying and bundling are considered *per se* harmful. In EU law, the assessment of separateness does not in any way prejudice the later assessment of the ability of the bundle to foreclose competition. [46]

On a similar note, the third concern is also a non-issue at this stage of the inquiry. Under the second *Microsoft* criterion, an 'anti-competitive bundle' by definition requires the manufacturer to be dominant in the tying product market. If there is no demand for the tying product without the tied product, then there is no competition issue, as it is impossible to be dominant on the market for a product for which there is no demand.

As made apparent by the discussion above, the current separateness criterion is characterized by a low threshold and a tendency to relegate a lot of substantive issues towards the more advanced steps of the analysis. This is commendable: in case of innovative products, the presence on the market of which we have little experience with, it will be difficult to understand at first glance why the demand for the standalone tied products offered by competitors is decreasing. If the threshold is set too high, it may produce unjustifiable overall results, as the Commission will never get around to assess whether the demand for the tied product, insofar as it has dwindled, has done so because the tied

---

[44] Commission Communication (n 1) para 19-20.
[45] Holzweber (n 11) 358.
[46] See 4th *Microsoft* criterion in Part 1.2 above.

product is no longer considered as an efficient standalone, or because of an abusive practice that may cause market foreclosure. On the other hand, the low separateness threshold itself is not capable of harming competition by freezing innovation because it does not in the least prejudice the finding of anti-competitive effects.

B. Coercion/inducement

The big novelty brought forth by the developments on the digital markets is the acknowledgement by the CJEU of a different, less exacting concept of consumer coercion.

In *TetraPak II,* [47] the CJEU insisted on coercion in its narrower sense: a consumer was considered coerced if they were given *no choice* but to buy A-B or to go without one of the products they needed. This concept of coercion will sometimes be sufficient in the context of digital markets as well, as exemplified by the recent Commission's decision in *Google Android* – it would appear from the Commission's press conference statement that its main issue with Google's commercial practices in relation to Android smartphones is that it made the licensing of Google Play to OEMs conditional on the purchase of other Google applications.[48] In such a case, insofar as Google Play is considered a 'must-have' software and the sale of Android smartphones without it would not be commercially viable, the OEMs can be considered coerced.

However, in *Microsoft*, the Commission, subsequently endorsed by the CFI, developed a wider approach to coercion, holding that *mere inducement* of consumers may, in certain

[47] Case C-333/94P *Tetra Pak II* [1996] ECR I-5951, para 132.
[48] European Commission Press release 'Antitrust: Commission fines Google EUR 4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine', IP/18/4581, 18th July 2018.

circumstances, leave them with no choice but to accept the tied product. [49]

What does inducement mean and how intensive does it have to be to constitute abuse? The CFI established no precise threshold. It based its judgment on the facts that 1) Microsoft only sold Windows 98 with WMP to OEMs; 2) WMP was made immune from uninstallation via the standard Uninstall function; 3) market analysis showed that the abovementioned state of facts led the vast majority of consumers to use WMP as opposed to other media players available on the market, even though there was no obvious difference in quality between WMP and the competing products. [50] Hence, it would appear that the test of inducement is now three-fold: 1) a dominant undertaking must adopt a practice of tying or bundling; 2) the practice must make the tied product more attractive to consumers; 3) the consumers must actually be influenced by this commercial practice at the expense of competing products.

On one hand, such an approach is intrusive. It raises an urgent question: to what extent should competition law seek to protect consumers? If they wilfully opt to remain inert and use whatever it is that dominant undertakings choose to offer them, should that decision be theirs to make?

The CFI seemed to place significant emphasis on the fact that back in 2007 when *Microsoft* was decided, the market research compiled by the Commission indicated that a large number of consumers felt uncomfortable about downloading software from the internet. [51] However, more recent research indicates that not all of these concerns are still relevant. According to the Office for National Statistics, in 2017 almost all persons aged 16 to 35 (99%) used internet. Internet use by retired adults has dramatically

[49] *Microsoft* (n 9) para 960-975.
[50] ibid.
[51] *Microsoft* (n *31*) recital 865.

increased from 61% in 2011 to 86% in 2017. [52] The revenue generated by UK businesses from e-commerce has nearly doubled between 2009 and 2017, suggesting that internet purchases are no longer something that the majority of the population fears or avoids. [53]

Perhaps reflecting these changes, the Commission in *Google Search (Shopping)* focused on a different issue: the fact that consumers are not willing to spend time researching the options that are not readily visible. [54] But should the dominant undertaking bear responsibility for what some might dismiss as laziness of consumers? If yes, why and under what conditions?

