

Issue [4] 2015



OXFORD UNIVERSITY
UNDERGRADUATE LAW JOURNAL



OXFORD UNIVERSITY

UNDERGRADUATE LAW JOURNAL

The views expressed by the contributors are not necessarily those of the Editorial or Honorary Board of the Oxford University Undergraduate Law Journal. Whilst every effort has been made to ensure that the information contained in this journal is correct, the Editors cannot accept any responsibility for any errors or omissions, or for any consequences resulting therefrom.

© 2015 Individual authors

-The-
4th Editorial Board
of the Oxford University Undergraduate Law Journal
2014-2015

Editors-in-Chief

Wei Jian Chan
Harris Manchester College

Yu Jie Wu
Magdalen College

Editors

Vincent Ooi
Trinity College

Denise Lim
Lincoln College

Associate Editors

Matthew Chan
Exeter College

Simon Herzberg
Harris Manchester College

Daniel Judd
Jesus College

Ern Huei Lim
Lady Margaret Hall

Esther Mak
University College

Rachel Tan
New College

Rachel Wong
Corpus Christi College

Sheryl Yeo
St Hilda's College

General Editors

Sophie Scholl
Exeter College

Aidan Coghlan
Harris Manchester College

-The-
Honorary Board

Sir Nicolas Bratza

Professor Michael Bridge

Donald Findlay QC

Professor Christopher Forsyth

Ian Gatt QC

The Right Hon the Lord Judge

The Right Hon the Lord Kerr of Tonaghmore

Michael Mansfield QC

The Right Hon the Lord Neuberger of Abbotsbury

The Right Hon the Lord Phillips of Worth Matravers

Lord Pannick QC

His Honour Judge Gordon Risis CB

Dinah Rose QC

Sir Konrad Schiemann

The Right Hon the Lord Wilson of Culworth

CONTENTS

i.	Preface <i>Paul S Davies, St Catherine's College</i>	1
ii.	Introduction <i>Wei Jian Chan and Yu Jie Wu, Editors-in-Chief</i>	2
iii.	Contributory Negligence and Intentional Trespass to the Person: Rethinking Pritchard & the Section 4 Definition of Fault <i>Joshua Pike</i>	3
iii.	Finding Principle in Illegality: Reflections on <i>Tinsley v Milligan</i> <i>Matthew Chan</i> Joint Winner of the Norton Rose Fulbright Prize for the Best Article in Contract, Tort, Trusts, and Land Law	13
iv.	The Right Approach to Wrongful Conception <i>Shaun Elijah Tan</i>	28
v.	Dividing the Single Indivisible Transaction: Balancing the Interests of Mortgagees and Innocent Occupants <i>Vincent Ooi</i> Joint Winner of the Norton Rose Fulbright Prize for the Best Article in Contract, Tort, Trusts, and Land Law	37
vi.	Prisoners' votes, referendums, and common law rights: A case note on <i>Moohan v Lord Advocate</i> <i>Paul Fradley</i>	53
vii.	A Failure to Protect: The Shortcomings of the United Kingdom's Judicial Response to the Privatisation of the Public Realm <i>Owen Nanlohy</i>	63
viii.	Something for Nothing: Explaining Single-Sided Contract Variations <i>Andrew Hill</i>	75
ix.	Broad, Inflexible and Redundant?: Fixing The Anti-Avoidance Rule in Section 75A Finance Act 2003 <i>Vincent Ooi</i>	96

Preface

Paul S. Davies, St Catherine's College

My students often remind me that studying law as an undergraduate in Oxford is not an easy option. Much is (rightly) expected of students in tutorials, and it would be easy for undergraduates to lament that simply getting through the reading lists can consume all their time and energy. Yet Oxford students show remarkable resilience, and the range of further activities in which students participate whilst keeping on top of their academic work is truly impressive. It seems that anything a student would like to do whilst at Oxford is possible – including being part of the Oxford University Undergraduate Law Journal.

For students to “be on the law review” is a typically American phenomenon. And in America, those students are invariably graduate students. In this jurisdiction, few journals are put together by students at all, let alone by undergraduate students. The editors and others behind this journal deserve tremendous credit for the huge amount of work that has gone into putting together this edition.

The Oxford University Undergraduate Law Journal serves a useful function in allowing students to engage in extended writing beyond the confines of the undergraduate course. It is encouraging to see students keen to explore the law further than the Jurisprudence syllabus requires, and the results are impressive. It is to be hoped that the authors have enjoyed the process associated with publishing their work, and that they will continue to engage in scholarly writing for other academic journals once their studies are over. The articles in this volume deserve a wide audience, and will be of particular interest to all those currently studying law as undergraduates. It is testament to the quality of the undergraduate community here that the contents are so worthwhile, and I look forward to reading next year's edition already.

Introduction

Wei Jian Chan and Yu Jie Wu, Editors-in-Chief

It brings us great satisfaction to introduce the fourth edition of the Oxford University Undergraduate Law Journal ('OUULJ').

We believe strongly that law is genuinely intellectually interesting, and not simply a means to an end. This journal reflects this sentiment. Since its humble beginnings in 2011, the OUULJ has sought to provide a platform for students to reach beyond the boundaries of the undergraduate law syllabus at Oxford. We have tried to foster debate about current legal issues, on the parts of both our contributors and readers. We hope that in years to come, the OUULJ will continue to be a useful platform for intellectual exploration and debate.

In 2015, the OUULJ organised its first outreach event. We are grateful to Professor Peter Mirfield and Associate Professor Jeremias Prassl for taking time out of their busy schedules to share their experiences in the legal publishing world with us. This event, we hope, will be the first of many to come, and we hope that the OUULJ will become a catalyst for academic writing in all forms.

No publication can be completed without its editors. We would like to thank our editorial board for their hard work in preparing the fourth edition of this Journal. In particular, we would like to give credit for the tireless efforts of our excellent Editors, Denise and Vincent; without their contributions, this edition would not have been possible. Lastly, we would like to thank Professor John Cartwright and Associate Professor Paul Davies for their thorough review of the journal and for helping to select the winners of the Norton Rose Fulbright Prize.

If you read this introduction as an undergraduate law student, we invite you to share your legal opinions with the OUULJ. It would be our pleasure to read about your views, and to share them with the world.

Contributory Negligence and Intentional Trespass to the Person: Rethinking *Pritchard* & the Section 4 Definition of Fault

Joshua Pike*

I. INTRODUCTION

Where once there were tentative suggestions in the textbooks that the Law Reform (Contributory Negligence) Act 1945 (the '1945 Act') permitted the apportionment of damages for contributory negligence in cases of intentional trespass to the person,¹ such tentative suggestions have now given way to conclusive statements to the contrary.² This change in position has been brought about by the Court of Appeal decision in *Pritchard v Co-Operative Group Ltd.*³ The purpose of this article is to attempt to challenge both the legal and normative validity of this new position that now appears to have become relatively unchallenged orthodoxy.

The attempt will comprise three parts. First, it will be shown that as a matter of pure statutory construction there is nothing in the 1945 Act that prevents its application to cases of intentional trespass to the person. It will be shown that this remains the case despite the judicial gloss placed on the section 4 definition of fault by the House of Lords in *Reeves v Commissioner of Police of the Metropolis.*⁴ This judicial gloss (the 'dual limb' approach) splits the section 4 definition in two, restricting each half to the actions of the defendant and the claimant respectively. It will be suggested that section 4 in the light of *Reeves* entails nothing more than importing the common law test for contributory negligence into the statute. Emphasis is placed on it solely importing the common law test, namely whether the claimant took reasonable care for his or her own safety,⁵ rather than also restricting the section's scope to cases where the common law outcome of the defence would have previously applied, i.e. where it absolved the defendant of all liability as a complete defence. Maintaining this distinction between the substantive legal test that relates to the actions of the claimant and the actual impact of the defence on the claim for damages is crucial.

Second, it will be argued that the Court of Appeal in *Pritchard*, in purporting to apply the same dual limb approach to the section 4 definition of fault as adopted in *Reeves*, fundamentally misunderstood this approach when it held that the new statutory doctrine of contributory negligence only applied in cases where the prior common law defence operated to absolve the defendant of all liability, thereby refusing to apply it in cases of intentional trespass to the person. It will be shown that this approach is irreconcilable with the facts of *Reeves* because it fails to maintain this crucial distinction between the test at common law and the outcome at common law. Only the incorporation of the former into section 4, and not the latter, is compatible with *Reeves*. It will be argued that this approach is not necessarily contrary to Parliament's intention in enacting the 1945 Act.

Third, it will be argued that the apportionment of damages in section 1(1) of the 1945 Act, regardless of Parliament's intention, has fundamentally altered the substance of contributory negligence as part of a wider trend in the law of tort towards shared responsibility. The approach of *Pritchard*, in its arbitrary adherence to pre-1945 case law, fails to take account of this fundamental change in emphasis and overlooks the possibility of a more just and equitable approach to shared fault in cases of intentional trespass to the person.

* Worcester College, Oxford. I am grateful to Mr Donal Nolan for his comments on an earlier draft. All errors remain my own.

¹ See, e.g., W.V.H Rogers (ed), *Winfield and Jolowicz on Tort* (18th edn, Sweet & Maxwell 2010) 365.

² W.E. Peel and J. Goudkamp (eds), *Winfield and Jolowicz on Tort* (19th edn, Sweet & Maxwell 2014) 708.

³ [2011] EWCA Civ 329, [2010] QB 320.

⁴ [2000] 1 AC 360 (HL).

⁵ *Nance v British Columbia Electric Railway Co Ltd* [1951] AC 601 (PC), 611 (Viscount Simon).

II. THE STATUTORY DEFINITION OF FAULT AND THE REEVES INTERPRETATION

Section 1(1) of the 1945 Act provides that the court shall reduce any award of damages ‘to such extent as the court thinks just and equitable having regard to the claimant’s share in the responsibility for the damage.’ This apportionment is to be applied ‘where any person suffers damage as the result partly of his own fault and partly of the fault of any other person or persons.’ The meaning of fault is given by section 4. It says fault means ‘negligence, breach of statutory duty or other act or omission which gives rise to a liability in tort or would, apart from this Act, give rise to the defence of contributory negligence.’ On an ordinary reading it is difficult to see why the defence is automatically precluded from applying to cases of intentional trespass to the person: section 4 appears to do nothing more than import standard meanings of fault into the Act. The phrase ‘other act or omission which ... would, apart from this Act, give rise to the defence of contributory negligence’ would appear to simply serve a residual function, ensuring that the preceding definitions cover all scenarios that had previously given rise to the common law defence. This would have been a conventional reading of the section, given the truism that statute overrides common law, but the courts have chosen not to take this approach.

The House of Lords case that establishes the current approach to section 4 is *Reeves*. In *Reeves* the claimant (the deceased’s estate) brought an action in negligence against the police after the deceased committed suicide whilst in police custody. The deceased had created a ligature through an open spy hole and thereby hung himself. The spy hole was only accessible due to the negligence of the police. The question arose as to whether there could be an apportionment of damages between the two parties for contributory negligence. *Reeves* is a landmark case because the House held that contributory negligence could also be established upon an intentional act of the claimant, rather than exclusively on a negligent one. However, for present purposes, it is the House’s general approach to the section 4 definition of fault that is of significance. It was said that section 4 should be split into two constituent limbs, one applying to the defendant and one applying to the claimant.⁶ The part relating to liability arising in tort ‘is concerned with fault on the part of the defendant.’⁷ The question to be asked in relation to the claimant was framed by Lord Hoffmann as ‘whether, apart from the [1945] Act’, the actions of the claimant ‘would have given rise to a defence of contributory negligence’.⁸ Thus section 4, once this judicial gloss is added, effectively reads:

Fault, in relation to the defendant, means negligence, breach of statutory duty or other act or omission which gives rise to a liability in tort.

Fault, in relation to the claimant, means an act or omission which would, apart from this Act, give rise to the defence of contributory negligence.

Despite this gloss placed on the section it is still difficult to see how this precludes the application of contributory negligence to cases of intentional trespass to the person. An example will help to illustrate this point. A physically assaults B, causing a fight. B hits A back, wounding A. A sues B in tort for battery. B admits liability but contends that both parties are at fault and asks the court to apply the 1945 Act and apportion damages accordingly. On the dual limb approach to section 4 adopted by *Reeves*, B could not rely on A’s prior assault being an act that gives rise to liability in tort in order to establish fault for the purposes of section 1. This part of the definition is restricted to the actions of B alone. However, the second limb refers to the actions of the claimant that, prior to the Act, would have

⁶ *Reeves* (n 4) 369 (Lord Hoffmann), 382 (Lord Hope).

⁷ *ibid* 369.

⁸ *ibid*.

given rise to the defence of contributory negligence. The test for determining whether an act of the claimant constituted contributory negligence at common law is whether the claimant failed to take reasonable care for his or her own safety.⁹ The initiation of a fight through physical assault by the claimant, as in the above illustration, could easily constitute a failure to take such reasonable care, depending of course on the particular facts. Thus it can be seen that the dual approach to the definition of fault as adopted in *Reeves* does not *prima facie* preclude the application of section 1 to cases of intentional trespass to the person where ‘apart from this Act, [would have given] rise to the defence of contributory negligence’ is taken to refer to the common law *test*. This is where it becomes crucial to maintain the distinction between the substantive common law *test* and its *outcome* as regards the award of damages. It is in *Pritchard* that this distinction is ignored.

III. THE PRITCHARD DEFINITION OF FAULT AND INTENTIONAL TRESPASS TO THE PERSON

Pritchard was a Court of Appeal case involving the tort of battery. The claimant was an employee of the defendant company. Having been on sick leave for two weeks the claimant attempted to persuade her manager to give her an additional day off as she was still feeling unwell. When the manager refused, the claimant, along with her sister and a friend, went to the store to confront the manager. After a heated exchange, in which the claimant was verbally abusive towards her manager, the manager laid hands on the claimant in an attempt to forcibly remove her. A struggle ensued whereby the claimant bit the manager. The claimant brought an action for damages, alleging that the assault by the manager had given rise to a near complete psychiatric breakdown, depression and agoraphobia and caused her inability to work. The defendant company admitted liability but sought to have damages reduced for contributory negligence. The Court of Appeal held that contributory negligence under the 1945 Act was not applicable to cases of intentional trespass to the person.

In his judgment, Aikens LJ purported to adopt the dual limb approach employed by *Reeves*. He said that fault, in relation to the claimant, meant an act or omission ‘which would at common law and but for the 1945 Act have given rise to the complete defence of “contributory negligence” to the claim being made by the claimant.’¹⁰ Aikens LJ went on to conduct a comprehensive review of the case law prior to 1945 before concluding that, at common law, contributory negligence was ‘not a defence in the case of an “intentional tort” such as assault and battery.’¹¹ The conclusion that there was no common law defence of contributory negligence to an intentional tort will not be contested here. The reasoning which follows from this is also sound, given the interpretation of section 4 that the Court of Appeal adopted. The approach appears to be this: (i) the courts will examine the case law prior to the 1945 Act; then (ii) if the facts under consideration would not have given rise to the complete defence of contributory negligence at common law then the 1945 Act, and consequently the defence, does not apply. On this view the 1945 Act does nothing more than reform the *effect* of the existing common law defence through the introduction of apportionment; the Act does not affect the *substance* of the defence.

⁹ *Nance* (n 5) 611 (Viscount Simon).

¹⁰ *Pritchard* (n 3) 330.

¹¹ *ibid* 338.

This approach relies heavily on the reasoning of Lord Hoffmann in *Standard Chartered Bank v Pakistan National Shipping Corporation and others (Nos 2 and 4)*.¹² This was the second House of Lords decision, after *Reeves*, to adopt the dual limb approach to section 4. Here the claimant was claiming damages for deceit and the defendants sought to rely on the 1945 Act, alleging that the claimant was also partly at fault for their loss and that therefore any damages should be reduced for contributory negligence. Lord Hoffmann framed the question as ‘whether at common law [the claimant’s] conduct would be a defence to its claim for deceit.’¹³ His Lordship, with whom the House unanimously agreed, concluded that it would not have been. The result was that the 1945 Act could not be applied to the facts at hand. Lord Hoffmann considered this approach to be in accordance with the purpose of the Act, ‘which was to relieve plaintiffs whose actions would previously have failed and not to reduce the damages which previously would have been awarded against defendants.’¹⁴ It would seem, therefore, that the decision in *Pritchard* is completely sound legally, backed up by no less than two decisions of the House of Lords. However, the approach in both *Pritchard* and *Standard Chartered Bank* loses sight of the fundamental distinction between the substantive common law test and the outcome it had on any award of damages. That this is so can be shown by reference back to the case of *Reeves*.

IV. THE CORRECT INTERPRETATION? COMPARING PRITCHARD WITH REEVES

Whilst the Court of Appeal clearly thought that it was applying established law it will be argued that *Pritchard*, upon closer inspection, amounts to a misinterpretation of the dual limb approach. As above, the House in *Reeves* held that the second limb of the definition, the limb that relates to the actions of the claimant, required an act or omission that would, apart from the Act, give rise to the defence of contributory negligence.¹⁵ As explained above, this can be read as entailing nothing more than asking whether the claimant took reasonable care of his or her safety, the common law test, and does not preclude the application of the 1945 Act to intentional trespass to the person. However, the question that the Court of Appeal posed in *Pritchard* was whether the act or omission would have given rise to the *complete* defence of contributory negligence.¹⁶ Admittedly the difference appears initially to be mere semantics but this is to confuse the important distinction already mentioned between the substantive content of the common law test and the effect that the defence had on liability.

In short there are two possible interpretations of the second limb of the section 4 definition of fault. Either the one proposed above is adopted or section 4 is read as not only incorporating the pre-existing test for contributory negligence but also restricting the use of that test, and thus the statute and apportionment of damages, to cases in which the defendant would have been absolved of all liability. The latter reading incorporates both the pre-existing test *and* the pre-existing outcome of the common law defence. The approach in *Pritchard*, by using the phrase ‘complete defence’, follows the latter interpretation. As will be shown, this interpretation is incompatible with the facts of *Reeves*. In order to demonstrate this it will be helpful to apply the *Pritchard* interpretation to the facts of *Reeves* itself to assess whether the approach of Aikens LJ amounts to a substantive departure from the approach taken by the House of Lords.

¹² [2002] UKHL 43, [2003] 1 AC 959.

¹³ *ibid* 966.

¹⁴ *ibid* 965.

¹⁵ *Reeves* (n 4) 369.

¹⁶ *Pritchard* (n 3) 330.

As *Reeves* and other cases demonstrate the police owe a duty of care to those in their custody. They are liable if they negligently breach this duty of care, as was the case in *Reeves* itself. The issue is that applying the *Pritchard* approach to section 4 will necessarily obviate such duties of care. If section 4, as this article proposes, incorporates solely the common law test into its definition then the question to be asked is whether the claimant failed to take reasonable care for his or her safety. In *Reeves* the House clearly felt that the claimant, after deciding that intentional acts can constitute contributory negligence, had satisfied this test. However, per the approach in *Pritchard* the claimant must be held to have failed to take reasonable care for his or her own safety *and* it must be a case in which prior to the Act the defendant would have been absolved of all liability. On the facts of *Reeves* the latter requirement is not met. Otherwise an absurd and unjust result is reached: namely that the police owe a duty of care to detainees but only up to the point *before* that detainee commits suicide.

Just as it is contradictory to simultaneously maintain that the police have a duty to prevent a detainee from committing suicide and to hold that suicide constitutes a *novus actus interveniens*,¹⁷ so too is it contradictory to hold that the complete defence of contributory negligence could apply in conjunction with such a duty. Thus, had *Reeves* been decided prior to 1945 the courts must have held that the common law defence did not apply or abandon the notion that police owe an effective duty of care towards detainees. This demonstrates the misunderstanding of the dual limb approach that *Pritchard* amounts to, having failed to separate the substantive common law test from the outcome of successfully pleading the defence; the approach adopted by the Court of Appeal, whilst *prima facie* a valid one, is actually irreconcilable with the facts of the very House of Lords case it is based upon.

This same reasoning can be applied to cases of intentional trespass to the person. To return to the simple illustration above, where A physically antagonises B into a fight in which A is injured sufficient for A to sue B in tort, whilst the common law test for contributory negligence might be capable of being satisfied the court could not apply the defence without obviating B's duty to respect A's bodily integrity. Thus the common law defence was barred from applying in actions for battery or assault. Of course, the same issue arises in the standard negligence cases where the common law defence did apply, but here the courts were willing to live with the injustice caused by the operation of the complete defence. The crucial point is that it is no longer necessary to choose between these two extremes given the introduction of apportionment. The *Pritchard* approach to section 4 refuses to acknowledge this change by restricting the scope of the 1945 Act in the way that it has.