If the consumer lazily chooses to purchase goods on the basis of little to no research on available competing offers, then the consequences of her decision are her own to bear. It would be unpracticable to force each consumer to spend a determined amount of time researching before making a purchase. If they are adults and have capacity for independent judgment, then their freedom to determine their market behaviour ought to be respected, even if said behaviour is not always rational.

However, the fact that consumers sometimes act lazily should not distract us from the crux of the issue: the behaviour of the dominant undertaking. It is a principle underpinning the entirety of EU competition law that dominant undertakings have a special responsibility to refrain from damaging competition. [55] If a dominant undertaking deliberately presents a consumer with

---

[52] 'Internet Users in the UK: 2017' (Office for National Statistics, 2017) <https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2017> accessed 23 March 2019.

[53]'Internet advertising of businesses – statistics on usage of ads', (Eurostat, 2016) <https://ec.europa.eu/eurostat/statistics-explained/index.php/Internet_advertising_of_businesses_-_statistics_on_usage_of_ads> accessed 23 March 2019.

[54] *Google Search* (n 9) 124.

[55] Commission Communication (n 1) para 1.

two products, one of them popular and desirable, and the other unknown and/or unpopular, and tells her that she cannot have one without the other, then it is the dominant undertaking who has made the first step towards harming the competitive process and has disregarded its special responsibility.

Does the consumer have to use the tied product? No. However, the marketing team of the dominant undertaking is well aware that she is likely to use it – it is why they chose the marketing strategy which involves bundling. While all marketing strategies are designed to induce particular consumer behaviour, in these circumstances it is the undertaking's dominant position in the tying market which is used to induce the consumer and which allows the tied product to bypass competition. Perhaps the tied product is just as good as the competing products and the consumer is acting perfectly reasonably in continuing to use it. We will never know because the tied product, by virtue of the bundle, has never had to compete with its substitutes on merit. Hence, by taking into account the way consumers actually behave in the market, the Commission in cases like in *Microsoft* and *Google Android* strives to better protect the competition process itself, and in so doing enhance the consumer's opportunity to make a rational economic choice. Whether she ultimately makes use of that choice or not is then for herself to decide.

4. Refusal to supply

In *Oscar Bronner,* [56] the CFI held that for refusal to supply to be abusive, the following conditions would have to be made out:

> 1. the product marketed in market A needs to be indispensable for the entry of a new product for which there is a potential consumer demand in the separate market B;

[56] Case C-7/97 *Oscar Bronner* [1998] ECR I-7791, para 32 and 37.

2. the refusal must be incapable of being objectively justified; and

3. the refusal must be likely to eliminate all competition.

Just like in cases of tying and bundling, refusal to supply cases related to digital markets have also generated some controversy. The terms 'new product', 'separate market', and 'indispensability', as well as the requisite degree of remaining competition have all undergone developments and will be considered in turn below.

A. Entry of a new product in a separate market

In the early refusal to supply case law such as *Magill* [57] or *Bronner,*[58] the complainants' projects of introducing a new product to the market were clearly defined. However, with the development of technology, the concept of a 'new product' became difficult to pin down.

If we consider *Magill,* the product at issue was a novel type of TV programme guide which would have been broader in scope than pre-existing guides and would have allowed consumers to plan their weekly TV schedule. [59] In comparison, the more recent *Microsoft* decision [60] was concerned with novelty in a broader sense. There was no single product that was about to be introduced. Rather, the Commission and the CFI were concerned about a multitude of potential future products which might be developed and might be valuable to consumers. Hence, the new criterion for 'new entry' fixed by the CFI was that of a new product *or* technological innovation. Interestingly, no indication was given by the Commission or the CFI as to the likelihood that these products would be developed or that they

[57] Case C-241-242/91P *Magill TV Guide* [1995] ECR I-743.
[58] *Oscar Bronner* (n 56).
[59] *Magill TV Guide* (n 57) para 10.
[60] *Microsoft* (n 9), para 632; in relation to software interoperability, *not* to the Windows-WMP tie assessed in Part 1.2.

would actually enjoy any consumer demand. This raises a question whether competition law should be protecting hypothetical products and the general idea of innovation, without any clue as to what that innovation might look like.