However, there are a number of issues with adopting the proposed interpretation of section 4 over that adopted in *Pritchard*. The first, and perhaps the greatest, obstacle is the *Standard Chartered Bank* case itself. The approach of Aikens LJ effectively mirrors the approach of the House of Lords in that case, except in relation to intentional trespass to the person rather than deceit. There are two important points to make here. The first is that *Standard Chartered Bank* is also based on the dual limb approach in *Reeves* and so consequently the arguments raised against *Pritchard* above apply equally to *Standard Chartered Bank*. The second point is that Lord Hoffmann did not go as far as to say, as Aikens LJ did, that section 4 required the act or omission of the claimant to give rise to the *complete* defence at common law. Rather he used the same language as in *Reeves*. This means that it is still possible to accommodate *Standard Chartered Bank* within a new interpretation of section 4.

The second issue with the proposed interpretation is the argument that it was clearly Parliament's intention that the effect of the 1945 Act was to be so restricted to cases where the defendant would previously have been absolved of all liability. This is how Lord Hoffmann conceived of Parliament's intention in *Standard Chartered Bank*, holding that the purpose of the Act was to

¹⁷ See Donal Nolan, 'Negligence Liability for Suicides in Custody' (2000) 8 Tort L Rev 91.

‘relieve plaintiffs whose actions would previously have failed and not to reduce the damages which previously would have been awarded against defendants.’¹⁸ Aikens LJ relied on this conception of Parliament’s intent in his judgment in *Pritchard*. Lord Hoffmann puts forward little evidence for this contention, however. The only part of the statute that could be relied upon is section 1(1) where it states ‘a claim in respect of that damage shall not be defeated by reason of the fault of the persons suffering the damage’ but it does not necessarily follow that the defence should be restricted accordingly to exclusively those cases. Just as there were cases where the claimant’s action failed, leading to unjust results, so too, as *Reeves* demonstrates, would there be cases where the defendant’s defence would have to fail. The introduction of apportionment renders both of these extreme outcomes unnecessary. There is evidence that can be adduced to suggest that Parliament did in fact intend contributory negligence to be broader than its common law ambit. This can be seen by reference to the Torts (Interference with Goods) Act 1977.

Section 11(1) of the Torts (Interference with Goods) Act 1977 states that contributory negligence is no defence in cases of intentional trespass to goods; Aikens LJ referred to this provision in *Pritchard*, stating that ‘there was no doubt that there was no such defence in relation to intentional trespass to goods’ at common law and that s 11(1) ‘merely makes the position plain’.¹⁹ However, applying the presumption that Parliament never legislates in vain, if the position at common law was as clear as Aikens LJ claims then that points not to the conclusion that s 11(1) ‘merely makes the position plain’ but that the 1945 Act is not automatically precluded from applying where previously the complete defence at common law would not have arisen on the facts. Otherwise s 11(1) is superfluous. There are clearly doctrinal differences between interference with goods and interference with bodily integrity but the point is that the provision should not have been necessary if Parliament’s intention in respect of the 1945 Act was as clear as *Pritchard* and *Standard Chartered* would suggest. Whilst admittedly this argument may seem overly technical and remote it is equally the case that the intention of Parliament was presumed rather than sufficiently evidenced in *Pritchard* and *Standard Chartered Bank*.

Ultimately, however, what Parliament intended is now a moot point. This is because, it will be argued, the introduction of apportionment fundamentally altered the doctrine of contributory negligence. In its misplaced preoccupation with pre-Act case law *Pritchard* has failed to reflect this fundamental change. In doing so, the Court of Appeal overlooked the new normative arguments in favour of allowing the apportionment of damages in cases involving intentional trespass to the person.

V. THE NEW LEGAL LANDSCAPE: NORMATIVE JUSTIFICATIONS FOR ALLOWING APPORTIONMENT IN CASES OF INTENTIONAL TRESPASS TO THE PERSON

Where before it would have been unthinkable to deny the claimant any and all damages in cases of intentional trespass to the person, apportionment allows for a just and equitable reflection of corresponding fault within monetary awards. This change in the law has had a fundamental effect not just on the defence of contributory negligence but also on the law of tort as a whole. As Jenny Steele has said, the 1945 Act reflects a growing trend towards the ‘exorcism of absolutes’.²⁰ It represents a

¹⁸ *Standard Chartered Bank* (n 12) 965.

¹⁹ *Pritchard* (n 3) 337-8.

²⁰ Jenny Steele, ‘Law Reform (Contributory Negligence) Act 1945: Collisions of a Different Sort’ in Arvind and Steele (eds), *Tort Law and the Legislature: Common Law, Statute, and the Dynamics of Legal Change* (Hart 2013) 160.

fundamental shift towards the ‘sharing of responsibility and denial of ideas of sole responsibility or sole cause’.²¹ This gives rise to a wealth of new normative justifications for extending the defence of contributory negligence beyond its common law ambit; arguments which the arbitrary approach to section 4 in *Pritchard* fails to take account of. That this is so can be seen even in *Pritchard* itself. Smith LJ, though she felt compelled to follow the reasoning of Aikens LJ, did so ‘with regret’ because she believed that ‘apportionment ought to be available to a defendant who has committed the tort of battery where the claimant has, by his misconduct, contributed to the happening of the incident, for example by provocative speech or behaviour.’²²

At the heart of this frustration that Smith LJ felt is the court’s illogically strict adherence to pre-Act case law. Lord Rodger in *Standard Chartered Bank*, when coming to the conclusion that contributory negligence was not available as a complete defence at common law to the tort of deceit, remarked that such a state of affairs would have been an ‘extreme doctrine’ given that it would ‘[absolve] the fraudulent defendant of all liability.’²³ However, to then hold that modern contributory negligence should not therefore apply is illogical and fails to reflect this fundamental shift in tort law towards shared responsibility. It is illogical because the reason for the introduction of apportionment is also given as the reason for restricting its use despite the fact that apportionment provides the means of rectifying the unjust results the complete defence would have otherwise resulted in.

As Sedley LJ remarked in *Vellino v Chief Constable of the Greater Manchester Police*,²⁴ there is no ‘substantial justice ... in sacrificing a judicial apportionment of responsibility on the altar of a doctrinal refusal to adjudicate.’²⁵ Instead, apportionment ‘[exemplifies] an idea that shared responsibility is the most appropriate and most nuanced approach’.²⁶ The case law prior to 1945 that *Pritchard* and *Standard Chartered Bank* cling to is devoid of this nuance. This makes it ‘impossible’, in the words of Lord Wright, ‘to get a true theory of contributory negligence’ prior to the 1945 Act.²⁷ This is because the cases are ‘apt to be unsatisfactory and misleading’²⁸ by virtue of suffering from a kind of tunnel vision: a binary choice between full responsibility or no responsibility at all. It is probably true that this fundamental shift towards shared responsibility ‘demonstrably goes beyond the intentions of those who devised it’,²⁹ but the law was changed fifty years ago to the year with no move by Parliament to reverse it.

Not only does this shift towards a nuanced approach to shared responsibility make the *Pritchard* interpretation, with its narrow focus on prior case law, seem outdated, but it also outmodes many of the normative arguments against applying contributory negligence to cases of intentional trespass to the person. This can be seen by re-evaluating some of the normative arguments that Aikens LJ uses in his judgment to justify excluding the defence. The principal normative argument was drawn from Professor Glanville Williams and will be addressed directly. Williams has argued that the exclusion of the defence is a ‘penal provision aimed at repressing conduct flagrantly wrongful.’³⁰ It is also, he says, the result of ‘ordinary human feeling that the defendant’s wrongful intention so outweighs the [claimant’s] wrongful negligence as to efface it altogether.’³¹ Whilst these are two distinct arguments they both appear to be premised on the idea that the presence of intention makes these torts sufficiently different from the general tort of negligence to warrant excluding the defence.

²¹ Steele (n 15).

²² *Pritchard* (n 3) 343.

²³ *Standard Chartered Bank* (n 12) 975.

²⁴ [2001] EWCA Civ 1249, [2002] 1 WLR 218.

²⁵ *ibid* 229.

²⁶ Steele (n 20) 162.

²⁷ Lord Wright, ‘Contributory Negligence’ (1950) 13 MLR 2, 10.

²⁸ *ibid*.

²⁹ Steele (n 20) 161.

³⁰ Glanville Williams, *Joint Torts and Contributory Negligence* (Stevens 1951) 198, cited in *Pritchard* (n 3) 332.

³¹ *ibid*.

Admittedly this conception of intentional torts does indeed appeal to ‘ordinary human feeling.’ Generally both the layman and the lawyer would regard something done intentionally as a greater wrong than something done carelessly or by pure accident, at least at first glance. Indeed the criminal law’s hierarchical approach to *mens rea* reflects this. However, as Dr Goudkamp has argued, such an approach fails to take into account the actual legal nature of these intentional torts.³² The tort of battery, for instance, does not require an intention to injure, merely an intention to trespass, i.e. to make physical contact.³³ Thus battery ‘may range from innocuous physical contact to brutal beatings, rape, and murder.’³⁴ This makes it very difficult to argue that, in the case of tort liability, the presence of intention necessarily reflects a certain degree of wrongdoing. It also undermines the argument that excluding the defence of contributory negligence is an effective punitive measure as this could lead to rather disproportionate results. Take, for example, the facts of *Pritchard* itself. The claimant in that case verbally harassed the store manager *en masse* with her sister and friend and then, after the battery occurred against her (which amounted to the store manager holding her arms out in front of her), proceeded to bite the manager.

It seems very unrealistic to say that excluding the defence of contributory negligence provides a useful punitive function here: just as it may appeal to ‘ordinary human feeling’ that intention connotes a greater degree of wrongdoing so too does it appeal to ordinary feeling that both parties are at fault in this kind of case, with the potential for the claimant to be more at fault than the defendant. Whilst it is possible to argue that the seriousness of the physical contact is irrelevant because the presence of intention automatically alters the moral position of the wrongdoer vis-à-vis the victim, such arguments, setting aside the desirability or otherwise of incorporating morality into the law, seem to stray too far into issues generally left to the criminal law. The change of position, if any, affected by the presence of intention plays a much larger role in criminal responsibility (e.g., the ‘thin-skull’ rule) than tort law, which is focused more on issues of fault and responsibility for loss rather than moral culpability. It is suggested that it is more due to the fact that the courts were forced to choose between two extremes, i.e. a full award of damages or none at all, prior to apportionment that gave the impression that intentional torts were in some way conceptually distinct from other torts to warrant the exclusion of the defence. Now that the 1945 Act has been passed and a choice between two extremes is no longer necessary it is easier to see that this is not the case and that there is a fairer alternative available.

If there is nothing conceptually distinct about intentional torts to warrant excluding the defence of contributory negligence out of hand, then, with the introduction of apportionment, it seems a much fairer approach to allow the defence to be applied. Where both parties are at fault for the claimant’s loss then damages are apportioned accordingly. There appears little justification, now that contributory negligence is no longer a complete defence, to not apply this approach just because intentional trespass to the person is involved. The criminal law allows even the crime of murder to be relegated to one of manslaughter where the defendant lost control due, *inter alia*, to things the victim said or did.³⁵ Whilst tort law is not interested in the characteristics of the individual parties or with human frailty it is interested in regulating the private interference with rights, and it does so by ascribing fault. The claimant who provokes the defendant with the intention of starting a fight can be said to be at fault, in other words to have failed to take reasonable care for his or her own safety per the proposed interpretation of section 4 above. That the defendant then commits the tort of battery, which as mentioned can cover an incredibly varied range of conduct, does not justify the complete exclusion of the claimant’s actions from consideration.

³² James Goudkamp, ‘Contributory Negligence and Trespass to the Person’ (2011) 127 LQR 519.

³³ *Wilson v Pringle* [1987] QB 237 (CA), 249 (Croom-Johnson LJ).

³⁴ Goudkamp (n 32) 520.

³⁵ Coroners and Justice Act 2009, ss 54-55.

Prior to *Pritchard* the courts seemed to be moving towards this approach. In *Murphy v Culbane*³⁶ the defendant had assaulted and beat the deceased by hitting him on the head with a plank. It was alleged that this assault occurred during a criminal affray that the deceased and others had initiated with the joint intent of harming the defendant. Lord Denning MR said, *obiter*, that had the deceased's widow been entitled to damages they might fall to be reduced under the 1945 Act because 'the death of her husband might be the result partly of his own fault and partly of the defendant' within the meaning of sections 1 and 4.³⁷ In *Barnes v Nayer*³⁸ two families had for a prolonged period of time subjected each other to serious abuse. The two families one day goaded their respective sons into a fight. The defendant, having been so provoked and having also been assaulted himself, went on to kill the deceased. May LJ said he saw, *prima facie*, 'no reason why, ... given the facts, a defendant to a claim for damages for assault cannot rely upon the Law Reform (Contributory Negligence) Act [1945].'³⁹ These cases illustrate the capacity for contributory negligence to provide for a more just reflection of blame, in line with the idea of shared responsibility, in cases of intentional trespass to the person.

There are a number of additional concerns that still need to be addressed, however. First, is the argument that it is not the presence of intention in intentional torts that justifies the exclusion of the defence but rather the concept, expressed in, among other cases, *Collins v Wilcock*,⁴⁰ that 'every person's body is inviolate'.⁴¹ However, if it were this doctrine, rather than the presence of intention, that precludes the application of contributory negligence then this would seem to be inconsistent with allowing the defence to apply in cases of ordinary negligence. By the very nature of personal injury claims a person's bodily integrity has been violated and yet prior to apportionment the courts were even willing to absolve the defendant of *all* liability in negligence cases where the defence applied. If anything the right to complete bodily integrity explains the vast range of physical conduct that can constitute battery but it does not follow that the doctrine precludes the sharing of responsibility for the infringement of that right. This again seemingly leaves the presence of intention as the underlying rationale behind distinguishing negligence from trespass to the person for the purposes of contributory negligence but as above this does not seem a convincing distinction, especially considering negligence, strictly speaking, can include intentional conduct anyway.

The second additional concern is that, even if it is desirable to allow contributory negligence to apply to cases of intentional trespass to the person, it is a matter best left for Parliament not for the judiciary. The reason why this argument is unconvincing is two-fold. Primarily, as explained earlier in relation to the dual limb approach to section 4, it is perfectly possible to read the definition of fault as accommodating cases of intentional trespass to the person. No more violence is done to the language of the statute on this approach than the alternative approach taken in *Pritchard*. On this view the courts would merely be applying the statute and, crucially, in a way that is a natural development of a general trend. This is the second reason. As argued above, the 1945 Act can be seen as part of a wider rethinking of the concept of fault in tort law, this rethinking being founded on the language of shared responsibility. Applying contributory negligence to cases of intentional trespass to the person would not constitute an exercise in unrestrained and spontaneous judicial creativity but would rather be a natural development in line with this trend, allowing a more nuanced – and consequently fairer – approach to cases of intentional trespass to the person.

³⁶ [1977] 1 QB 94 (CA).

³⁷ *ibid* 99.

³⁸ [1986] CA Transcript No 1085.

³⁹ *ibid* 6.

⁴⁰ [1984] 1 WLR 1172 (QB).

⁴¹ *ibid* 1178 (Goff LJ).

VI. CONCLUSION

Whilst at first the legal reasoning employed in *Pritchard* seems sound, and has previously gone largely unchallenged, an attempt has been made to show that the foundations of the decision are not as solid as they appear. Ultimately *Pritchard* represents a reading of the 1945 Act that arbitrarily restricts the operation of the contributory negligence defence to the law as it was over half a century ago. However, the law has moved on significantly since then with the introduction of apportionment, as part of a wider shift in the law of tort towards shared responsibility for loss. It is better to adopt an approach to the section 4 definition of fault that allows for this change to be reflected in a natural development of the law, namely by reading into the Act solely the requirement that the actions of the claimant constitute a failure to take reasonable care for his or her own safety, as opposed to also arbitrarily restricting its scope to pre-1945 situations. This is indeed the only interpretation of section 4 that is consistent with the facts of *Reeves* and is at least capable of accommodating cases such as *Standard Chartered Bank*. Neither does such an interpretation necessarily have to be at odds with Parliamentary intention if it is accepted that Parliament must have been open to the idea of contributory negligence developing beyond its prior common law scope. Regardless, of greater importance is that the law of tort should be allowed to continue this trend towards shared responsibility, and thus more just and equitable awards of damages, by allowing contributory negligence to apply in cases of intentional trespass to the person.

Finding Principle in Illegality: Reflections on *Tinsley v Milligan*

Matthew Chan¹

I. INTRODUCTION

In a lecture to the Chancery Bar Association in 2012,² Lord Sumption stated that ‘the law of illegality is an area in which there are few propositions, however contradictory or counter-intuitive, that cannot be supported by respectable authorities at the highest levels.’³ The prescience of his statement was borne out by a series of recent decisions handed down by the Supreme Court, revealing various fundamental differences within the judiciary as to how the doctrine of *ex turpi causa*, or the law pertaining to the ‘illegality defence’, should be applied in practice. Accordingly, there is considerable judicial uncertainty as to the residual significance, if any, of the House of Lords decision in *Tinsley v Milligan*.⁴ Most recently, Lord Sumption himself held in *Jetivia SA v Biltta (UK) Ltd*⁵ that the ‘reliance test’ expounded in *Tinsley v Milligan* epitomised the inflexible approach that should generally be taken with respect to the illegality defence. In contrast, Lords Toulson and Hodge downplayed the significance of *Tinsley*, saying that the application of the illegality defence was not based on a rigid application of the *Tinsley* test, but that it was necessary to consider the policy underlying the illegality defence in order to decide whether it should defeat a given claim.

Indubitably then, the *Tinsley* test is at once central to a discussion of the illegality doctrine, and productive of many conceptual and practical difficulties concerning the same. This article seeks to do three things. Firstly, having examined the House of Lords judgment in *Tinsley*, it will be argued that its direct precedential reach is considerably attenuated in areas of the law not directly falling within its factual scope. Secondly, it will be considered if the retention of the *Tinsley* test in areas where it necessarily remains binding authority is, on balance, normatively justified. Finally, having argued that it is not, the possibility and implications of reform in the area will be discussed.

II. TINSLEY V MILLIGAN

The essence of the doctrine of *ex turpi causa non oritur actio* is perhaps best encapsulated by Lord Mansfield CJ in *Holman v Johnson*:⁶

No court will lend its aid to a man who founds his cause of action upon an immoral or an illegal act. If, from the plaintiff's own standing or otherwise, the cause of action appears to arise *ex turpi causa*, or the transgression of a positive law of this country, there the court says he has no right to be assisted.

While Lord Mansfield thus plainly presented the doctrine to be simply a public policy of judicial abstention in the presence of illegality, the policy effectively, albeit incidentally,⁷ operates as a defence insofar as a defendant may plead that, notwithstanding any substantive merit to the claimant's claim, the claimant should by reason of his own illegality be unable to succeed in his claim.

¹ Exeter College, Oxford.

² Lord Sumption, ‘Reflections on the law of illegality’ (Chancery Bar Association, London, 23 April 2012). The lecture was subsequently published in [2012] RLR 1.

³ Lord Sumption [2012] RLR 1, 1.

⁴ [1993] UKHL 3, [1994] 1 A.C. 340.

⁵ [2015] UKSC 23, [2015] 2 W.L.R. 1168.

⁶ (1775) 1 Cowp 341 (KB) 343.

⁷ See, e.g., the decision of the Singapore High Court in *ANC Holdings Pte Ltd v Bina Puri Holdings Bhd* [2013] SGHC 97 [78].

Although Lord Mansfield's formulation of *ex turpi causa* was stated simply and categorically, the twentieth century saw a mass of litigation that obfuscated the application of the doctrine. In a bid to explicate and resolve this issue of conflicting authority, the Court of Appeal sought to treat the whole body of authority as having ensued from an essentially discretionary process. Thus in *Euro-Diam Ltd v Bathurst*,⁸ the illegality defence was held to apply when, in all the circumstances, it would be an 'affront to the public conscience' to allow the plaintiff's claim.⁹ It was in this context that *Tinsley v Milligan* came before the House of Lords. Ms Tinsley and Ms Milligan contributed to the purchase of a home together, but had the legal title conveyed to Ms Tinsley alone, so that Ms Milligan could make fraudulent claims to social security benefits. After the parties fell out with each other, Ms Milligan sought a declaration that the property was held by Ms Tinsley on trust for both parties. Having applied the 'public conscience' test as set out above,¹⁰ the Court of Appeal found in favour of Ms Tinsley. However, the House of Lords unanimously rejected the public conscience test. Lord Goff noted that:¹¹

the adoption of the public conscience test... would constitute a revolution in this branch of the law, under which what is in effect a discretion would become vested in the court to deal with the matter by the process of a balancing operation, in place of a system of rules, ultimately derived from the principle of public policy enunciated by Lord Mansfield CJ in *Holman v Johnson*...

Having sounded the death knell for the 'public conscience' test and any substantial role of discretion in the application of the illegality defence, the court was then divided as to what its replacement was, or should be. The majority favoured a 'reliance test', whereby a party to an illegality could recover by virtue of a legal or equitable property interest he could establish his title with *relying* on his own illegality.¹² It had been established in *Bowmakers v Barnet Instruments*¹³ that a plaintiff could enforce *legal* property rights provided that he did not need to rely on an illegal contract for any purpose other than to provide the basis of his claim to a property right;¹⁴ thus '[i]f the law is that a party is entitled to enforce a property right acquired under an illegal transaction... the same rule ought to apply to any property right so acquired, whether such right is legal or equitable'.¹⁵ Applying the reliance test to the facts, the majority thus found in favour of Ms Milligan. Given her financial contribution to the purchase of the house, a presumption of resulting trust arose which was not rebutted; thus Ms Tinsley was held to be holding the house on trust for both parties in equal shares.

The minority found instead for Ms Tinsley. Lord Goff held that the 'so-called *Bowmakers* rule' did not apply in the present case, (*inter alia*) because a claimant who had not come to a court of equity with 'clean hands' could not obtain the assistance of the court, even if he or she could *prima facie* establish a claim without recourse to the underlying fraudulent or illegal purpose.¹⁶ Thus, having argued that the 'clean hands' maxim was 'more broadly based' than the rule laid down in *Bowmakers*, Lord Goff (with whom Lord Keith agreed) held that Ms Milligan's claim must fail given that her claim was tainted by virtue of her illegal agreement with Ms Tinsley.¹⁷

⁸ [1990] 1 QB 1 (CA).

⁹ *ibid* 35.

¹⁰ [1992] Ch 310 (CA) 321.

¹¹ *Tinsley* (n 4) 363.

¹² *ibid* 375 (Lord Browne-Wilkinson).

¹³ [1945] KB 65 (CA).

¹⁴ *Tinsley* (n 4) 370.

¹⁵ *ibid* 371.

¹⁶ *ibid* 358.

¹⁷ *ibid* 362.

It may thus be noted that the ratio in *Tinsley* comprised two distinct elements. Firstly, the House of Lords (unanimously) *rejected* the public conscience test as it had been applied by the Court of Appeal in cases such as *Euro-Diam* and *Tinsley* itself. Secondly, the House of Lords (by a bare majority) *formulated* a ‘reliance principle’ as applying at equity as well as in law.

The distinction between the two is, for present purposes, analytically important. As will be evident, it is not seriously doubted by anyone that the ‘reliance test’ *per se* is *not* a universal test for determining whether the illegality defence applies. On the other hand, doubt persists amongst the judiciary as to whether the rejection of the public conscience test necessarily entails that a court is to have no discretion in applying the illegality defence, and whether this is to be the case regardless of the context in which the claim is brought (e.g. in contract, tort, or trusts).

This doubt was manifest in the judgment of *Jetivia*. The directors of Bilta (UK) Ltd had caused it to enter into a series of carousel frauds with various parties, including Jetivia SA, between April and July 2009. After Bilta was compulsorily wound up in November 2009 pursuant to a petition presented by HMRC, its liquidators brought proceedings against its directors and Jetivia, claiming that the parties were parties to an unlawful means conspiracy to injure Bilta by a fraudulent scheme. Jetivia argued, *inter alia*, that the illegality defence applied on the facts to defeat Bilta’s claim.

Although the Supreme Court was unanimous in dismissing Jetivia’s appeal and held that the illegality defence did not apply *on the facts* because the wrongful activity of Bilta’s directors simply could not be attributed to Bilta in the proceedings, this result belied the considerable difference in opinion as to the *general* basis of the illegality defence. Lord Sumption held that the illegality defence was based on a rule of law which the court was required to apply if and only if it applied. As had recently been confirmed by the Supreme Court in *Les Laboratoires Servier v Apotex*,¹⁸ the application of the defence was not a discretionary power, nor was it ‘dependent upon a judicial value judgment about the balance of the equities in each case’.¹⁹ He concluded that in this sense, the House of Lords decision in *Tinsley* remained good law. In contrast, Lords Toulson and Hodge felt that the applicability of the illegality defence was not based on a rigid application of the *Tinsley* ‘reliance test’; in fact, it was not based on a rigid approach at all. Citing Lord Wilson’s judgment in *Hounga v Allen*²⁰ with approval, they held that it was necessary to consider the policy underlying the illegality defence in order to decide whether it should defeat the claim at hand.²¹ The thorniness of the issue is best illustrated by the reluctance of the majority to comment substantively on the basis of the defence. Lord Neuberger (with whom Lords Clarke and Carnwath agreed) noted that the proper approach which should be adopted to a defence of illegality was a ‘difficult and important’ topic,²² but refrained from discussing it further: there had been no real argument in the topic, and the issue of what approach to take to illegality was not determinative of the outcome.²³ Lord Mance exhibited similar restraint, remarking only the need for ‘further examination’ of the issue if fuller argument was provided in future cases.²⁴

¹⁸ [2014] UKSC 55, [2015] A.C. 430.

¹⁹ *Jetivia* (n 5) [62].

²⁰ [2014] UKSC 47, [2014] 1 W.L.R. 2889.

²¹ *Jetivia* (n 5) [171].

²² *ibid* [13].

²³ *ibid* [15].

²⁴ *ibid* [34].

III. AN EXAMINATION OF THE CASE LAW

It is evident from the foregoing discussion that despite recent and copious Supreme Court debate on the basis of the illegality defence, the issue remains frustratingly unresolved. An examination of the case law that pertains to the illegality defence will reveal that – perhaps unsurprisingly – the ‘reliance test’ has minimal application in areas of law other than trusts. It will further be submitted that although *Tinsley* has been interpreted in subsequent cases to stand for a *general* proposition that judges are not to apply discretion in considering the illegality defence, as a matter of positive law, the account of Lords Toulson and Hodge in *Jetivia* is a better description of the law on illegality than that of Lord Sumption, inasmuch as courts have habitually considered policies underlying the illegality defence to determine if it should apply in each case, even if they have at times ostensibly accepted the authority of *Tinsley*.

A. THE RELIANCE TEST

In the law of tort, the courts have largely departed from the ‘reliance test’ in deciding if illegality applies on the facts. Notably, the ‘inextricable link’ test in *Cross v Kirkby*²⁵ is commonly cited as an alternative to the *Tinsley* test in the tortious context.²⁶ The parties in *Cross* had gotten into an altercation initiated by the claimant, who eventually suffered a skull fracture after being hit on the head with a baseball bat by the defendant. The Court of Appeal held that the claimant’s claim in assault and battery failed because it was defeated by the illegality defence; where the claimant was behaving unlawfully or criminally, his claim was liable to be defeated *ex turpi causa* if it was established that the facts which gave rise to it were ‘inextricably linked’ with his criminal conduct.²⁷

Subsequently, the test was refined, albeit materially retained, by the House of Lords in *Gray v Thames Trains*²⁸ in the context of negligence. The claimant had suffered post-traumatic stress disorder (PTSD) after being injured in a major railway accident caused by the defendant’s negligence. He later killed a man and was convicted of manslaughter by reason of diminished responsibility. It was not held to be in doubt that he would not have committed the offence but for the PTSD. He sought damages from the defendant, *inter alia* for loss of earnings and loss of liberty. His action failed before the Lords. Lord Hoffmann stated that the illegality defence had a narrower and a wider manifestation. In its narrower form, a civil court would not award damages to compensate a claimant for the injury or disadvantage which a criminal court had imposed on him by way of punishment for a criminal act. In its wider form, it held that one could not recover for damages in respect of the consequences of one’s own criminal act.²⁹ In considering when the wider form of the defence was to apply, Lord Hoffmann considered variants of the test espoused in *Cross*, but thought that ‘metaphors’ such as ‘inextricably linked’ were ‘unhelpful’.³⁰ He then concluded that the ordinary test of causation should be adopted when determining whether the illegality defence should apply.³¹

This distinction, between causing something and merely providing the occasion for someone else to cause something, is one with which we are very familiar with in the law of torts... Can one say that, although the damage would not have happened but for the tortious conduct of the defendant, it was caused by the criminal act of the Claimant? [...] Or is the position that although the damage would not have happened without the criminal act of the Claimant, it was caused by the tortious act of the Defendant?

²⁵ [2000] EWCA Civ 426.

²⁶ See eg. *Hounga* (n 20) [31].

²⁷ *Cross* (n 25) [103] (Judge LJ); see also [76] (Beldam LJ).

²⁸ [2009] UKHL 33, [2009] 1 A.C. 1339.

²⁹ *ibid* [51].

³⁰ *ibid* [54].

³¹ *ibid*.

As it stands, the causation approach generally remains good law in areas of tortious liability. Certainly, lower courts have regarded it as the definitive approach in negligence cases.³² Accordingly, the reliance test is not generally regarded as relevant in tort. Although Lord Hoffmann had considered arguments relating to *Tinsley* in *Gray*, he shortly dismissed its relevance to the facts: '[t]he questions of fairness and policy are different and the content of the rule is different. One cannot simply extrapolate rules applicable to a different kind of situation'.³³

A more explicit qualification of the ambit of the reliance test may be found in *Stone & Rolls Ltd v Moore Stephens*.³⁴ *Stone & Rolls Ltd* had been wholly owned by its director, who fraudulently siphoned company assets away and falsified its accounts. Following its liquidation, its creditors, acting in the name of the company, sued the company auditors for failing to detect the fraud. The House of Lords, by a bare majority, held that the auditors could not be sued by the company's liquidator. Although the complexities of the case are substantial and cannot be wholly addressed in this article, it suffices for present purposes to note that the House of Lords narrowly construed the ambit of the *Tinsley* reliance test:³⁵

The House in *Tinsley v Milligan* did not lay down a universal test of *ex turpi causa*. It was dealing with the effect on illegality on title to property... The House did not hold that illegality will never bar a claim if the claim can be advanced without reliance on it. On the contrary, the House made it plain that where the claim is to enforce a contract the claim will be defeated if the defendant shows that the contract was for an illegal purpose, even though the claimant does not assert the illegal purpose in making the claim...

It is submitted that the House of Lords' decision to limit the ambit of the reliance test in *Gray* and *Stone & Rolls* can be defended both doctrinally and practically. Firstly, purely as a matter of doctrine, the majority decision in *Tinsley* was squarely justified in terms of the passing of *proprietary* rights. Lord Browne-Wilkinson sought to elide the position on equitable proprietary rights with that on legal proprietary rights; in doing so, he did not explicitly purport to be laying down a rule that would cover situations where merely *personal* rights were involved.

Secondly, it is from a practical viewpoint clear that the language of 'reliance' is inapt and unhelpful in cases where personal rights are involved. By way of illustration, a court might reasonably hold that the claimant in *Cross* had to 'rely' merely on the fact that the defendant had assaulted him with a baseball bat and thus award him damages in accordance with ordinary tort principles; but equally, given that this latter assault had been provoked by the claimant himself, the court could find that the claimant had to 'rely' on his own illegal conduct for the purpose of establishing his claim, and thus deny him relief. This ambiguity demands that the court reach its conclusion on factors extrinsic to amorphous notions of 'reliance'. As such, it is highly questionable whether the latter can ever be of much discriminating value in cases where proprietary rights are not concerned.

With that having been said, recognising that the reliance test is not (and should not be) universally applicable in respect of the illegality defence does not begin to resolve the fundamental disagreement amongst the bench in *Jetivia*. Lord Sumption did not dispute that the reliance test was so limited in application. For instance, it is evident from his judgment in *Apotex* that he recognises that the *Tinsley* reliance test is not applicable in tort.³⁶ However, what he contends, in contradistinction to Lords Toulson and Hodge, is that *Tinsley* conclusively rejected the existence of any discretionary element in courts' invocation of the illegality defence. It is thus to this element of *Tinsley* that we must turn to now.

³² See *Delaney v Pickett* [2011] EWCA Civ 1532, [2012] 1 W.L.R. 2149; *Joyce v O'Brien* [2013] EWCA Civ 546, [2014] 1 W.L.R. 70.

³³ *Gray* (n 28) [31].

³⁴ [2009] UKHL 39, [2009] 1 A.C. 1391.

³⁵ *ibid* [21].

³⁶ *Apotex* (n 18) [19].

B. THE ROLE OF DISCRETION

In *Jetivia*, Lord Sumption referred to *Apotex* in support of his contention that the illegality defence is ‘based on a rule of law... not a discretionary power on which the court is merely entitled to act’.³⁷ In *Apotex*, the appellants, Servier, held patents in UK and Canadian law for a drug called perindopril erbumine. The respondents, Apotex, were a Canadian group that began to import and sell generic perindopril erbumine tablets in the UK. Servier obtained an interim injunction against Apotex by giving a cross-undertaking for damages, promising to compensate the latter for any loss caused by the injunction if it later turned out that it should not have been granted. Apotex became entitled to compensation when the court found the UK patent to be invalid. Servier challenged the award for damages on the basis of illegality, arguing that it was contrary to public policy for Apotex to recover damages when the manufacture of the product in Canada would have been unlawful in infringing their Canadian patent.

The Court of Appeal held³⁸ that the infringement of Servier’s Canadian patent was not a relevant illegality for the purposes of the defence, because in dealing with the illegality defence, the court was entitled ‘to take into account a *wide range of considerations* in order to ensure that the defence only applies where it is a just and proportionate response to the illegality involved in the light of the *policy considerations* underlying it’³⁹ (italics added). The Supreme Court reached the same result but, by a majority, rejected the reasoning of the lower court. Crucially, Lord Sumption, with whom Lord Neuberger and Lord Clarke agreed, held that the Court of Appeal had been wrong to treat the question as depending on the culpability of the illegality, the proportionality of the application of the defence or the general merits of the particular case.⁴⁰ As in his judgment in *Jetivia*, he pointed out that the House of Lords decision in *Tinsley* had rejected the ‘public conscience’ test with respect to the illegality defence on the ground that it imported a discretionary element into what was in reality a rule of law;⁴¹ thus, the Court of Appeal decision was inconsistent with *Tinsley*. In the minority, Lord Toulson also dismissed the appeal but dissented from the majority reasoning. Rather, he agreed with the Court of Appeal that public considerations *should* be taken into account in the determination of the illegality defence, given that the defence is based on public policy.⁴²

While it is thus largely unarguable that the majority reasoning in *Apotex* affirms the relevance of *Tinsley*, *Hounga v Allen* appears to be a considerable obstacle to Lord Sumption’s position in *Jetivia*. The respondents offered to employ the claimant as a home help in the UK in return for schooling and £50 per month, and they helped her obtain false identity documents with which she entered the UK and obtained a six-month visitor’s visa. Later the respondents evicted the claimant from the house, dismissing her from employment. The claimant issued proceedings for unlawful discrimination in relation to her dismissal.⁴³ In allowing the claimant’s appeal, the Supreme Court held that the illegality defence did not defeat the complaint of discrimination. Lord Wilson, delivering the lead judgment, held that:⁴⁴

[t]he defence of illegality rests upon the foundation of public policy [...] So it is necessary, first, to ask “What is the aspect of public policy which founds the defence?” and second, to ask “But is there another aspect of public policy to which application of the defence would run counter?”

³⁷ *Jetivia* (n 5) [62].

³⁸ [2012] EWCA Civ 593.

³⁹ *ibid* [73] (Etherton LJ).

⁴⁰ *Apotex* (n 18) [19].

⁴¹ *ibid* [18].

⁴² *ibid* [62].

⁴³ Contrary to the Race Relations Act 1976.

⁴⁴ *Hounga* (n 20) [42].

On the facts, the illegality defence was held to be inapplicable because while there were scarcely any public policy reasons why the defence should be applied to defeat the claim, to *reject* the defence's application would serve to uphold a governmental policy to combat human trafficking and protecting its victims.⁴⁵

Prima facie, *Hounga* appears to be starkly inconsistent with Lord Sumption's assertion that courts generally eschew a 'discretionary weighing of the equities'⁴⁶ in applying the illegality defence, in accordance with *Tinsley*. Indeed, while Lord Wilson explicitly acknowledges the *Tinsley* test,⁴⁷ he then found⁴⁸ that the effect of the reliance test was 'soften[ed]' by the House of Lords decision in *Stone & Rolls*. His subsequent approach of considering the public policy factors for and against applying the illegality defence in the present case thus viably supports the contention of Lords Toulson and Hodge that the application of the illegality defence depended on a consideration of the policy factors underlying it.

Indeed, in addressing *Hounga*, Lord Sumption in *Jetivia* had to concede that there were exceptions to the applicability of *Tinsley*; in some cases 'an examination of competing policies may be required, and that is where a competing public policy... requires the imposition of civil liability notwithstanding that the claim is founded on illegal acts'.⁴⁹ Thus, he said that '[t]he court [in *Hounga*] was not purporting to depart from *Tinsley v Milligan* without saying so. It simply recognised the case before it in which a competing public policy required that damages should be available even to a person who was privy to her own trafficking'.⁵⁰ Accordingly, he felt sceptical about the significance of *Hounga* as a statement of principle of general application.⁵¹

In response to Lord Sumption's explanation of *Hounga*, two things may be remarked. Firstly, by conceding that there could be 'public policy' exceptions that displace a strict application of the *Tinsley* approach, Lord Sumption's view becomes considerably more difficult to distinguish from that of Lords Toulson and Hodge. After all, to say that in certain (admittedly exceptional) cases countervailing policy considerations may cause the illegality defence to be disapplied is to presuppose that the courts must in such cases consider and weigh the policies *for* and *against* the applicability of the defence; this must be true if not all 'competing public polic[ies]' are to be allowed to trump the policy underlying the defence. It is clear that this is a short way from holding that public policy considerations are *generally* to be considered when deciding on illegality.

Secondly, it is questionable if *Hounga* is as exceptional as Lord Sumption represented it to be. In particular, it is argued that in some cases, despite ostensible adherence to the no-discretion principle in *Tinsley*, courts in fact engage in a discretionary 'balancing of equities' exercise in deciding whether or not the illegality defence should apply.

An eminent example is the approach set out by Lord Hoffmann in *Gray*. In *Apotex*, Lord Sumption held that neither the 'narrower' nor the 'wider' limbs of the illegality defence in *Gray* depended on the court's assessment of the significance of the illegality, the proportionality of its application or the merits of the case; the narrow rule 'operated automatically' while the wider rule 'was simply a question of causation'.⁵² With respect, however, he was being overly sanguine about the

⁴⁵ *Hounga* (n 20) [45], [52].

⁴⁶ *Jetivia* (n 5) [99].

⁴⁷ *Hounga* (n 20) [28].

⁴⁸ *ibid* [30].

⁴⁹ *Jetivia* (n 5) [101].

⁵⁰ *ibid* [102].

⁵¹ *ibid*.

⁵² *Apotex* (n 18) [19].

implications of the causation approach. It is trite that the concept of causation in tort is profoundly knotty. Policy considerations have on multiple occasions come to the fore in courts' attempts to refine the concept of factual causation in certain exceptional contexts,⁵³ whereas the concept of legal causation is by definition premised on the interplay between policy factors adjudged to be relevant in determining if liability should arise in a given case.⁵⁴ As Lord Mance remarked extra-judicially on the causation approach adopted in *Gray*, the distinction between 'causing' and 'occasioning' is capable of being inconsistently applied; it is 'not a matter of mathematics, but ultimately of judgment by the court as to the relative weight which ought to be attached to the different events, and in that sense one of policy'.⁵⁵

Unsurprisingly, the problems with the causation approach have not been lost on the lower courts. The Court of Appeal judgment in *McCracken v Smith*⁵⁶ (incidentally delivered on the same day as the Supreme Court delivered theirs in *Jetivia*) presents an illuminating example. The respondent suffered a brain injury while riding pillion on a trial bike that collided with a minibus, driven by the appellant. The rider, his friend, did not have a valid driving licence or insurance, and neither party was wearing a helmet. At first instance,⁵⁷ it was held that both the friend and the appellant had been negligent and were liable to the respondent in damages, albeit lessened to the extent of the respondent's contributory negligence. The Court of Appeal considered an appeal by the appellant that the judge at first instance had wrongly declined to apply the illegality defence with respect to the respondent's claim against him, given that the respondent had engaged in a joint enterprise with his friend to ride the bike dangerously. Delivering the leading judgment, Richards LJ professed difficulties with applying the causation approach:⁵⁸

[T]he situation cannot be accommodated neatly within the binary approach of Lord Hoffmann in *Gray*... The accident had two causes, properly so called – the dangerous driving of the bike and the negligent driving of the minibus – and it would be wrong to treat one as the mere “occasion” and the other as the true “cause”. [The respondent's] injury was the consequence of both, not just of his own criminal conduct and not just of [the appellant's] negligence.