A similar question arises in relation to the criterion of separate markets. First, in *IMS Health*, the Court accepted as sufficient the existence of a potential or a hypothetical market. It remained determinative that two different interconnected stages may be identified and that the upstream product be indispensable for the supply of the downstream product, but it was no longer required that the upstream product be marketed as such. [61] In *Microsoft,* the CFI did make a finding of separate markets for operating system ('OS') and for software using the OS as a platform. However, *IMS Health* was cited with approval and to that extent remains good law. [62] Hence, we are now in a situation where Art 102 TFEU appears to be invokable for the protection of hypothetical products on hypothetical markets.

This development is to be welcomed, though it may at first glance appear strange. As noted in Part 2, digital markets are particularly influenced by network effects. The trouble is that where a platform technology affected by network effects – such as an OS or a general search – becomes dominant and then chooses to refuse access by competitors' products to the platform, it becomes impossible for *any* products, other than those created by the dominant undertaking, to enter the market. Closely related is the fact that *no* related new downstream markets can be developed now or in the future – or they can only be developed with the dominant undertaking as the monopolist fulfilling the entirety of the consumers' demand. This is actually a much graver situation than in *Magill* where the dominant undertaking would only have control over a very specific (and on the facts of that case quite niche) product and product market. Therefore, a test that can detect threats to hypothetical markets is

[61] Case C-418/01 *IMS Health* [2000] ECR I-5039, para 44-45.
[62] *Microsoft* (n 9) para 335.

essential for preserving the possibility of entry into digital markets.

B. Indispensability of product and effective competition

While indispensability of the upstream product and the existence of effective competition started out as two distinct criteria, it is suggested that they have become closely intertwined and that their interpretation has been influenced by identical concerns. In particular, it will be shown that the Court has moved away from the concept of 'indispensability for the creation of a new product' and towards 'indispensability for commercial viability of the product', where this viability is necessary to maintain effective competition.

Let us begin by analysing the *Oscar Bronner* decision and its indispensability criterion. In *Bronner*, the CFI was adjudicating on a case concerned with distribution of daily newspapers. The question was whether access to a home-delivery scheme operated by the dominant undertaking was indispensable for a new entrant to the daily newspapers market. The Court refused to recognise access as indispensable, as there existed other means through which the seller of newspapers could reach its customers – e.g. through sale in shops and kiosks or delivery by post. The CFI emphasized that while these options may be less advantageous than home-delivery, their existence and their actual use for the purpose of newspaper delivery precluded the finding of indispensability of access. The threshold that the CFI used was that refusal to supply must not make it impossible or unreasonably difficult for the competing undertaking to enter and operate on the relevant market. [63]

In comparison, the Commission and the CFI in *Microsoft* chose to interpret indispensability as a test based on commercial viability rather than the practical possibility of an alternative

[63] *Oscar Bronner* (n 56) para 43-44.

access route. [64] The Court recognized that different OS such as Linux do exist on the market and could be relied on by software manufacturers to provide a platform on which their software might function and might be used by consumers. It also recognized that the data about Windows that was publicly accessible did provide a limited possibility for ensuring interoperability between Windows and non-Microsoft software. These opportunities for market access were less favourable for software manufacturers than the one demanded by Commission. However, it is indisputable that there were alternative means through which competitors could enter into the relevant market, and therefore there is a strong argument that, on orthodox application of *Oscar Bronner* the Commission's case should have failed.

On a similar note, *Microsoft* dispensed with the *Oscar Bronner* criterion of elimination of *all* competition, holding instead that elimination of *all effective* competition would suffice for an Art. 102 TFEU violation. Having considered the market shares of the different OS and the potential for development of software without disclosure of interoperability data, the CFI concluded that while some competition would remain possible, it would be so niche as to be incapable of countervailing Microsoft's dominant position. [65]

An analogous approach was taken by the Commission in *Google Search (Shopping)*. The Commission's decision addressed Google's practices in marketing its specialized shopping search. At the time the Commission began its investigation, Google Shopping used to be positioned above normal search results in a special unit alongside other specialized Google products such as Google Maps, which made it more visible to users than competing specialized search engines. Furthermore, Google Shopping was exempt from the influence of an algorithm that Google Search used to rank websites according to relevance to

[64] *Microsoft* (n 9) para 632.
[65] *Microsoft* (n 9), para 563.

the search query, which resulted in Google Shopping offers being consistently displayed at the top of general search page, in contrast with competing offers, which would typically be displayed around page 4. [66]