Having concluded that the fact that the criminal conduct was one of the two causes was not a sufficient basis for the illegality defence to apply, Richards LJ held that the right approach was to reject the illegality defence but to reduce the respondent's damages in accordance with principles of contributory negligence,⁵⁹ an approach that served the public interest insofar as it accounted for the fault of both the appellant and the respondent.⁶⁰

McCracken thus gives the lie to Lord Sumption's pronouncement on the causation approach. In holding both that there were two 'causes' of the injury and that the illegality defence should not apply, Richards LJ clearly took into account issues such as the significance of the illegality, proportionality and the merits of the case. As he pointed out, any broad test of causation was almost by definition satisfied on the facts;⁶¹ it would thus have been difficult for him to mechanically apply the causation approach to reach a decision on the illegality defence *without* any reference to merits.

⁵³ See, e.g., *Fairchild v Glenhaven Funeral Services Ltd* [2002] UKHL 22, [2003] 1 A.C. 32; *Sienkiewicz v Greif* [2011] UKSC 10, [2011] 2 A.C. 229.

⁵⁴ See, e.g., *The Wagon Mound (No 2)* [1967] 1 AC 617 (PC).

⁵⁵ Lord Mance, 'Ex turpi causa – when Latin avoids liability' [2014] *Edinburgh Law Review* 175, 184.

⁵⁶ [2015] EWCA Civ 380.

⁵⁷ [2013] EWHC 3620 (QB) (Keith J).

⁵⁸ *McCracken* (n 56) [51].

⁵⁹ *ibid* [52].

⁶⁰ *ibid* [53].

⁶¹ *McCracken* (n 56) [54].

In contract, the courts have evinced a similar willingness to look at policy factors in applying the illegality defence. Notably, despite ostensible reference to the *Tinsley* test, the Court of Appeal in *Parkingeye v Somerfield Stores*⁶² also tied their decision to a consideration of policies that underlay the illegality defence. The respondent operated an automated car parking system in car parks operated by the appellant. When the appellant terminated their contract prematurely, the respondent sued for damages for breach of contract. The appellant claimed that the contract was void for illegality because of the unlawful means the respondent had used to collect some of the fines. In finding that the illegality defence should not apply, Sir Robin Jacob, with whom Laws LJ agreed, found it significant that the ‘facts of the case, considered with a sense of proportionality, [did not] involve such an invasion of any of the policy rationales as to deprive ParkingEye of its remedy’.⁶³ Interestingly, Sir Robin had no difficulty with concluding that a proportionality-based consideration of relevant policy concerns was wholly consistent with the authority of *Tinsley*.⁶⁴

In applying the “disproportionate” test I do not think I am exercising a judicial discretion. It was settled by *Tinsley v Milligan* that a defence of illegality point cannot be solved by applying a discretion based on public conscience. Proportionality as I see it is something rather different. It involves the assessment of how far refusal of the remedy furthers one or more of the specific policies underlying the defence of illegality.

Even if we accept that Sir Robin was not reviving the ‘public conscience’ test that was rejected in *Tinsley*, it remains, with respect, difficult to see how he was not exercising a discretion. His proportionality approach entailed weighing policy considerations that supported applying the illegality defence against the considerations that militated against such application. Given that these considerations are hardly amenable to precise or fixed valuation, the court surely had to make value judgments as to whether and to what extent certain policies should be considered significant, which patently involves the exercise of judicial discretion. While it can at least be accepted that a consideration of proportionality differs *notionally* from a ‘public conscience’ test, it remains *practically* indistinguishable from the approach furthered by Lords Toulson and Hodge.

In conclusion, it is highly doubtful that *Tinsley* remains influential in the law on illegality, insofar as it appeared to reject the role of judicial discretion in determining if the illegality defence was to apply. *Hounga* is an obvious and prominent example of the courts explicitly reasoning by balancing conflicting policies; but *Gray* and *Parkingeye* are further instances of courts reasoning in a discretionary fashion, despite appearances to the contrary (e.g. the veneer of a hard-edged causation test, or an explicit admission of the binding authority of *Tinsley*). All that can be said with some certainty is that, as a matter of positive law, the *Tinsley* reliance test still applies to cases within the direct factual context of trusts. It remains to be discussed whether the reliance test is a normatively justifiable approach in areas where it *does* indubitably apply.

IV. TINSLEY: A CRITIQUE

A. CRITICISM

As is evident from a survey of the authorities, it is highly doubtful that *Tinsley* remains the test for illegality in all or even most circumstances. Crucially, however, aside from being dismissive of its relevance to the particular circumstances, courts have often been critical of the formulation of the reliance test itself.

⁶² [2012] EWCA Civ 1338, [2013] Q.B. 840.

⁶³ *ibid* [40].

⁶⁴ *ibid* [39].

By far the main criticism of the *Tinsley* reliance test has been the exclusive focus on the procedural issue of whether the claimant was required to plead his or her illegality. No heed is given to 'substantive' issues, such as whether and to what extent the claimant's claim is otherwise meritorious in any way, or whether the policies underlying the illegality defence would be furthered by its application in the given case. As the High Court of Australia observed in *Nelson v Nelson*:⁶⁵

Regard is had only to the procedural issue; and it is that issue and not the policy of the legislation or the merits of the parties which determines the outcome. Basing the grant of legal remedies on an essentially procedural criterion which has nothing to do with the equitable positions of the parties or the policy of the legislation is unsatisfactory, particularly when implementing a doctrine which is founded on public policy.

Given that the *Tinsley* test is intrinsically blind to the general 'merits' of the claim and thus in that sense operates 'arbitrarily', it has unsurprisingly drawn much flak for '[having] the potential to force the courts into unjust decisions'.⁶⁶ As we have already seen, the result in *Tinsley* itself hinged on the operation of a presumption inherent in trusts law that had nothing to do with the merits of the case.

Admittedly, it is at least arguable that applying the reliance principle in *Tinsley* gave rise to a fair result. Finding that the illegality defence operated to bar Ms Milligan's claim would have left Ms Tinsley with the full benefit of Ms Milligan's money although the two were equally culpable in entering into the illegal transaction. Indeed, Lord Goff, who would have defeated Ms Milligan's claim, confessed that his approach would have produced an unduly harsh result for her.⁶⁷ In the light of this, it might be argued that the majority achieved the 'fairer' result in resorting to the reliance principle. (Indeed, Lord Sumption has suggested extra-judicially that such principle was contrived precisely to avoid the 'distasteful' result that Lord Goff's approach would have achieved.⁶⁸) Nevertheless, the result as mentioned owed everything to the automatic operation of the presumption of resulting trust in that instant case; following the House of Lords' formulation of the reliance principle, there is nothing to prevent the principle being applied to reach unjust results. Indeed, Lord Browne-Wilkinson distinguished the facts of *Tinsley* from a case where the presumption of advancement would apply: on a transfer from a man to his wife or children, equity presumes an intention to make a gift. In such cases, then, the claimant would generally have to adduce evidence to rebut such presumption and in doing so would normally have to plead his underlying illegal purpose where it existed.⁶⁹

*Collier v Collier*⁷⁰ was just such a case where the presumption of advancement applied. Mr Collier put his night-club and recording studio premises in his daughter's name to hold for him and thus keep it from falling into the hands of his creditors threatening his bankruptcy. However, his daughter later refused to return it when he asked. Given that the presumption of advancement operated and Mr Collier was unable to show the real reason for the transfer of assets without pleading his own illegal purpose, his claim failed, and his daughter, a joint partner in his original illegal plan, gained a windfall. Whatever one might think of the merits of Mr Collier's claim, it is surely disturbing to consider that his claim would probably have succeeded *but for* the fact that he was the father of his partner in crime, thus triggering the presumption of advancement instead of the presumption of resulting trust. Indeed, in deciding the case in the Court of Appeal, Mance LJ (as he then was) was palpably dissatisfied with the fact that the application of the *Tinsley* test led to an opposite result to that in *Tinsley* itself, despite uncanny similarities on the facts:⁷¹

⁶⁵ (1995) 184 CLR 538, 609 (McHugh J).

⁶⁶ Law Commission, *The Illegality Defence* (Law Com No 320, 2010) para 2.15.

⁶⁷ *Tinsley* (n 4) 363.

⁶⁸ Lord Sumption (n 3) 4.

⁶⁹ *Tinsley* (n 4) 372.

⁷⁰ [2002] EWCA Civ 1095.

⁷¹ *ibid* [105].

[I]t can hardly be suggested that any real difference in opprobrium exists between the conduct of the successful Ms Milligan and the father in [...] the present case.

He however professed to being bound by *Tinsley* as a matter of precedent, although he suggested reform of the law in future.⁷²

Most tellingly, even though Lord Sumption has consistently posited that the *Tinsley* test generally remains the definitive approach to illegality as a matter of law, it should not be thought that he is unaware, or supportive, of the unideal formulation or effects of the test. In his lecture to the Chancery Bar, he dryly noted that following *Tinsley*, '[t]he test of relevance thus came to depend on the effect of presumptions devised by equity for a very different purpose, and on the incidence of the burden of proof.'⁷³ Having considered the application of the reliance test in *Collier*, a case where 'the moral equities were the same', he noted that '[i]n an area of law which turns on principle and public policy, this concentration on form over substance seems difficult to justify'.⁷⁴

B. IS TINSLEY DEFENSIBLE?

In the light of the ostensibly overwhelming criticism of the *Tinsley* test, stemming both from its 'merit-blind' nature and from its propensity to compel the courts to do injustice in certain circumstances, one is left wondering if *Tinsley* is at all normatively defensible. Two 'defences' of the *Tinsley* reliance test shall be examined here.

Firstly, it has been observed that the criticism of the decision in *Tinsley* has thus far been substantially based on the interplay between the presumption of resulting trust and the presumption of advancement, and the fact that their application had been premised on what is presently considered archaic and sexist assumptions (e.g. that a husband or father generally intends to make gifts to his wife or children, while a wife or mother lacks this intention). Given that the presumption of advancement has now been abolished by statute,⁷⁵ it might be argued that the criticism of *Tinsley* is now *pro tanto* less warranted. Indeed, in deciding not to implement the suggestions of the Law Commission to conduct statutory reform of the illegality defence as it applies in trusts law,⁷⁶ the Government stated that they were 'not satisfied that there is a sufficiently clear and pressing case for reform', citing the abolition of the presumption of advancement as a key reason for this view.⁷⁷ However, while criticism of the presumptions should indeed be analytically decoupled from criticism of the reliance test in general, it is surely erroneous to conclude that the presumption of advancement was the only, or even the predominant, problem with the reliance test. McHugh J's criticism of the test in *Nelson* was centred around the broader *irrelevance* of applicable public policy and the merits of the case, a malady of which the overriding *relevance* of presumptions is but a symptom. It is a *non sequitur* to hold that just because the existence of the presumption of advancement gives rise to particularly egregious applications of the reliance test, its abolition will completely or substantially resolve the deeper problems plaguing the test.

⁷² *Collier* (n 70) [106].

⁷³ Lord Sumption (n 3) 5.

⁷⁴ *ibid.*

⁷⁵ s 199(1) Equality Act 2010: 'The presumption of advancement (by which, for example, a husband is presumed to be making a gift to his wife if he transfers property to her, or purchases property in her name) is abolished.' However, it should be noted that the Equality Act 2010 has never been brought into force.

⁷⁶ See Section 5 below.

⁷⁷ Ministry of Justice, *Report on the implementation of Law Commission proposals* (2012) 14-15.

Secondly, as Lord Sumption pointed out in *Jetivia*, the purely procedural nature of the reliance test appears to be far more amenable to clear and certain application than its main rivals, insofar as it was capable of devising ‘principled answers which are no wider than is necessary to give effect to the policy stated by Lord Mansfield and are certain enough to be predictable in their application’.⁷⁸ Indeed, the ‘merit-blind’ nature of the *Tinsley* test may paradoxically be a strength even if it is also a weakness, insofar as one is generally able to tell in advance how the test would apply to the facts of a given case. This may be contrasted with the ‘inextricable link’ test in *Cross* and even the causation approach in *Gray*. As was noted by Lord Hoffmann in *Gray*, the metaphor of ‘inextricable link’ is unhelpful; the phrase merely begs the questions of what sort of ‘links’ would have to exist between the claimant’s criminal conduct and his claim, and how ‘inextricable’ such a link would have to be. Further, the judgment in *McCracken* amplifies any doubts one might have as to whether the causation approach constitutes a substantial improvement over the ‘inextricable link’ test where clarity and certainty of application are concerned. As such, if the *Tinsley* test is to be faulted for being blind to the policy considerations underlying the illegality defence, it may tenably be rebutted that the courts applying the ‘causation’ or ‘inextricable link’ tests have indeed nothing *but* (necessarily open-textured) policy considerations to go on with.

Like the first argument, however, the second stops far short of a complete justification of the *Tinsley* test. The importance of clarity and certainty must of course not be understated, but it is tristic that the courts have no business in applying a test that does not further desirable judicial aims, regardless of how clear or certain it may be. Given that it is common ground that the illegality defence is based on considerations of public policy, it would be inexplicable to hold that courts are to cleave rigidly to a test that only occasionally and incidentally produces results that uphold that policy. This argument derives greater force when it is remembered that the policy underlying the defence ‘is not based upon a single justification but on a group of reasons, which vary in different situations’:⁷⁹ if a single rigid test is unable to consistently uphold a single discrete public policy, it is puzzling how it can adequately give weight to multiple such policies at the same time, especially when their furtherance would lead to conflicting results in a given case. Finally, the need to examine policy underpinning the illegality defence is made all the more crucial in cases where policy factors *extrinsic* to the defence may seem to compel its disapplication; the dilemma that faced the Supreme Court in *Hounga* provides a paradigm example. As we have seen, Lord Sumption’s dismissal of *Hounga* as exceptional is perforce unsatisfactory, because it begs the question of *how* exceptional it was, and *why*. Specifically, why in *Hounga* was a ‘competing’ public policy allowed to disapply the illegality defence where it otherwise would have applied? The notion of ‘competition’ is here instructive: the offending public policy must be weighed against the policy underlying the defence, and it is preferable that the courts do this transparently, as Lord Wilson did in *Hounga*, rather than opaquely.

A final example will demonstrate that explicit discussion of policy considerations can be profitable to the law on illegality and to the general law, and paradoxically lead to greater clarity than if a single inflexible test were employed. In the recent Court of Appeal decision in *R (Best) v Chief Land Registrar*,⁸⁰ the Government had appealed against a decision that a squatter could rely upon his adverse possession of a residential property in order to claim title to the property even though his occupation amounted to a criminal offence under section 144 of the Legal Aid, Sentencing and Punishment of Offenders Act 2012 (LASPOA). The court noted that:⁸¹

⁷⁸ *Jetivia* (n 5) [62].

⁷⁹ *Gray* (n 28) [30] (Lord Hoffmann).

⁸⁰ [2015] EWCA Civ 17.

⁸¹ *ibid* [52] (Sales LJ).

...there is not one single rule with blanket effect across all areas of the law. Instead, there are a number of rules which may be identified, each tailored to the particular context in which the illegality principle is said to apply.

Having referred closely to Lord Wilson's approach in *Hounga*, the court thus considered it necessary to consider the aspect of public policy which founded the illegality defence, and then weigh this against the countervailing public policy considerations that ran counter to the defence. After considering the respective public policy considerations underlying the law on adverse possession – as embodied in the provisions of the Land Registration Act 2002 (LRA) – and s 144 LASPOA, the court concluded that the defence should not apply in the present case. It is submitted that this explicit sensitivity to the respective regimes of LASPOA and adverse possession is welcome, insofar as it allowed the court to acknowledge and weigh the demands of public policy openly, something that would necessarily have been precluded by a mechanistic application of the *Tinsley* test.

V. IDEAS FOR REFORM

It is by now hopefully evident that it is time for reform to free the common law from the straitjacket of *Tinsley*. Quite like in *Tinsley* itself, however, there is a far stronger consensus on what should *not* be law than on what *should* be.

A study of possible reforms would be remiss if it did not discuss the proposals of the Law Commission in this area. A consultation paper published in 2009⁸² (“CP 189”) and a report presented to Parliament in 2010⁸³ (“LC 320”) make evident the Commission's distaste for the *Tinsley* reliance test, for many of the reasons discussed above. Accordingly, in CP 189, the Commission initially suggested statutory reform for the illegality defence as and where it applied, but later resiled from that in LC 320, stating that recent House of Lords judgments in *Gray* and *Stone & Rolls* had indicated a judicial move away from *Tinsley* in areas such as contract and tort, such that statutory reform no longer appeared to be required in cases that were factually dissimilar to *Tinsley*.⁸⁴ As such, statutory reform was recommended only for trusts cases which fell squarely within the ambit of *Tinsley*, wherein courts were to be granted a structured discretion to apply the illegality defence.⁸⁵ This discretion would thus take into account a non-exhaustive list of factors comprising: (1) the conduct and intention of the parties; (2) the value of the equitable interest at stake; (3) the effect of allowing the claim on the criminal purpose; (4) whether refusing the claim would act as a deterrent; and (5) the possibility that the person from whom the equitable interest was being concealed may have an interest in the value of the assets of the beneficiary.⁸⁶

In a subsequent article,⁸⁷ Andrew Burrows lauded the Commission for their efforts in catalysing reform in the law of illegality. Noting piecemeal judicial reforms initiated by the House of Lords in *Gray*, and the Court of Appeal in *Apotex* and *Parkingeye*, he concluded that the law:⁸⁸

⁸² Law Commission, *The Illegality Defence: A Consultative Report* (Law Com CP No 189, 2009).

⁸³ Law Commission, *The Illegality Defence* (Law Com No 320, 2010).

⁸⁴ *ibid* para 1.18. See generally Part 3: Illegality and Other Claims.

⁸⁵ *ibid* para 1.17. See generally Part 2: Illegality and Trusts.

⁸⁶ Law Com No 320 (n 83) para 2.80.

⁸⁷ Andrew Burrows, ‘Alternatives to Legislation: Restatements and Judicial Law Reform’ in Gullifer and Vogenauer (eds), *English and European Perspectives on Contract and Commercial Law* (Hart Publishing 2014).

⁸⁸ *ibid* 50.

...has therefore been put on a more flexible basis that allows the courts to cut through to the underlying policies and to reach a proportionate result. That has been done without the legislation that I, for one, thought would be needed.

With the benefit of hindsight, it must surely now be clear that the law has *not* clearly evolved in the way contemplated either by the Commission or by Burrows. The Court of Appeal decision in *Apotex* of course went on to be overruled in the Supreme Court after the publication of Burrows' article, and our foregoing analysis of *Jetivia* has confirmed that the law of illegality remains unsatisfactorily turbid, not just with respect to cases involving trusts. For instance, Lord Neuberger's ambivalence over the applicability of *Tinsley* in that case⁸⁹ was not intended to be limited to trusts law, and even in *Parkingeye*, Sir Robin reached his decision on the footing that the reliance test was applicable in the law of contract.⁹⁰ Thus, reform in this area of the law necessarily remains illusive, but – if we are to accept the criticisms thereof – necessary.

How then should reform proceed, and along what lines? If we accept that judicial efforts have not succeeded in satisfactorily reforming the illegality defence over the last five years, the necessity of legislative reform may seem ineluctable, even in cases that are not ostensibly caught by the precedential ambit of *Tinsley*. Certainly, other jurisdictions have enshrined in statute rules that are analogous to the UK illegality defence. For instance, in New Zealand, the Illegal Contracts Act 1970 governs the effect of illegality in the law of contract. Under the Act, every 'illegal contract'⁹¹ 'shall be of no effect and no person shall become entitled to any property under a disposition made by or pursuant to any such contract'.⁹² The court is given an unstructured remedial discretion in such cases: it has the power to grant relief '...as the court in its discretion thinks just'.⁹³

Crucially, however, it is submitted that the discussion above makes it evident that an unbounded discretion such as that granted in the 1970 Act is undesirable, at least in the UK context. Given that the illegality defence is based on considerations of public policy, it must be evident that enabling courts with a formless statutory discretion does not clearly lead them to vindicate those considerations. It is instructive to note that implementing such discretion by legislation would effectively amount to raising (and immortalising) the spectre of the public conscience test that had confronted the House of Lords in *Tinsley*; as Lord Goff implied⁹⁴ in that case, the discretion afforded by the public conscience test was not likely going to be capable of giving sufficient effect to the principle of public policy enunciated by Lord Mansfield in *Holman v Johnson*.

While the Commission's recommendation of a 'structured' discretion might thus appear to be preferable to both a rigid technical test and a conferral of unfettered discretion, insofar as it compels explicit judicial consideration of public policy, it must be queried whether a statutory solution helps in any way in the vindication of applicable public policy factors. Quite obviously, while the Commission has recommended several factors that courts should take into account in applying the illegality defence, not all of the suggested factors *directly* reflect public policy considerations: for instance, the first factor (conduct and intention of relevant parties) is stated to *derive* from multiple policies that justify the illegality defence, particularly deterrence and the protection of the integrity of the legal system.⁹⁵ However, if courts are ultimately to have regard to the underlying public policy factors when applying statutory factors (as they would inevitably be compelled to do, in the absence of further statutory elaboration), it is unclear whether the statute would do any work at all. On the contrary, any net effect

⁸⁹ *Jetivia* (n 5) [17].

⁹⁰ *Parkingeye* (n 62) [36].

⁹¹ As defined by the common law: Illegal Contracts Act 1970 (NZ), s 3.

⁹² Illegal Contracts Act 1970 (NZ), s 6.

⁹³ Illegal Contracts Act 1970 (NZ), s 7.

⁹⁴ *Tinsley* (n 4) 363.

⁹⁵ Law Com No 320 (n 83) para 2.67.

of the statute must surely be obfuscatory, given that it might detract from the ultimate and paramount need to apply the defence *if and only if* a balance of public policy considerations demands it. Furthermore, as the Commission concedes,⁹⁶ the policy rationales underlying the defence are multiple and varied; given that there can be no hierarchy of rationales, it is more than a little unsatisfactory that the Commission suggested statutory factors that were designed to uphold *six* discrete policy rationales⁹⁷ without substantially explaining why these should be considered and vindicated to the exclusion of the others. Although the Commission clearly stated that the factors were intended to be non-exhaustive,⁹⁸ it would be unsurprising, indeed natural, if courts were nevertheless to deem that the named factors were more significant than other public policy factors for no other reason than the fact that they were named, even if other factors should otherwise be deemed to be more relevant in a given case. Finally, statutory reform also inhibits the development of new rationales for the illegality defence – indeed, the possibility that the law may be frozen in a way that makes desirable change difficult⁹⁹ is regrettably endemic to all instances of legislative reform.

In summary, it is argued that the Law Commission was right in advocating reform with respect to the illegality defence, especially given the inflexibility of the *Tinsley* test and the lack of clarity as to its precedential scope. However, it was too sanguine in noting that reform in areas outside of trusts was well underway following the cases of *Gray* and *Stone & Rolls*; even if the courts had taken promising steps toward desirable reform in those cases, they have surely since backtracked. Furthermore, it failed to note that the attractiveness of legislative reform in this area was merely specious. In an area of the law so thoroughly justified by multiple and context-sensitive public policy considerations, it is at best unhelpful, and at worst distracting, to prescribe factors that the courts must take into account when considering the defence, even if such factors are intended as proxies for said considerations. The courts must be the engineers of reform; they must recognise that reform outside the area of trusts law is no less pressing than reform inside it; and they must not shy away from the need to explicitly tackle policy considerations in deciding whether or not to apply the illegality defence.

VI. CONCLUSION

In the second of a series of celebrated lectures,¹⁰⁰ Benjamin Cardozo stated shortly that '[t]he rule that misses its aim cannot permanently justify its existence'.¹⁰¹ This teleological approach is eminently applicable to the law relating to the illegality defence, which has since *Holman v Johnson* been firmly and fundamentally grounded in public policy. Thus even insofar as *Tinsley v Milligan* operated to straitjacket the defence by rejecting judicial discretion and imposing a value-free test founded on procedural reliance, courts have often escaped its rigidity by explicitly distinguishing the case (as in *Gray*) or implicitly departing from it (as in *Parkingeye*). However, this subversion of *Tinsley* has not been wholly desirable: trusts cases that are factually similar to *Tinsley* are ineluctably bound by it, and alternative approaches to the illegality defence remain unsatisfactory vehicles of the public policy considerations that underpin the illegality defence. The current rules that apply to the defence are eminently capable of missing their aim; one should thus not be overly reluctant to conclude that their existence is not, and cannot, be justified.

⁹⁶ Law Com No 320 (n 83) para 2.5.

⁹⁷ *ibid.*

⁹⁸ *ibid* para 2.78.

⁹⁹ See Burrows (n 87) 43.

¹⁰⁰ Benjamin Cardozo, *The Nature of the Judicial Process* (Yale University Press 1921).

¹⁰¹ *ibid* 66.

The Right Approach to Wrongful Conception

Shaun Elijah Tan¹

1. INTRODUCTION

A case of ‘wrongful conception’ was defined in *McFarlane v Tayside Health Board*² as ‘an action by parents of an unwanted child for damage resulting to them from the birth of the child’.³ Lord Steyn here suggested that ‘instinctively, the traveller on the Underground would consider that the law of tort has no business to provide legal remedies consequent upon the birth of a healthy child, which all of us regard as a valuable and good thing’.⁴ However, the majority of the Australian judges sitting in the High Court of Australia, in the case of *Cattanach v Melchior*,⁵ were certainly no travellers on the Underground. *Cattanach* decided—contrary to the decision in *McFarlane*—that the parents of a child born as a result of a doctor’s negligence are entitled to recover damages for the costs of raising the child until adulthood. The purpose of this essay is twofold. First, the various analyses on actionable damage and public policy from these two cases will be outlined and then explored in greater depth. I will argue that these two cases, *McFarlane* and *Cattanach*, were not decided on sound bases, whether on grounds of public policy or legal principle. Ultimately, I suggest that instead of viewing wrongful conception as a tort of negligence, the ‘right approach’ is to view wrongful conception as a nominate tort, akin to the trespassory torts, which are actionable *per se*.

II. UNDERSTANDING MCFARLANE AND ‘WRONGFUL CONCEPTION’

Mr and Mrs McFarlane had four children in 1989. They purchased a bigger house and needed a larger mortgage. To meet the growing financial needs of the family, Mrs McFarlane decided to return to work. The couple decided that Mr McFarlane would undergo a vasectomy, because they did not want any more children. On 16 October 1989, a consultant surgeon performed the operation on Mr McFarlane at a hospital operated by Tayside Health Board. The operation was successful. A consultant surgeon later wrote to Mr McFarlane on 23 March 1990, informing him that ‘[his] sperm counts are now negative and [he] may dispense with contraceptive precautions’. The couple relied on this advice. However, in September 1991, Mrs McFarlane became pregnant and gave birth to a healthy daughter, Catherine, on 6 May 1992.

The legal proceedings started in Scotland. The couple sued Tayside Health Board in delict and the claim can be split into two parts. First, Mrs McFarlane claimed £10,000 for the pain, suffering

¹ St Edmund Hall. I am thankful to Sarah Green, Dr Benjamin Spagnolo, Benjamin Ong, Matthew Chan, Zhang Weiran and Denise Lim for their incisive comments. Any errors and solecisms remain solely my own. This essay is dedicated to all my tutors, in particular, Professor Adrian Briggs and Dr Aileen Kavanagh.

² [2000] 2 AC 59 (HL).

³ *ibid* 76 (Lord Steyn). ‘Wrongful conception’ is different from ‘wrongful life’. A ‘wrongful life’ claim is made by a child for damage to himself arising from the very fact of his birth. This type of claim involves the child arguing that he should never have been born. The English courts have rejected ‘wrongful life’ claims because they violate the sanctity of human life: *McKay v Essex Area Health Authority* [1982] 2 All ER 771 (CA). This essay focuses exclusively on ‘wrongful conception’. Further, it focuses on a healthy mother wrongfully conceiving a healthy baby.

⁴ *McFarlane* (n 2) 82.

⁵ [2003] HCA 38. This decision, however, has been qualified in some States. Queensland, for example, through the *Justice and Other Legislation Amendment Act 2003*, inserted sections 49A and 49B into the *Civil Liability Act 2003*, disallowing claims for ‘costs ordinarily associated with rearing or maintaining a child’ (s 49A(2)). However, the legal principles expounded in *Cattanach* are still relevant to our present purposes.

and distress resulting from the unwanted pregnancy. Second, the couple claimed a sum of £100,000 for the cost of bringing up Catherine. At first instance, the Lord Ordinary, Lord Gill, dismissed the action with regard to both heads of claim. The thrust of Lord Gill's judgment was that 'the privilege of being a parent is immeasurable in monetary terms (...) the benefits of parenthood transcend any patrimonial loss'.⁶ However, the Inner House allowed a reclaiming motion and reversed Lord Gill's order, holding that both heads of claim are recoverable on conventional principles of delict.⁷ On appeal, the House of Lords unanimously held that the maintenance costs for a healthy child were not recoverable. On the issue of Mrs McFarlane's pain, suffering and distress arising from the unwanted pregnancy, the House of Lords, by a 4:1 majority, held that the claim was recoverable.

Cattanach, a similar case heard by the High Court of Australia,⁸ revolved mainly around the same issues. Mr and Mrs Melchior, satisfied with the size of their family, decided to stop having more children. When Mrs Melchior first consulted Dr Cattanach, she told him that her right ovary and fallopian tube had been removed. Dr Cattanach then went on to perform a tubal ligation on Mrs Melchior, attaching a clip only to Mrs Melchior's left fallopian tube. Eventually, Mrs Melchior discovered that she was pregnant and gave birth to her son, Jordan. It transpired that her right fallopian tube had not been removed. The trial judge found that Dr Cattanach was negligent because he had too readily accepted Mrs Melchior's assertion, without any further investigation. The High Court of Australia, by a 4:3 majority, held that the couple could recover damages for the potential maintenance costs of Jordan.

The House of Lords and the High Court of Australia had to grapple with two major issues, *inter alia*, in *McFarlane* and *Cattanach* respectively, in order to determine the damages that the couples were entitled to. First, what is the actionable damage in such cases? Second, what exactly was the scope of the duty of care that the surgeons owed to the couples? To determine the exact scope of the duty of care, most of the judges sitting in both cases resorted to policy arguments. Their Lordships' and their Honours' answers to these two questions will now be subject to closer examination.

III. VARYING CONCEPTIONS OF ACTIONABLE DAMAGE

It is necessary to be clear about what the actionable damage is and about what is 'consequential' on it. As we shall observe, the damages that the claimants are entitled to turn on a construction of actionable damage. If the maintenance costs can be considered as consequent on physical injury, then the economic loss will be allowed. However, if the maintenance costs are considered pure economic loss, it will not be recoverable unless it falls into an exceptional category.⁹ This position is evinced in the case of *Spartan Steel & Alloys v Martin & Co*,¹⁰ where the Court of Appeal held that economic loss 'consequent' on physical damage is recoverable but economic loss 'independent' of the physical damage is not.¹¹

⁶ *McFarlane v Tayside Health Board* (1997) SLT 211 (SC (OH)), 216.

⁷ *McFarlane v Tayside Health Board* (1998) SCLR 126 (Court of Session, Inner House (Second Division)).

⁸ The High Court of Australia is the highest appellate court in Australia.

⁹ Pure economic loss is not generally recoverable in the English and Australian law of negligence. One exception is *Hedley Byrne v Heller* [1964] AC 465 (HL) type of liability for negligent misstatements. Another exception is a *Junior Books v Veitchi* [1983] 1 AC 520 (HL) scenario, where the high degree of proximity between the parties was said to make the relationship 'akin to contract'.

¹⁰ [1973] QB 27 (CA).

¹¹ *ibid* 39 (Lord Denning).

The High Court of Australia and the House of Lords offered varying interpretations of actionable damage. However, all except one of these interpretations should be rejected because they are logically untenable.

One view advanced is that, while the mother has suffered personal injury, the maintenance costs are not consequential on it. The maintenance cost is thus categorised as pure economic loss and, therefore, unrecoverable. Two arguments were presented in *McFarlane* to achieve this very end.

The first was presented by Lord Hope. His Lordship attributed the claim for maintenance cost to the father, and concluded that 'this is a claim for economic loss' because Mr. McFarlane 'does not claim that he suffered physical or mental injury'.¹² By this analysis, the maintenance costs would no longer be consequential on the physical injury of the mother. It follows that the claim is one of pure economic loss. However, the artificiality of this analysis is plain. As Hoyano rightly pointed out, this analysis 'obviously has limited utility where the mother is a single parent or is in paid employment'.¹³

The second was to divide the duty of care, owed by the doctor to the patient, into several parts. Lord Slynn held that 'the doctor undertakes a duty of care in regard to the prevention of pregnancy: it does not follow that the duty includes also avoiding the costs of rearing the child if born and accepted into the family'.¹⁴ Whilst His Lordship would grant compensation for the 'physical effects of the pregnancy', the doctor 'does not assume responsibility for those economic losses'.¹⁵ The maintenance costs, it follows, would be detached from physical injury to the mother. This line of reasoning was not well-received in *Cattanach*. Most of the judges in *Cattanach* thought that if there was physical injury inflicted on the mother by the doctor's negligence, it follows that the maintenance costs were 'consequential'. Moreover, Hoyano also points out that 'there was no attempt in *McFarlane* to explain why or how a case of failed sterilisation is different from other cases of negligently performed surgery where a separate duty of care analysis is not required for future financial losses, such as loss of income'.¹⁶

Another view was suggested by Gleeson CJ in his dissenting judgment in *Cattanach*. His Honour thought that if the Melchioris had suffered damage 'it is because of the creation of that relationship and the responsibilities it entails'.¹⁷ His Honour asserted that the relationship between parent and child 'is the immediate cause of the anticipated expenditure which the respondents seek to recover by way of damages'.¹⁸ The corollary is that the court was dealing with 'a claim for recovery of pure economic loss arising out of a relationship' rather than 'a claim for financial loss consequential upon personal injury to a plaintiff, or damage to a plaintiff's property'.¹⁹ This therefore meant that the claim for maintenance costs was unrecoverable. Gleeson CJ's argument was persuasively criticised by two of the judges in the majority, McHugh and Gummow JJ. Their Honours suggested that Gleeson CJ's argument examined the case from the 'wrong perspective'.²⁰ Their Honours thought that the actionable damage is the 'expenditure that they have incurred or will incur in the future, not the creation or existence of the parent-child relationship'.²¹

¹² *McFarlane* (n 2) 89.

¹³ Laura Hoyano, 'Misconceptions about Wrongful Conception' (2002) 65 MLR 883, 886.

¹⁴ *McFarlane* (n 2) 76.

¹⁵ *ibid.*

¹⁶ Hoyano (n 13) 887.

¹⁷ *Cattanach* (n 5) [26].

¹⁸ *ibid.*

¹⁹ *ibid* [30].

²⁰ *ibid* [67].

²¹ *ibid.*

It is submitted that Gleeson CJ's argument, though creative, should ultimately be rejected. First, Gleeson CJ's analysis smacks of artificiality. How can the law deny that such a relationship arises as a consequence of pregnancy? Gleeson CJ, in his judgment, failed to offer a convincing argument on why and how we can deny the link between the process of pregnancy and the relationship that follows inexorably as a corollary. Second, if Gleeson CJ's argument were adopted in the English courts, it would introduce uncertainty into an area which was hitherto thought clear. It is foreseeable that if Gleeson CJ's approach were to be adopted, courts in the future might imply such relationships between the different parties in order to convert a putative consequential economic loss to pure economic loss based on the courts' intuition.

The most plausible view, advanced by majority in *Cattanach*, is that maintenance costs for the child flowed naturally from the claim for the pain and suffering of pregnancy and was not, thus, a claim for pure economic loss. The maintenance costs of raising the child would therefore be recoverable, on a straightforward application of legal principle. McHugh and Gummow JJ, for example, held that the damage is the 'burden of the legal and moral responsibilities which arise by reason of the birth of the child that is in contention'.²² This perspective was also echoed by Kirby and Hayne JJ (although the latter eventually dissented on the basis of public policy). Kirby J was unwavering in his tone when he suggested that because Mrs. Melchior suffered physical injury, 'she would be entitled to recover on normal principles without disqualification'²³ the maintenance costs of rearing the child.

Even though this analysis is the most principled amongst all the others—since it adheres to the conventional understanding of consequential and pure economic loss—there is something distasteful about it. This approach necessarily leads us to conclude that the birth of the child is the actionable damage. Lord Millett agreed with Lord McCluskey that assigning a monetary value to the child's existence is 'as difficult and unrealistic as it is distasteful'.²⁴ His Lordship very perceptively points out that 'the exercise must either be superfluous or produce the very result which is said to be morally repugnant. If the monetary value of the child is assessed at a sum in excess of the costs of maintaining him, the exercise merely serves to confirm what most courts have been willing to assume without it. On the other hand, if the court assesses the monetary value of the child at a sum less than the costs of maintaining him, it will have accepted the unedifying proposition that the child is not worth the cost of looking after him'.²⁵

IV. POLICY: CONTROLLING OR LOOSING THE REINS OF THE UNRULY HORSE?

Public policy was once likened by Burroughs J to 'a very unruly horse and when you get astride it you never know where it will carry you. It may lead you from the sound law'.²⁶ However, this view stands in stark contrast to Lord Denning's opinion in *Enderby Town FC v Football Association*.²⁷ Lord Denning disagreed and suggested that 'with a good man in the saddle, the unruly horse can be kept in control'.²⁸ With respect to Lord Denning, it will be argued here that, in this area of law, Burroughs J was right; the unruly horse has led the courts into the wilderness of inconclusive arguments. In *McFarlane*, most of the judges sought to rely on public policy arguments in order to justify non-recoverability of the maintenance costs. These arguments will be explored in greater depth in the section below.

²² *Cattanach* (n 5) [68].

²³ *ibid* [149].

²⁴ *McFarlane* (n 2) 111.

²⁵ *ibid*.

²⁶ *Richardson v Mellish* (1824) 2 Bing 229, 252.

²⁷ [1971] Ch 591 (CA).

²⁸ *ibid* 606.

One such policy argument can be labelled as ‘moral argument’. This argument focuses mainly on the sanctity of life and the impact on the growth of the child should he or she find out that he or she was unwanted by the parents. Lord Millett in *McFarlane* put the point across very strongly in arguing that the law should regard the birth of a normal, healthy baby to be ‘a blessing, not a detriment’.²⁹ Lord Millett argued that though in truth, the birth of a child is a ‘mixed blessing’—since parenthood comes with its ups and downs—the ‘society itself must regard the balance as beneficial’.³⁰ *Pace* Lord Millett, though, this argument does not take us very far. Kirby J pointed out in *Cattanach* that this argument ‘represents a fiction which the law should not apply to a particular case without objective evidence that bears it out’.³¹ The notion of ‘blessing’ is an amorphous one. It conceals a whole series of important questions: by whose standard should this ‘blessing’ be measured? How can the intangible blessings be properly measured? It is intrusive for the law to make value judgments like these. By answering these questions, the law will inevitably foist its understanding of ‘blessing’ on the parents – when they plainly do not consider parenthood to be a blessing – and this fails to respect the couples’ reproductive choices and autonomy.

Second, distributive justice has also been offered as a public policy argument. Lord Steyn in *McFarlane* invoked this in highlighting that the matter ‘requires a focus on the just distribution of burdens and losses among members of a society’. His Lordship mentioned that if the question—whether maintenance costs in wrongful conception should be recoverable—were to be asked to commuters on the Underground, ‘an overwhelming number of ordinary men and women would answer the question with an emphatic ‘No.’ And the reason for such a response would be an inarticulate premise as to what is morally acceptable and what is not’.³²

It is interesting to note that Lord Steyn has been a vocal proponent of resorting to moral considerations in the law. In *Smith New Court Securities v Citibank N.A.*,³³ Lord Steyn said: ‘I make no apology for referring to moral considerations. The law and morality are inextricably interwoven. To a large extent the law is simply formulated and declared morality’.³⁴ However, this policy argument has been subjected to trenchant criticism from the judiciary and academic commentators alike. Hoyano argued that distributive justice ‘permits the judiciary to abdicate its responsibility to identify and explain intellectually rigorous and coherent principles as the basis for decisions, in favour of an empirically untested appeal to public opinion, yielding unpredictable results which invite reversal at each level of appeal, depending on each judge’s subjective and avowedly instinctive notions of what justice requires’.³⁵ This has also been echoed in the High Court of Singapore. Choo Han Teck J, in *ACB v Thomson Medical Pte Ltd*,³⁶ remarked that ‘if public policy is a quicksand, philosophies of distributive justice are sinkholes’.³⁷ McHugh and Gummow JJ also cited Hale LJ’s criticism of Lord Steyn’s judgment with approval.³⁸ Hale LJ, in *Parkinson v St James and Seacroft University Hospital*,³⁹ highlighted that ‘the fact that so many eminent judges all over the world have wrestled with this problem and reached different conclusions might suggest that the considered response would be less emphatic and less unanimous’.⁴⁰

²⁹ *McFarlane* (n 2) 114.