The Commission did not announce the theory of harm on which it based its *Google Search (Shopping)* decision. [67] Notably, it distinguished *Oscar Bronner* on the facts as a case involving mere refusals to supply, whereas Google was found to have taken active steps to prevent equal access, such as algorithm modification. [68] However, that is not the only ground of divergence. There are elements in the Commission's decision which clearly relate to indispensability of access, and which equate indispensability with commercial viability. For instance, the Commission took into account the existence of other search engines such as Yahoo or Bing. [69] It also considered Google's argument that no competitors were completely excluded from the general search – in other words, supply was never *per se* refused. [70] Yet, it finally concluded that these facts could not preclude a finding of indispensability. The market shares of other search engines across Europe were so small that any provider of specialized shopping search who wanted to operate on more than a niche market, would have to advertise on Google. [71] Moreover, Google would have to subject its own services, and services provided by competitors, to the

[66] *Google Search* (n 9).
[67] For two contrasting views on this issue see in particular Edward Iacobucci and Francesco Ducci, 'The Google search case in Europe: tying and the single monopoly profit theorem in two-sided markets' (2018) European Journal of Law and Economics, accessed online from Springer Link, and Pinar Akman, 'The Theory of Abuse in Google Search: A Positive and Normative Assessment under EU Competition Law' [2017] Journal of Law, Technology and Policy 302.
[68] *Google Search* (n 9) recital 650.
[69] ibid recital 306.
[70] ibid 35.
[71] ibid recital 306.

same processes and methods for positioning and display in its search results pages. [72]

As a matter of evidence, particular importance was attributed by the Commission to analyses which showed that the majority of traffic is distributed between the first three general search results, with only very few users actually perusing the links at the bottom of the first page and little to no users ever reaching page 4. [73] Furthermore, the traffic enjoyed by Google Shopping increased dramatically since Google exempted it from the algorithm and adorned it with appealing design features (including prominent positioning on the page and images of offered items). [74]

The move from strict differentiation between indispensability and presence of *any* competition towards a more unified concept of indispensability of new products/innovation for the maintenance of *effective* competition is commendable for its coherence with the broader logic of competition law: competition law is concerned with dominance affecting the market as a whole, not small groups of consumers with niche preferences. It will be of little solace to a non-dominant undertaking if it can reach a thousand consumers if it needs to be reaching hundreds of thousands in order to develop a strong enough network and survive on the market. Similarly, it will not be enough to maintain effective competition if the non-dominant undertaking is only granted access, due to a biased algorithm, to page 4 of a search engine. Although access to the platform will not have been technically refused in such a scenario, what is truly indispensable in digital markets is not access to an advertising platform such as Google Search per se, but rather access to a visible place on the platform – for which the downstream

[72] ibid, recital 699-700.
[73] ibid 124.
[74] ibid 123.

undertakings must be allowed to compete on merit if effectiveness of competition is to be preserved.

5. Conclusion

This article sought to assess the changes brought to the interpretation of Art 102 TFEU by the emergence of novel digital markets. It has been suggested that these markets, influenced as they are by network effects, are particularly susceptible to abuse by tying and bundling, and refusal to supply, necessitating the Commission and the CJEU to re-assess the interpretation of Art 102 TFEU, so as to better protect the irrational consumer.

The legal approach to both types of abusive conduct has undergone considerable developments, with consumer behaviour playing an increasingly central role in the interpretation of Art 102 TFEU. Notably, the gateway criteria for both tying and bundling, and refusal to supply, have been lowered: low levels of consumer demand will suffice for satisfying that 'separate products' limb of the anti-competitive tying and bundling test; hypothetical products and hypothetical markets will suffice for satisfying the 'new product' and 'separate market' limbs of the anti-competitive refusal to supply. These indicate, on the part of the CJEU, the willingness to explore the potential of each instance of allegedly abusive conduct to harm the competitive process and, by necessary implication, the consumer, even when the conduct is entirely novel and would not have easily fit into the old interpretations of Art 102 TFEU.

In addition, the limbs of the test that do directly address the potential for harming competition – coercion in case of tying and bundling; indispensability in case of refusal to supply – have also been interpreted to become more sensitive to consumer behaviour and to consumer irrationality in particular. These developments are to be welcomed insofar as they accurately represent the market realities and better preserve the competitive

process, which enables the consumer to choose between products competing on their merits.

UNIVERSITY · OF · OXFORD