³⁰ *ibid.*

³¹ *Cattanach* (n 5) [148].

³² *McFarlane* (n 2) 82.

³³ [1997] AC 254 (HL).

³⁴ *ibid* 280.

³⁵ Hoyano (n 13) 905.

³⁶ [2014] SGHC 36.

³⁷ *ACB* (n 36) [8].

³⁸ *Cattanach* (n 5) [82].

³⁹ [2002] QB 266 (CA).

⁴⁰ *ibid* [82].

The problems with this policy argument are immediately apparent: it is a bald assertion on his Lordship's part, which requires empirical validation. In fact, the decision in *Cattanach* serves as a good counterpoint to Lord Steyn's assertion, *viz.* that 'an overwhelming number of ordinary men and women' would not grant child-rearing costs as compensation. But even if this could be empirically proven, it would not follow that the judges should apply the majority view through the law. Stevens pointed out that 'if our rights are to be overridden whenever there is a consensus opinion that they are not deserving of respect, we do not have rights worthy of the name. The law is not, and should not be, determined by the judge's best guess of majority public opinion'.⁴¹ Finally, there are many competing conceptions of distributive justice in the sea of literature.⁴² Lord Steyn's version of distributive justice is merely one of a spectrum. This gives rise to a crucial and unresolvable question: why should his Lordship's account be adopted over the many others?

Last, the courts in England have also justified the decision of *McFarlane*—albeit an *ex post facto* rationalisation—on the grounds of protecting the funds of the National Health Service ('NHS'). Lord Bingham in *Rees v Darlington Memorial Hospital NHS Trust*⁴³ suggested that one of the policy considerations underlying the decision in *McFarlane* was 'a sense that to award potentially very large sums of damages to the parents of a normal and healthy child against a National Health Service always in need of funds to meet pressing demands would rightly offend the community's sense of how public resources should be allocated'.⁴⁴ His Lordship also emphatically supported Kirby J's suggestion in *Cattanach*, where His Honour pointed out that 'concern to protect the viability of the National Health Service at a time of multiple demands upon it might indeed help to explain the invocation in the House of Lords in *McFarlane* of the notion of "distributive justice"'.⁴⁵ This was also reiterated by Lord Nicholls in *Rees*, where his Lordship pointed out that to argue that the 'National Health Service should pay all the costs of bringing up the child' seems like 'a disproportionate response to the doctor's wrong'.⁴⁶

However, this policy argument does not appear to be of much succour. First, this argument raises another question: why should the cost rightly fall on the claimant rather than the NHS? If the answer lies in the quantum of damages, the courts will face an uphill task trying to draw a bright line between when to compensate and when not to. Second, based on empirical research into allocation of public health resources, it has been suggested that 'decisions are made on the basis of clinical and cost effectiveness at the expense of ethical inquiry into what is acceptable'.⁴⁷ *Ex hypothesi*, Chico argues that negligence claims must be judged on their cost effectiveness rather than merit, and there is no *a priori* reason why negligence victims should be disentitled to a claim in damages.⁴⁸ Additionally, Chico perceptively highlights the incoherence in this argument, with regard to the court's disregard to the extension of negligence liability against NHS in other areas. She points out, for instance, that the decision of *Chester v Afshar*⁴⁹ does not chime with this policy argument when applied in the *McFarlane* context. The House of Lords allowed a novel claim, in *Chester*, by relaxing traditional causation rules. She rightly questions 'if the NHS can afford *Chester* why can it not afford *McFarlane*?'⁵⁰

⁴¹ Robert Stevens, *Torts and Rights* (OUP 2007) 311.

⁴² Peter Cane, 'Distributive Justice and Tort Law' (2001) N.Z.L.Rev. 401, 405.

⁴³ [2003] UKHL 52.

⁴⁴ *ibid* [6].

⁴⁵ *Cattanach* (n 5) [178].

⁴⁶ *Rees* (n 43) [16].

⁴⁷ Victoria Chico, 'Wrongful Conception: Policy, Inconsistency and the Conventional Award' (2007) MLI 8(2) 139, 146.

⁴⁸ *ibid*.

⁴⁹ [2004] UKHL 41.

⁵⁰ Chico (n 47) 149.

By examining some of the policy arguments offered by the courts, it is hoped that the problems of resorting to policy, in this area of the law, are thoroughly exposed. Far from arriving at a reasoned answer, the courts have been led into the wilderness by this unruly horse.

By the end of these two sections, it appears as though we have hit somewhat of a dead end. If we cannot decide the case on conventional negligence principles, are we then to wave our white flags, sink back in our seats and resign to our fate?

V. A FRESH START

A fresh start is needed. In this section, I argue that the tort of wrongful conception should be seen as an actionable tort *per se*, similar to the trespassory torts, rather than as a part of the tort of negligence. First, I intend to explore the reasons why the torts of trespass are actionable *per se*. Next, I will explore the advantages and disadvantages of reconceptualising wrongful conception as a discrete tort on its own, and then conclude that only by reshaping wrongful conception as an actionable tort *per se* can we achieve greater analytical clarity in this area of law.

Why are some torts, especially the trespassory ones, actionable *per se*? What does it mean for a tort to be actionable *per se*? As Nolan and Davies explain ‘some torts are actionable only if harm is caused, others are actionable without proof of damage’ (*per se*). The distinction is an accident of history but can be rationalised. Torts actionable without proof of damage are either torts where damage can be presumed, or they are torts whose function is to protect particular rights from any invasion, whether or not damage has resulted. Nominal damages may be given to signify that a right has been invaded although no harm has been done.’⁵¹

Why is it advantageous to view wrongful conception as an actionable tort *per se*, rather than a tort of negligence?

First, it can be argued that, by doing so, the law escapes the preceding discussion on actionable damage and public policy. The problems with seeing wrongful conception as a tort of negligence are plentiful. First, we have to conclude that the maintenance costs of bringing up the child are consequential economic losses, by ordinary legal principles. By that line of reasoning, we have to conclude that the birth of the child is an actionable damage. This is a conclusion which the judges in *McFarlane* tried to escape from, by magically holding that the maintenance costs are pure economic loss, without any clear explanation. Further, it has been pointed out that the policy arguments applied are weak and inconclusive. Instead of helping us come to a conclusion on whether there is a duty of care these policy arguments pull us further away from it by further complicating matters. These problems of viewing wrongful conception as a tort of negligence are serious – they are serious enough to warrant a rethink of the classification of wrongful conception.

⁵¹ Donal Nolan and John Davies, ‘Torts and Equitable Wrongs’ in Andrew Burrows (eds), *English Private Law* (OUP 2013).

Second, if we view wrongful conception as an actionable tort *per se*, we redirect the focus back on the rights of the mother, rather than on the fault of the doctor. This has been the emphasis of Lord Bingham and Lord Millett in *Rees* and it is a sensible approach. Lord Millett, for instance, suggested that the modest conventional sum awarded in *Rees* was for ‘the denial of an important aspect of their personal autonomy, *viz.* the right to limit the size of their family’.⁵² Lord Bingham, in the same case, spoke of the legal wrong as depriving the mother of ‘the opportunity to live her life in the way that she wished and planned’.⁵³ If we transform wrongful conception into a nominate tort, into an actionable tort *per se*, the gist of the action consists in the wrong, in infringing the claimant’s freedom to live her life in the way that she had so planned. This is a wrong that is serious enough to garner protection from any potential invasion, regardless of whether legally recognisable actionable damage arises.

Third, by doing so, the conventional sum⁵⁴ can be justified on the basis of the nominal sum. If wrongful conception is seen as a tort of negligence, it is difficult to justify the conventional sum, as Lords Steyn, Hope and Hutton held as the minority in *Rees*. The minority opposed this conventional sum for a few reasons. Lord Steyn objected because it is ‘forbidden territory’ that the judges should not enter into and that ‘there are limits to permissible creativity for judges’.⁵⁵ Additionally, his Lordship also pointed out that the sum is ‘a backdoor evasion of the legal policy enunciated in *McFarlane*’.⁵⁶

Lord Hope sustained a vigorous critique of the majority’s judgment. First, his Lordship held that the compensatory principle would be broken if ‘the conventional sum was intended to give [the parents] something for their financial loss’ because ‘it would deny them the opportunity of attempting to establish the true value of that part of their claim according to the compensatory principle’.⁵⁷ Second, his Lordship pointed out the lack of ‘any consistent or coherent ratio’ for the conventional sum. His Lordship pointed out that Lord Bingham’s claim, *viz.* that the conventional sum is not compensatory in nature, and that it deviates from the normal approach to the assessment of damages. Additionally, Lord Hope questions: if the sum is not based on compensation, ‘what basis can there be for it?’⁵⁸ Last, his Lordship points out that the sum might be deemed as ‘derisory’ by the parents.

Lord Hope points out that the courts are in ‘uncharted waters’ and questions ‘how is one to measure the loss of the right to limit the size of one’s family against an award of that kind, bearing in mind the far-reaching and long-lasting effect that the birth of the uncovenanted child will have on the life of the parent?’⁵⁹ The onslaught, executed elegantly by the minority on the conventional sum, is valid. But, it is only valid if the conventional sum is thought of as compensatory, since compensatory damages are usually awarded to claimants who successfully establish negligence on the part of the defendants. However, if wrongful conception were instead seen as a tort which is actionable *per se*, very much like the tort of trespass, it follows that the court can grant nominal damages due to the wrong committed by the defendant on the claimant. The court need not ‘compensate’ for any loss or damage; it only needs to give a sum to signify that a right has been infringed. The conventional sum approach favoured by the majority in *Rees* can therefore be justified under the actionable tort *per se* approach.

⁵² *Rees* (n 43) [123].

⁵³ *ibid* [8].

⁵⁴ The majority in *Rees*, by the slight margin of 4-3, decided to put a ‘gloss’ (as termed by Lords Bingham and Nicholls) on the *McFarlane* judgment by granting a conventional sum of £15,000 to the mother (or the couple jointly) in cases of wrongful conception.

⁵⁵ *Rees* (n 43) [46].

⁵⁶ *ibid*.

⁵⁷ *ibid* [73].

⁵⁸ *ibid* [74].

⁵⁹ *ibid* [75].

VI. SUMMARY: THE ROAD NOT TAKEN

Robert Frost, in a poignant poem, wrote ‘Two roads diverged in a wood, and I – I took the one less travelled by, And that has made all the difference’. The law of torts has been, in recent decades, sauntering along the path that the tort of negligence paved. One can see the beguiling appeal in doing so; after all, the tort of negligence, with its duty of care test, comes closest to unifying all the other torts and transforming the English Law of Torts to the English Law of Tort. Yet as appealing as this may be, when it comes to wrongful conception, the path less travelled—by reconceptualising wrongful conception as an actionable tort *per se*—is the ‘right’ approach. It will restore intellectual clarity back to a confused and muddled area of law. It will signify a protection of the mother’s autonomy by focusing on it more directly, rather than placing the doctor’s negligence in the limelight. It will, additionally, avoid the quagmire on actionable damage and public policy that the tort of negligence engenders. If the law of torts chooses to walk down this path, foreign though it may be, it will make all the difference.

Dividing the Single Indivisible Transaction: Balancing the Interests of Mortgagees and Innocent Occupants

Vincent Ooi¹

I. INTRODUCTION

When a mortgagee provides the funds necessary to purchase a property, the law protects the priority of the mortgagee's security interest from any other interests in the property created after that property was purchased (unless consented to by the mortgagee). The House of Lords affirmed this principle in *Abbey National Building Society v Cann* ('*Cann*'),² and rejected the technical argument based on a *scintilla temporis*³ in favour of the policy argument concerning the need to ensure that housing loans are available and affordable. The effect of this principle is to allow the mortgagee to enforce the security interest in the property in the event of the mortgagor's default, even if there are individuals with an equitable interest in the property and who are in actual occupation.

The loan, the purchase of the property and the creation of the mortgage on the property are collectively considered by the law to be a single indivisible transaction, preventing any equitable interest from arising at any point between these stages. This is the first sense in which the transaction is said to be indivisible. The transaction is also indivisible in another sense, for the burden of the mortgagor's default on the loan is subject to no apportionment in the event that he fails to meet his payment obligations. The mortgagee is allowed to enforce the security interest and the equitable interests of the occupants of the property are irrelevant at this stage. Yet, there are situations where one may question whether this leads to a fair outcome. Where the default on the mortgage occurs due to the fraud of the mortgagor, and the mortgagee ought to have known or suspected this, it is arguable that the 'relatively innocent' holder of the equitable interest should not have to suffer the full burden of the consequences of the default. This was the situation facing the House of Lords in *Re North East Property Buyers Litigation* ('*Scott*').⁴

The term 'one indivisible transaction' was used multiple times in *Scott*. It is reasonable to think that their Lordships were concerned with the application of priority rules when they use the term. This article focuses on the two senses of 'indivisibility' of the single indivisible transaction and the *obiter dicta* of Baroness Hale in *Scott*, where she advanced arguments for adopting a less rigid approach to these two issues. For indivisibility in the first sense, Baroness Hale argued that the indivisibility of the transaction 'depends on the facts'.⁵ As for indivisibility in the second sense, her Ladyship introduced the idea of 'innocence as a comparative concept' and insisted on a 'middle group'.⁶ This suggests her desire for the law to develop and allow division of the burden on the basis of relative fault.

¹ Trinity College. I am very grateful to Elizabeth Drummond for all her insightful comments and guidance. I would also like to thank the editors of the Oxford University Undergraduate Law Journal for reviewing this article.

² [1991] 1 A.C. 56 (HL).

³ Literally: 'a spark of time'. The legal meaning of the term will be described in detail later in the article.

⁴ [2014] UKSC 52, [2014] 3 W.L.R. 1163. The case is also commonly referred to as *Scott v Southern Pacific Mortgages Ltd*, being the only case of ten which was still outstanding when it reached the Supreme Court.

⁵ *Scott* (n 4) [115].

⁶ *ibid* [122].

Ultimately, it will be shown that while her Ladyship's arguments may seem appealing, they represent a considerable deviation from the law at the present time. The apportionment of the burden of the mortgage default is essentially the key issue of concern as the 'single indivisible transaction' itself remains a valuable policy tool. This article will propose two statutory⁷ measures to allow the courts the discretion to spread the consequences of the fraud between the mortgagee and the holder of any equitable interest in the property. It is hoped that this will enable the courts to achieve fairer outcomes in such situations. However, it is acknowledged that until there is statutory intervention, there is little that the courts can do to help the innocent occupants within the existing framework of the law. *Scott* is unlikely to be the last word on this issue and it remains to be seen how 'indivisibility' in these two senses will be challenged and developed by both statutory and judicial means.

A. THE BACKGROUND: CANN

The key facts of *Cann* are straightforward enough.⁸ George Cann obtained a loan from the Abbey National Building Society ('AN') to buy a leasehold house. This loan was secured by a mortgage on the property. Daisy (George's mother) and Abraham Cann then moved into the house. Eventually, George defaulted on the payments of the loan and AN sought possession of the house. By way of defence, Daisy claimed that due to her financial contribution to a property previously bought jointly by George and herself, and given George's assurance that she would always have a roof over her head, she had an equitable interest in the property. She then claimed that, by virtue of her actual occupation of the property, her equitable interest took priority over AN's security interest.⁹ The House of Lords decided that the loan, the purchase of the property and the creation of the mortgage on the property were all a single indivisible transaction,¹⁰ rejecting the claim that Daisy's equitable interest arose before the house was encumbered by the mortgage.

It is reasonable to think that their Lordships considered the potential consequences of their decision on the availability and affordability of home loans. While accepting that the *scintilla temporis* might logically exist, they chose to give effect to 'reality' and held that the acquisition of the legal estate and the legal charge (mortgage) were 'simultaneous and indissolubly bound together'.¹¹ The interests of the individual holding the equitable interest in the property were subordinated to the public interest in the certainty of home mortgage transactions and in the availability of such mortgages.

II. DEVELOPMENT OF THE SINGLE INDIVISIBLE TRANSACTION PRINCIPLE

The single indivisible transaction principle was developed as a response to the argument for the existence of a *scintilla temporis*. The starting point to understanding the principle would undoubtedly be the *Cann* case itself where arguments for both sides were advanced. Here, the House of Lords approved the single indivisible transaction principle as being consistent with 'reality' and declared the *scintilla temporis* to be no more than 'a legal artifice'.¹² With the authority of the House of Lords, the issue was largely settled. Attempts to work around it were largely limited to trying to distinguish *Cann* rather than overrule it. In this sense, *Scott* is no exception, with their Lordships unanimously approving the principle.¹³

⁷ It is noted that Baroness Hale herself seemed to envisage a purely statutory solution, and noted that she was glad that the Law Commission would review the Land Registration Act 2002. See *Scott* (n 4) [122].

⁸ *Cann* (n 2) 59-61.

⁹ Daisy invoked s 70(1)(g) of the Land Registration Act 1925, which is *in pari materia* with Schedule 3 Paragraph 2 LRA 2002. This will be discussed in detail later in the article.

¹⁰ *Cann* (n 2) 92.

¹¹ *ibid.*

¹² *ibid* 93.

¹³ *Scott* (n 4) [114].

A. THE OVERRIDING INTEREST OF ACTUAL OCCUPATION

Consider three parties: A, the mortgagor of a property; B, the mortgagee of a property; and C, the owner of a beneficial interest in the property. C may have gained the beneficial interest from A in a variety of ways, whether by contribution to the purchase price of the property, through proprietary estoppel, by contract, or various other means. In order to purchase the property, A takes a loan from B and grants B a legal charge over the property as part of the mortgage agreement. A then defaults on the payments of the loan and B seeks possession of the property.

The only way that C can defeat B's claim to possession is by showing that he has an interest which has priority to B's legal charge. However, s 29(1) of the Land Registration Act 2002 ('LRA 2002') has the effect of postponing the priority of all unprotected interests in the land at the point of the mortgage to immediately after the legal charge. Essentially, if an interest is not protected, the legal charge will have priority to it.

C's interest will be protected if it is a registered charge or the subject of a notice in the register.¹⁴ Alternatively, C's interest will also be protected if C is in actual occupation of the property and thus can prove that he has an overriding interest. The latter mechanism relies on s 29(2)(a)(ii) and Schedule 3, Paragraph 2 of the LRA 2002. If C can show that he has a proprietary interest,¹⁵ is in actual occupation of the property and does not fall foul of the additional exceptions in Schedule 3 Subparagraphs 2(b) and (c),¹⁶ his interest will not be postponed in priority to after the execution of the legal charge. This interest can potentially bind B and defeat his claim to possession of the property.

The crucial question then becomes: when did C's beneficial interest arise? If it arose after the legal charge was granted, then there is no question of changing its priority, and the legal charge has priority to it. If, however, it arose before the legal charge was granted, then the actual occupation mechanism as previously described may help maintain its priority as against the legal charge.

B. UNDERSTANDING THE SCINTILLA TEMPORIS

The House of Lords in *Cann* held that before the property was transferred to A, C could not possibly have had any proprietary interest in the property.¹⁷ A himself did not have any proprietary interest in the property and, consistent with the *nemo dat* principle,¹⁸ any interests he might have given to C were at best personal rights enforceable only against himself (and therefore incapable of binding third parties).¹⁹ It is only when A received the property that the equitable rights held by C could be 'fed' by the acquisition of the legal estate and become proprietary rights.²⁰ Depending on when the estoppel was 'fed', C might have a proprietary interest which had priority to the legal charge.

¹⁴ s 29(2)(a)(i).

¹⁵ *National Provincial Bank Ltd v Ainsworth* [1965] AC 1175 (HL) held that a purely personal right against A will not suffice.

¹⁶ The actual occupation mechanism will be defeated if the following can be shown under subpara 2(b) that an inquiry was made of C and he failed to disclose the right when he could reasonably have done so. Alternatively, under subpara 2(c), the mechanism is defeated if C's occupation would not have been obvious on a reasonably careful inspection of the land at the time of the disposition and if B did not have actual knowledge of his interest.

¹⁷ *Cann* (n 2) 92.

¹⁸ *Nemo dat quod non habet* (literally: 'no one gives what he does not have').

¹⁹ *Scott* (n 4) [50]. Lord Collins here explains the decision in *Cann*.

²⁰ *Cann* (n 2) 89–90. The mechanism is explained in detail in *Scott* (n 4) [71].

In *Cann*, counsel for C advanced the argument that ‘the transaction necessarily involved conveyancing steps which [...] must be regarded as taking place in a defined order, so that there was a ‘*scintilla temporis*’ between the purchaser’s acquisition of the legal estate and the creation of the society’s charge during which the estoppel could be fed’.²¹ In other words, there must have been a moment in time, however brief, where A held the fee simple of the property, unencumbered by the legal charge, after which the grant of the legal charge took effect. This moment in time was described as the *scintilla temporis* (the grant of the legal charge must necessarily have been executed after A received the fee simple of the property, for he could not grant it without first having a proprietary interest in the property).²²

If the *scintilla temporis* did indeed exist, then the estoppel would have been fed in that brief moment and C would have acquired a proprietary interest prior to the granting of the legal charge. The actual occupation mechanism as described earlier would then prevent s 29(1) LRA 2002 from postponing the priority of C’s interest to after that of the legal charge. C would be able to assert his proprietary rights against B, resisting B’s claim for possession.

C. REJECTION OF THE SCINTILLA TEMPORIS: THE SINGLE INDIVISIBLE TRANSACTION

Their Lordships in *Cann* acknowledged the ‘attractive legal logic’ of this argument but ultimately rejected it as ‘flying in the face of reality’.²³ Lord Oliver gave two main reasons for deciding that the acquisition of the legal estate and the legal charge were ‘simultaneous and indissolubly bound together’. The first reason was that the acquisition of the legal estate was entirely dependent upon the funds from B. It would arguably not be fair to subordinate B’s security interest to C’s equitable interests where C would never have obtained the equitable interest without B’s loan in the first place. The second reason given was that in most cases, conveyancing practice was that the documents for the legal charge would have been executed before the transfer of the legal estate to A (in terms of the sequence of execution of documents, not their legal effect). To say that the law logically required a different ordering of the stages would be to ignore what actually happens in practice and to insist on a technicality.

Lord Oliver then neatly summed up the effect of the single indivisible transaction, explaining that A ‘never in fact acquires anything but an equity of redemption’.²⁴ If A receives the legal estate when it was already burdened with the legal charge of B, then he cannot possibly go on to give any right to C which has priority over that of B’s legal charge. Accordingly, C has no interest which priority might be protected by the actual occupation mechanism. The legal charge has priority over C’s equitable interests and B’s claim for possession will succeed.

III. DEVELOPMENTS IN THE LAW AFTER CANN

A. THE BACKGROUND: SCOTT

In order to understand *Scott* fully, it is necessary to appreciate the context in which the cases arose. ‘Sale and rent back’ schemes (‘SRB’) are arrangements designed and marketed by private companies.

²¹ *Cann* (n 2) 90.

²² *ibid* 101.

²³ *ibid* 92.

²⁴ *ibid* 93.

The company arranges purchases a property from the homeowner using a loan from a bank secured by the property itself. As the company promises to give the homeowner an extended lease (sometimes a life interest) in the property, the homeowner is attracted by the prospect of obtaining quick money while still being able to reside in the property. In return, the company charges the homeowner a hefty 'fee', deducted from the purchase price of the property.²⁵

Even if the commercial viability of such schemes were not seriously questionable, the arrangement has some serious problems. First of all, in order to get the loan from the bank, the company is normally asked to declare that there are no outstanding leases on the property beyond 12 months.²⁶ This is standard bank policy and the companies would have to make a false declaration in order to get the loan in the first place. Secondly, there is a nasty tendency of the companies to suddenly abscond once they received their 'fee'.

Scott is a typical SRB case, with the original homeowner, Rosemary Scott ('RS'), effectively agreeing to sell her house to Ameer Wilkinson ('AW') at a price substantially below the market rate, having been promised the right to indefinite occupation at a discounted rent. Ameer took out a loan from Southern Pacific Mortgages Ltd ('SP'), using the house as collateral and then promptly defaulted on the payments and disappeared.

Our previous model involving A, B and C remains valid, though with some additional information. C is a homeowner who sells the property to A. The beneficial interest that C obtains from A in this case results from representations made by A that he will grant a lease to C so that he can stay in his own property even after the sale (raising a proprietary estoppel). As before, A then defaults on the mortgage payments and, in this case, disappears.

B. THE SUPREME COURT IN SCOTT CONSIDERS CANN

The facts in *Scott* were different from those in *Cann* and the Supreme Court considered two main questions:²⁷

- 1) Whether A could grant proprietary, as opposed to merely personal, rights to C before he acquired the legal estate; and
- 2) Even if C had proprietary rights, whether *Cann* applied to prevent him from asserting an equitable right that had arisen only on completion.

The key difference between the two cases lies in the contract signed in *Scott*. Counsel for RS in *Scott* argued that her interest arose pre-completion and that this enabled them to distinguish *Cann*. They effectively argued that since her interest existed during the period between contract and conveyance, it slipped in before the legal charge took effect.

The argument is based on the existence of a vendor-purchaser constructive trust. The roots of a vendor-purchaser constructive trust can be found in the case of *Lysaght v Edwards*,²⁸ where Jessel MR held that once there is a 'valid contract for sale, the vendor becomes in equity a trustee for the purchaser of the estate sold, and the beneficial ownership passes to the purchaser'.²⁹

²⁵ It is noted that since 1 July 2009, the Financial Conduct Authority has regulated firms that offer SRB. The interim scheme ran from 1 July 2009 and was replaced by the current scheme on 30 June 2010. Financial Conduct Authority, 'Sale and rent back' <<http://www.fca.org.uk/firms/financial-services-products/mortgages/sale-and-rent-back>> accessed 1 July 2015.

²⁶ *Scott* (n 4) [5].

²⁷ *ibid* [10].

²⁸ (1876) 2 Ch D 499 (Ch).

²⁹ *ibid* 506.

Applying this to the facts of *Scott*, AW would have had a beneficial (and proprietary) interest in the property before completion, with (rather counterintuitively on the facts) RS as the trustee of that interest. AW's promise to RS would then be capable of giving RS a proprietary interest, since AW would have had a sufficient interest in the property from which to make that gift at the time.

The argument based on a vendor-purchaser constructive trust was rejected by both Lord Collins and Baroness Hale, though their reasoning differed. Lord Collins held that the 'contract, conveyance and mortgage are indivisible'³⁰ and that 'it would be wholly unrealistic to treat the contract for present purposes as a divisible element in this process'.³¹ Thus, the interest granted by A to C can only be personal in nature until such point where A is conveyed the legal estate. Then the estoppel is 'fed' and C's beneficial interest gains proprietary character.³² Lord Collins then held that *Cann* would apply to *Scott*, with the effect that the acquisition of the legal estate and the grant of the charge would be one indivisible transaction.³³ Baroness Hale rejected the constructive trust argument on the basis that 'the purchaser of land cannot create a proprietary interest in the land, which is capable of being an overriding interest, until his contract has been completed'.³⁴

C. DOES SCOTT GO FURTHER THAN CANN? STILL AN INDIVISIBLE TRANSACTION

From this point onwards, everything their Lordships considered is strictly speaking *obiter dicta*, since the legal issues in *Scott* were resolved when Lord Collins held that A could not grant proprietary rights to C before he acquired the legal estate. The case was decided solely by answering the first question, rendering the answer to the second question (with all the discussion of indivisibility) inconsequential.³⁵ However, it will soon be apparent that the *obiter dicta* in this case provides much more fertile ground for analysis and development of the law than the *rationes*, which does little more than affirm *Cann*.

Lord Collins went on to consider whether the single indivisible transaction analysis extends to situations where the equitable interest is said to arise at the time of the contract of sale.³⁶ He affirmed the judgment of Aldous LJ in *Nationwide Anglia Building Society v Ahmed*³⁷ that it was implicit in *Cann* that contract, conveyance and mortgage are indivisible.³⁸ Lord Collins conceded that the contract of sale has separate legal effects, but argued that it would be wholly unrealistic to treat the contract for present purposes as a divisible element in the process.³⁹ Lord Collins thus seemed to think that there is in substance a single indivisible transaction which effectively prevents any beneficial interest of C from ever having priority over B's legal charge in such 'mortgage situations', regardless of the time between each individual component of the transaction.

³⁰ *Scott* (n 4) [85].

³¹ *ibid* [87].

³² *Scott* (n 4) [71]–[72].

³³ *ibid* (n 4) [79].

³⁴ *ibid* [122].

³⁵ It is noted that Lord Wilson and Lord Reed expressly acknowledged that the indivisibility of the contract from the conveyance and the mortgage was not part of the reasons for the decision. See *Scott* (n 4) [123].

³⁶ *Scott* (n 4) [70]. In doing so, he accepts that his statements are made *obiter*.

³⁷ (1995) 70 P & CR 381 (CA).

³⁸ *Scott* (n 4) [85].

³⁹ *ibid* [87].

IV. THE FIRST SENSE OF 'INDIVISIBILITY': DIVIDING THE TRANSACTION ON THE BASIS OF THE FACTS

Baroness Hale's *dicta* in *Scott* are by far the most interesting and have the most potential for development. As far as the *rationes* of the case went, she agreed with Lord Collins' judgment. However, with respect to Lord Collins' *dicta* on the single indivisible transaction, she preferred to adopt a different and more flexible approach.⁴⁰ It is noted that she had the support of the majority of the Supreme Court on this point, with Lord Wilson and Lord Reed supporting her⁴¹ and Lord Collins and Lord Sumption taking a different position.

There are two points being considered here. Baroness Hale seemed firm on the first point, namely, that the contract, conveyance and mortgage do not constitute a single indivisible transaction. She insisted that there is no 'tripartite transaction to which vendor, purchaser and lender are all party'.⁴² In addition, she considered that to insist that these factors constitute a single indivisible transaction would be to 'fly in face of the facts' and 'create confusion'.⁴³ On the other hand, she seemed willing to accept that the conveyance and mortgage may constitute a single indivisible transaction, albeit that this depends on the facts.⁴⁴ On the first point, Baroness Hale presented a persuasive argument⁴⁵ and this article will not address it further.

A. DEPENDING ON THE FACTS

On the second point, the single indivisible transaction rule in *Cann* was thought to apply to all cases where A took out a mortgage with B to purchase the property. However, Baroness Hale held that this need not be the case in every situation. She held that 'not all conveyances are indivisible: it depends on the facts'.⁴⁶ While we were not told what characteristics will persuade the court to depart from the general finding of an indivisible transaction, Baroness Hale provided us with at least one situation. She confirmed that 'if the mortgage takes place sometime after the conveyance, there may be a period during which the purchaser owns the land without encumbrances'.⁴⁷ The difficulty with this statement is that it provides little guidance as to how long 'sometime' might be.

Baroness Hale's assessment of the *Cann* line of cases may shed some light on what is required to show that there was no indivisible transaction. She held that in the cases of *Coventry*⁴⁸, *Woolwich*⁴⁹, *Piskor*⁵⁰ and *Cann* the conveyance and the mortgage were virtually contemporaneous and the mortgage loan was required to complete the transaction.⁵¹ Her Ladyship also accepted that in *Ahmed*⁵² there was an indivisible transaction because the transactions all took place on the same day and each of the participants knew what the terms of the arrangements were.⁵³

⁴⁰ *Scott* (n 4) [95].

⁴¹ *ibid* [123].

⁴² *ibid* [119].

⁴³ *ibid* [120].

⁴⁴ *ibid* [115].

⁴⁵ *ibid* [116], [119]–[120]. The most persuasive argument is that the contract to sell the land is an entirely separate matter from the mortgage and the conveyance. There is no reason for the default position to be that they all occurred simultaneously. Technically speaking, for the single indivisible transaction rule to exist, B and C need not even know that the other exists, let alone that A and B have made a contract together.

⁴⁶ *Scott* (n 4) [115].

⁴⁷ *ibid*.

⁴⁸ *Coventry Permanent Economic Building Society v Jones and Others* [1951] 1 All ER 901 (Ch).

⁴⁹ *Woolwich Equitable Building Society v Marshall* [1952] Ch 1 (HL).

⁵⁰ *Church of England Building Society v Piskor* [1954] Ch 553 (CA).

⁵¹ *Scott* (n 4) [109].

⁵² *ibid*.

⁵³ *ibid* [121].

There is also Baroness Hale's assessment that *Universal Permanent Building Society v Cooke*⁵⁴ may not have been wrongly decided, indicating that her Ladyship agreed with the principles laid down in *Cooke*. Therefore, in order to understand the factors that determine whether an indivisible transaction exists, it is useful to analyse *Cooke*.

The facts of *Cooke* are briefly stated as follows. Cooke leased out her flat to Gavine on a weekly tenancy before obtaining the fee simple of the property. Cooke was conveyed the fee simple of the property on 28 December 1948 and mortgaged the property the following day. She subsequently defaulted on her payments of the loan and the bank sought possession. Gavine argued that she had a tenancy by estoppel and sought to resist the order for possession. The Court of Appeal held that there was nothing to suggest that the conveyance and the mortgage were part of a single transaction.⁵⁵ It found itself bound on the facts of the case to deem the two transactions as being two independent transactions, the mortgage taking place a day after the conveyance.⁵⁶

Regarding the case of *Cooke*, Baroness Hale seemed to think that it can be argued that the transaction was not indivisible because the transaction was not contemporaneous (mortgage executed one day after the conveyance) and because there was no evidence that the mortgage was required for the purchase.⁵⁷ Baroness Hale did not commit herself to a position, though she noted that it was likely that the House of Lords in *Cann* knew about *Cooke* and did not expressly overrule it or even mention it.⁵⁸

This suggests three factors which militate towards a finding of an indivisible transaction: 1) The transactions took place contemporaneously (on the same day), 2) the mortgage loan was required to complete the transaction; and 3) each of the parties knew what the terms of the arrangements were.

The first possible distinguishing factor of *Cooke* is intellectually unsatisfying. To argue that a transaction is divided and not a single whole simply because the constituent procedures were completed a single day apart is a highly doubtful proposition. If the law is willing to accept that the *scintilla temporis* argument is a mere technicality, then it should similarly be willing to look past a single day's delay and assess the substance of the transaction. Lord Collins also seemed unconvinced that counting the number of days between the constituent procedures is a good way to go about assessing the transactions as well. He noted that while in *Ahmed*, the contract and conveyance were executed on the same day, the (indivisible transaction) analysis is not dependant on that.⁵⁹

The second possible distinguishing factor of *Cooke* may be more promising. If the mortgage is not required for the purchase of the property, then the conveyance and the mortgage will not be part of a single indivisible transaction. This factor seems to make perfect sense. The very reason for insisting on a single indivisible transaction in *Cann* itself was because the acquisition of the legal estate is entirely dependent on the provision of funds from the lender.⁶⁰ However, what this really suggests is not that *Cooke* is an exception to the single indivisible transaction rule in *Cann*, but that *Cooke* was never caught by the rule in the first place. *Cann* only applies where the mortgage is required to acquire the property. This point has been uncontroversially accepted since *Cann* itself.

⁵⁴ [1952] Ch 95 (CA).

⁵⁵ *Cooke* (n 54) 101.

⁵⁶ *ibid.*

⁵⁷ *Scott* (n 4) [110].

⁵⁸ *ibid.*

⁵⁹ *Scott* (n 4) [85].

⁶⁰ *Cann* (n 2) 92.

The third possible distinguishing factor is reasonable. If the innocent occupant in particular knew about the terms of the arrangements, one might reasonably expect him to raise the issue of his equitable interests to the mortgagee. Remaining silent about the false pretences on the basis of which the loan was obtained (if not for the false pretences, the loan would not be given in such circumstances) leaves open the accusation that one is an accomplice to a fraud. However, though knowledge of the terms of the arrangement is a strong factor for finding that there was a single indivisible transaction, it is not clear that the reverse applies. The existing framework and rationale for the single indivisible transaction is largely based on conveyancing practice and policy arguments, not fault arguments. It is only when there is sufficient fault to vitiate the contract that it becomes relevant. Even if the innocent occupant did not know the full terms of the arrangements, this does not affect the realities of conveyancing practice or policy justifications in any way. If fault arguments are to be introduced, it is argued that they are better considered at the stage of apportioning the burden of the transaction rather than when dealing with the quite technical point regarding whether the transaction is indivisible or not. These arguments will be considered later.

After analysing all the cases, Baroness Hale's statement that things will be 'dependent on the facts' really provides us with no more information than the statement itself. We have no guidance as to when this exception might apply or whether there might even be any cases where the exception could apply at all. At its narrowest interpretation, we might construe the statement as stating that the single indivisible transaction rule does not apply to cases where the mortgage is not required to purchase the property; a largely redundant statement.

Of course the statement can be interpreted in a far wider manner than that. With no guidance as to what 'dependent on the facts' mean, the courts now have the potential to make decisions in a more flexible manner. However, bearing in mind that Baroness Hale's statements are non-binding *dicta* and that they do not provide much guidance, it is highly unlikely that any subordinate court would be too keen to prioritise it over *Cann*, which is technically very much still good law. It seems that this attempt at 'division' of the single indivisible transaction is unlikely to have any significant impact on the law.

V. THE SECOND SENSE OF 'INDIVISIBILITY': DIVIDING THE BURDEN OF THE TRANSACTION

At the end of her judgment, Baroness Hale expressed some uneasiness about the decision. She noted that the law insisted on an 'all or nothing' approach.⁶¹ Either C could prove that there was some reason in equity to restrain B from enforcing the security as against C's right altogether, or B would be able to totally disregard C's rights and enforce its security. As far as equitable remedies in such situations go, the law has been unwilling to depart from this 'all or nothing' approach and apportion the burden of loss between the two (relatively) innocent parties once the third party has either absconded or been made insolvent. In *TSB Bank Plc v Camfield*⁶² the court insisted that it had no power to award partial rescission.⁶³ To the court, rescission was an all or nothing process and the right of the representee and not of the court. The court could not grant equitable relief to which terms may be attached.⁶⁴

⁶¹ *Scott* (n 4) [122].

⁶² [1995] 1 WLR 430 (CA).

⁶³ *ibid* 436.

⁶⁴ *ibid* 438-439.

A. UNDERSTANDING THE HARSHNESS OF THE LAW

In 2013, there were more than 220,000 claims for possession made in England and Wales⁶⁵ as mortgagors defaulted on their loans. Since *Cann* was decided, there have been other cases where the security has been successfully enforced by the banks with no compensation to the occupants of the properties with beneficial interests. While the loss of a family home is one which undoubtedly evokes a sense of sympathy, the courts have long since steeled themselves against prioritising the few individuals before them, aware of the consequences on the rest of society if they did so. It is thus interesting that the Supreme Court, the barristers involved in the case, and the majority of authors writing on this issue have felt that the victims in *Scott* and the ‘sale and rent back’ cases deserve special sympathy.

The special element which seems to evoke such sympathy is likely to be the fact that the homeowners affected by such schemes were victims of fraud. The victims were promised money and the right to stay for extended periods in their properties (sometimes for life). They never for a moment thought that they would be evicted from their own homes. There is a fundamental (and emotive) difference between on the one hand moving into a new property with the expectation of a proprietary right and then being denied performance of that expectation, and being evicted from one’s own home after receiving express assurances that this would not happen on the other. Lord Collins himself noted that even if the victims receive monetary compensation, it does not change the fact that they have lost their homes.⁶⁶

B. IS THE LAW REALLY THAT HARSH?

The full picture is a little more complicated than that. Thomas Tyson argues that the whole structure of a ‘sale and rent back’ transaction is a ‘preposterous’ and ‘eminently resistible’ one.⁶⁷ The vendor releases some equity in the property, but receives little in return.⁶⁸ Tyson compares the deal to one to exchanging a cow for some magic beans,⁶⁹ implying that the promised leases and/or bonus payments are simply too good to be true. One cannot help but feel that *caveat emptor* may be an apt expression in these cases.

As to monetary compensation, it is unlikely that the mortgagors who organised the whole scheme will be found or indeed be solvent on any level. But that does not mean that the victims are totally bereft of any remedy. Lord Collins noted that the victims may have claims against the Solicitors’ Compensation Fund⁷⁰ and it is always possible that professional negligence claims may be brought. It is possible that there will be some negligence or wrongdoing on the part of the solicitors, whether it lies in being accomplices to the mortgagors or failing to give proper advice to the victims. All things considered, there is probably still a need for the law to intervene to help these victims as these remedies are inadequate. However, they should be considered as part of the overall compensation framework.

⁶⁵ Ministry of Justice ‘Mortgage and Landlord Possession Statistics Quarterly: October to December 2013’ (Ministry of Justice Statistics Bulletin, 13 February 2014) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/279342/mortgage-landlord-possession-statistics-oct-dec-2013.pdf> accessed 1 July 2015.

⁶⁶ *Scott* (n 4) [24].

⁶⁷ Thomas Tyson ‘Magic beans for that cow? *Scott v Southern Pacific Mortgages* [2014] UKSC 52’ (Zenith Chambers News and Articles, 24 October 2014) <http://www.zenithchambers.co.uk/cms/document/scott_v_southern_pacific_article.pdf> accessed 1 July 2015.

⁶⁸ *ibid*

⁶⁹ *ibid*

⁷⁰ *Scott* (n 4) [24].

C. INNOCENCE AS A COMPARATIVE CONCEPT

When arguing for a middle ground approach between the ‘all or nothing’ approach of the current law, Baroness Hale advanced an approach that would use relative fault to apportion the burden of the loss between B and C.⁷¹ She argued that there might be a point where a vendor who has been *tricked* out of her property can assert her rights even against a subsequent purchaser or mortgagee.⁷² From the point of view of the lenders, she argued that there might be a point where the claims of lenders who have failed to heed the obvious warning signs that would have told them the borrower was not a good risk should be postponed to those of vendors who have been made promises that the borrowers cannot keep.⁷³ Baroness Hale then expresses her approval of the plans of the Law Commission to review the impact of fraud in the LRA 2002.⁷⁴ These statements make it apparent that her Ladyship was concerned mainly with fraud as the ground for granting equitable relief.

Baroness Hale argued that lenders may relatively be at fault when they fail to heed the ‘obvious warning signs’ that the borrower was not a good risk.⁷⁵ This statement has the potential to produce very controversial outcomes. It seems to suggest that even if B has no knowledge of the situation, vulnerabilities or even existence of C, B’s actions may constitute sufficient ‘fault’ to make it bear some of the losses. It is argued here that even if we are to use a concept of ‘comparative innocence’, any ‘fault’ on the part of the parties must be sufficiently linked to the other party they are splitting the losses with. So, it is less controversial to suggest that if B knew or should have known that C would potentially suffer loss from his actions, we might want to impose duties to C on B.⁷⁶ Undue influence works in a similar manner and we can understand how the concept of constructive notice works.⁷⁷ It is far more controversial, however, to suggest that B, potentially having no knowledge of the existence of C, should effectively owe a duty to C to heed the ‘obvious warning signs’ given by A. In equity (fiduciary duties) and in tort (negligence) it is rare for one to owe a duty to a party whose existence one is unaware of, even as part of a class of persons. Simply put, if B does not know that C exists, B can act solely in his own interests without having to worry about the interests of any third parties.

It would be a lot less controversial to suggest that Baroness Hale’s ‘comparative innocence’ concept should be restricted to cases where B knows or at least suspects the existence of C. On the other hand, while C inevitably knows of the existence of B in such cases, that need not be a requirement. One need not know of a duty’s existence to be owed a duty.

In practice, lenders are rather keen to ensure that they get good security for their loans. It is most unlikely that they will grant loans to borrowers who are ‘obviously not good risks’. Furthermore, standard mortgage terms generally permit only shorthold tenancies of fixed term no more than 12 months.⁷⁸ If there were clear signs that A was making a false declaration on the form or intending to break this term, it is unlikely that any bank would have granted him the loan. The prospect of not having the right to vacant possession is not a welcome one, and few lenders would knowingly walk into a transaction like that. The whole reason why the transaction occurred in the first place was due to the lack of full disclosure by A to B. Hence, the kind of transactions that are likely to arise in practice are cases of constructive notice, where the mortgagee should have known about the risks in granting the mortgage but still went ahead. This may occur due to poor design of the system, human error or just plain negligence (in the ordinary sense of the word).

⁷¹ *Scott* (n 4) [122].

⁷² Emphasis added.

⁷³ *Scott* (n 4) [122].

⁷⁴ Law Commission, *Twelfth Programme of Law Reform* (Law Com No 354, 2014) paras 2.13- 2.16.

⁷⁵ *Scott* (n 4) [122].

⁷⁶ Indeed this seems to be the rationale behind the exceptions in Schedule 3 paragraph 2, LRA 2002.

⁷⁷ See *Barclays Bank Plc v O'Brien* [1994] 1 A.C. 180 (HL) and *Royal Bank of Scotland v Etridge (No 2) and other appeals* [2001] UKHL 44, [2002] 2 AC 773. *Etridge* is the leading case.

⁷⁸ *Scott* (n 4) [5].

D. EXISTING COMMON LAW REMEDIES

The closest existing common law principles potentially available to protect the victims are those of unconscionability, undue influence and misrepresentation. However, none of them really seems applicable to the situation, and even so, on the authority of *TSB v Camfield*, all of them entail ‘all or nothing’ remedies. As there is no direct contractual relationship between B and C, and quite possibly no dealings at all between them at all, a successful plea of any of these vitiating factors would merely render the contract between A and C voidable. As the legal charge has already been granted to B, this would not assist C.

While the doctrine of constructive notice in undue influence may seem more promising in such cases, the fact remains that B and C have no contractual relationship and constructive notice has not yet been expanded to include cases where there is no contractual relationship between the victim and the party with supposed constructive notice. The contracts are between A and B, and between A and C. Even if B knows or ought to have known about C’s situation, there is no scope for the application of undue influence principles. What Baroness Hale was considering is a form of action which can work even when the only link between B and C is their link with A. The common law currently does not provide a remedy to C against B in such a situation. A statutory solution will indeed be required. A ‘proposed statute’ will be considered later in this article. Before it is covered in detail, the necessary characteristics of this hypothetical ‘proposed statute’ will be discussed.

E. APPORTIONMENT OF LOSSES: A LESSON FROM HISTORY

If such a situation seems familiar, it is because this is not the first time that the courts have had to deal with the issue of apportionment of the burden of loss between two (relatively) innocent parties. Roughly half a century ago, Devlin LJ expressed his dissatisfaction with the ‘all or nothing’ approach in *Ingram v Little*,⁷⁹ a case concerning mistake of identity involving a fraudster who then dropped out of the picture. As is currently being done with this issue, the issue then was taken up by the Law Commission, which published their response in their Twelfth Report (Transfer of Title to Chattels (Cmnd 2958, 1966)).⁸⁰ The idea of apportioning losses flowing from cases of mistake was rejected by the Law Commission. Diamond argues that the rejection was largely based on two grounds. The first ground was that a power of apportionment would result in uncertainty following from the grant of a wide and virtually unrestrained judicial discretion. The second ground was the insistence of the Law Commission that in cases where both parties were innocent, the loss should lie where it falls.⁸¹

A good half a century has passed since that report. Also, the subject matter here, being land, is likely to raise very different considerations.⁸² However, it is important for us to consider the reasons for the rejection of a very similar power to apportion losses in the law of mistake. The second ground of rejection of the proposal is unlikely to be an issue here. Baroness Hale implied that such a power should only be exercised where there is not only relative fault, but apparently a rather high level of relative fault on the part of the lenders.⁸³ As regards the first ground of rejection, certainty in the law is of paramount importance in ensuring that the home loans and mortgages system flows smoothly. The question we have to ask is: ‘will the proposed statute create an acceptable level of uncertainty this time?’

⁷⁹ [1961] 1 QB 31 (CA).

⁸⁰ Aubrey Diamond summarised and commented on the report under the Reports of Committees section of the *Modern Law Review*; see further (1966) 29 MLR 418.

⁸¹ *ibid* 414.

⁸² The difference in treatment of legal issues relating to land as compared with the other areas of private law is well established and (largely) well justified. For an in depth consideration of this point, see Peter Birks, ‘Before We Begin: Five Keys to Land Law’ in Bright and J Dewar (eds) *Land Law: Themes and Perspectives* (OUP 1998).

⁸³ *Scott* (n 4) [122].

It is tempting to argue that since ‘sale and rent back’ transactions are now regulated by the Financial Conduct Authority⁸⁴ and the market has in practice shut down,⁸⁵ we should compensate the victims of such transactions as it would be unlikely to affect the supply of home loans.⁸⁶ However, there are a significant number of fraud cases in property transactions each year. Any attempts to restrict the effect of the proposed statute to only ‘sale and rent back cases’ will be open to the fair criticism of being unprincipled. It is highly difficult to justify theoretically why victims of such schemes in particular will have a new remedy available to them while it is denied to the victims of other kinds of fraud in property transactions. The danger is that, in trying to do justice to a small group of people we end up distinguishing the law in unprincipled ways, or worse, changing the law such that it affects confidence in standard commercial transactions and does more harm than good.

On the other hand, if we were to expand the scope of the proposed statute to include fraud in property transactions in general, it would be difficult to control the level of uncertainty generated. Unlike transactions tainted by undue influence, it is impossible for lenders to effectively screen such transactions. In undue influence cases, there is at least some form of dealings between the lender and the victim making it potentially reasonable for us to attribute constructive notice to the lenders and thereby protect the victim’s interest should the lenders not discharge their duties in ensuring that the victim is properly legally advised. Fraud cases have a wide variety of forms and B and C need not even have met or had any dealings with each other at all in some of these situations.

F. STATIC AND DYNAMIC SECURITY

We are faced once again with the one of the core problems of land law, that of balancing static and dynamic security. Very briefly, static security involves a preference for protecting the title of the original owner against subsequent purchasers, whereas dynamic security involves protecting the subsequent purchasers instead. The decision to prioritise dynamic security was already made with the passing of the LRA 2002,⁸⁷ with Parliament weighing the options and accepting that there might be a few cases where property owners would be disadvantaged by the rules. The intention behind s 29 and, more broadly, the LRA 2002 itself, was that the old doctrines of notice (and particularly, constructive notice) should no longer apply.⁸⁸ Excepting the concession of overriding interests, buyers should simply be able to rely on the Register when making their decisions as to the property. The ‘relative fault’ that is considered when apportioning losses under the proposal is most likely to come from judgments as to what B knew or should have known about C and his situation.⁸⁹ This completely goes against the intention behind the 2002 Act as it effectively requires B to investigate C’s situation. This more or less resurrects the old doctrine of notice, something which was expressly removed by the 2002 Act.

⁸⁴ Since 2009, initially by the Financial Services Authority under s 19, Financial Services and Markets Act 2000. This is now governed by the Financial Conduct Authority.

⁸⁵ *Scott* (n 4) [2].

⁸⁶ Given that such transactions effectively no longer exist, even if the banks tightened their loans criteria for such transactions, it would not make much difference to the market.

⁸⁷ This is inherent in numerous provisions in the LRA 2002, including s 26, s 29, s 58. For a more in-depth consideration, see Amy Goymour [2013] CLJ 617.

⁸⁸ Law Commission, *Land Registration for the 21st Century: A Consultative Document* (Law Com No 254, 1998) paras 3.44 – 3.50.

⁸⁹ It was noted earlier that it would be controversial to allocate liability to B where B does not know of C’s existence.

It is noted that Baroness Hale and the Law Commission may have had different things in mind when her Ladyship made her comments about the Law Commission's plans. With the developments in how the courts now deal with the impact of fraud on the Land Register (see the *Malory*⁹⁰ and *Fitzwilliam*⁹¹ line of cases), it is likely that the focus of the Law Commission will be on fraud which results in void dispositions and its impact on the register and the application of Schedule 4 and Schedule 8 of the LRA 2002. Baroness Hale's statement may now prompt the Law Commission to look into this particular area, but it is likely that when the statement was made, the Law Commission intended to review the law on fraud pertaining to land registration rather than on the allocation of losses in cases of fraud. It is expected that the Report, when published, will still predominantly focus on the former issue.

In summary, it is argued that the fully understandable sympathy which Baroness Hale and the rest of the Supreme Court felt for the victims of 'sale and buy back' schemes should not be translated into any changes in the law. Parliament has already made its decision (twice, if we count the proposals on reforming the law of mistake) that an 'all or nothing' approach is justified on pragmatic grounds. The *dicta* of Baroness Hale do not seem to have added any substantial fresh ideas to this debate. One suspects that the Law Commission will make a critical and thorough assessment of the issue and come to the same conclusion which its predecessors did half a century ago.

G. MITIGATING THE HARSHNESS AND THE POWER TO MAKE DISCRETIONARY ORDERS

As noted earlier, the primary loss to the victims is likely to be their eviction from their property. Barring those cases where the solicitor lacked professional indemnity insurance or was truly free from any wrongdoing, the victims may receive compensation from either the Solicitors' Compensation Fund or as a result of professional negligence claims. This provides Parliament with an opportunity to assist these victims.

A potentially acceptable solution might be to confer upon the courts the power to exercise discretion to:

- 1) Freeze any attempts to take possession of or sell the mortgaged property until the conclusion of any proceedings by the victims to seek compensation from A and/or B (proceedings to be brought in a reasonable amount of time);⁹² and
- 2) Order a sale of the property to the victims at fair market value within a reasonable time from the conclusion of the abovementioned proceedings.

B should be able to claim a fair interest rate for the loan (which C should have to pay regardless of whether C eventually buys the property back or not), which C in turn should be able to claim in the compensation proceedings. The idea is for the victims to be in a position to use their compensation to purchase their property back. It is noted that since C did receive some money in the initial transaction (which has probably already been spent), C may need to take out a loan in order to make up for that shortfall and buy back the property. Unfortunately, if the victims were themselves substantially at fault and/or it is not possible to extract any compensation from the solicitors involved, the victims may have to bear the full losses themselves.

⁹⁰ *Malory Enterprises Ltd v Cheshire Homes (UK) Ltd* [2002] EWCA Civ 151, [2002] Ch. 216

⁹¹ *Fitzwilliam v Richall Holdings Services Ltd* [2013] EWHC 86 (Ch), [2013] 1 P. & C.R. 19

⁹² The solution would fundamentally be an equitable remedy and equitable laches would probably apply.

H. ALTERNATIVE MEASURES

While the idea of dividing the burden of the transaction between B and C has been largely rejected in this article, it is possible to envisage some models for doing so. There are two kinds of potential losses suffered. One is the loss of possession of the property (which cannot be apportioned), while the other is the pecuniary loss suffered due to the transaction. The two models based on tort and equity allow for C to be compensated by B, regardless of whether the solicitor is additionally liable for damages. C will not be allowed to recover an amount greater than his loss from B and the solicitor. Where relevant, B and the solicitor can then bring proceedings to apportion the losses between themselves.

Two measures will be proposed, one relying on a duty in equity and the other relying on a duty in tort. The two measures require different statutory provisions to be enacted. The proposed duty in equity would resemble the current approach used in undue influence cases. A statutory duty could be imposed on B; not to ensure that C was given independent legal advice as in undue influence cases, but to conduct reasonable investigations as to the situation of C and ascertain the nature of the contract between A and C. The duty would not apply where B was unaware of the existence of C, but B would be required to conduct reasonable investigations as to that point where he suspected that C (as a class of persons) might be involved.

If the duty was breached, the first step would be to determine the relative fault of the parties. Then, the losses due to each party should be calculated in proportion to such relative fault. A new statutory mechanism would be required for determining the relative fault and it is suggested that the existing Law Reform (Contributory Negligence) Act 1945 ('LRCNA') could be modified to fulfil this role. As it is impractical to expect the banks to retain the property indefinitely, the default position would be that C should get the fee simple of the property, but only if C pays B a sum corresponding to B's share of the losses. If C fails to pay this sum, then B would be allowed to enforce the security, but must then pay C a sum corresponding to C's share of the losses. This model requires three new statutory mechanisms: 1) the power to order the conveyance of the property to C; 2) the power and mechanism to determine relative fault; and 3) the power to apportion losses.

Alternatively, the tort model involves the creation of a statutory duty of care on the part of B. B would be under a duty to C to take reasonable steps to ascertain whether C exists if he has reason to suspect it, and then to conduct reasonable investigations as to the situation of C, including ascertaining the nature of the contract between A and C. The standards expected by the duty in equity and tort should be similar in practice.

If the duty was breached, the LRCNA 1945, s 1(1)⁹³ could be applied directly to the case. The interpretation of 'fault' in s 4⁹⁴ is wide enough to include relative fault as considered in this situation. It is unlikely that the facts of such cases will be sufficient to persuade the courts to grant an order of specific performance that B not act in contravention of C's beneficial interests in the property. Thus, a statutory provision is likely to be necessary to provide for discretion to exercise this power. As noted earlier, it is impractical to expect the bank to hold onto the property indefinitely. The default position would be that the courts would order that B convey the fee simple of the property to C; provided that C pays B a sum corresponding to B's share of the losses. If C fails to pay this sum, then B would be allowed to enforce the security, but must then pay C a sum corresponding to C's share of the losses. This model requires two new statutory mechanisms: 1) the creation of a statutory duty of care owed by B to C; and 2) the power to order the conveyance of the property to C.

⁹³ s 1(1): Where any person suffers damage as the result partly of his own fault and partly of the fault of any other person or persons, a claim in respect of that damage shall not be defeated by reason of the fault of the person suffering the damage, but the damages recoverable in respect thereof shall be reduced to such extent as the court thinks just and equitable having regard to the claimant's share in the responsibility for the damage.

⁹⁴ 'Fault' means negligence, breach of statutory duty or other act or omission which gives rise to a liability in tort or would, apart from this Act, give rise to the defence of contributory negligence.

The two models have some shortcomings. The equity model runs into the problem that *TSB v Camfield* still does not allow apportionment of loss, requiring express statutory intervention to get around this problem. The tort model has the advantage of utilising the well-established rules in the LRCNA but requires the creation of a statutory duty of care. In practice, their effects should be similar. Also, if the models are narrowly drafted such that they apply only to ‘sale and rent back’ transactions, they can be criticised on the grounds that they are unprincipled. If they are broadly drafted to include fraud in property cases in general, then they are likely to create too much uncertainty in the home loans and mortgages market. In addition, there are complications arising from the fact that both B and C expected A to hold onto the property and perform his obligations to the two of them. With A out of the picture, B cannot be expected to hold onto the property until C’s proprietary interest runs out. C similarly cannot be forced to buy the property. These models can deal with the allocation of pecuniary losses, but at best aid C in getting his property back. There are no guarantees at that point. Nevertheless, they represent models which could feasibly work if Parliament is willing to bear the costs mentioned above.

VI. CONCLUSION

At a superficial level, *Scott* seems to promise a departure from the orthodoxy of the single indivisible transaction. Where in the past conveyances and mortgages were part of an indivisible whole (for this category of cases), *Scott* suggests that they may now be divisible ‘depending on the facts’.⁹⁵ Where there was rigid insistence on an ‘all or nothing approach’, *Scott* now suggests a middle approach where burden of the losses of the transaction may be apportioned based on the grounds that ‘innocence is a comparative concept’.⁹⁶

However, once we look at the substance of those statements, we realise that they provide little by way of guidance for their application. As mere *obiter dicta*, it is unlikely that subordinate courts will attempt to rely on these statements in distinguishing the cases before them from the orthodoxy approved by no less than the House of Lords.⁹⁷ The lack of clear guidance as to the meanings of these statements coupled with very plausible narrow interpretations only makes their task more difficult. This means that *Scott* as it currently stands is unlikely to have any impact on the law. The single indivisible transaction seems to remain indivisible as ever.

It may be just as well that this is so. From both theoretical and pragmatic standpoints, the Supreme Court seems to have arrived at a reasonable decision in their judgments. One can feel a great sympathy for the victims involved while recognising that sympathy need not translate into legal rights or even evoke feelings of injustice. Apportioning losses between (largely) innocent parties is not an easy task and what the Law Commission and Parliament think as to its feasibility has already been established.

Objectively speaking, if the victims can obtain compensation for their losses,⁹⁸ then the loss of their property (while serious) may not constitute a grave injustice. If Parliament is willing to intervene, it is submitted that their focus should be on giving the victims an opportunity to buy their houses back (suffering some losses), rather than on attempting to apportion the losses between the parties. The former approach may actually stand a chance of being enacted into law. It is certain that we have not heard the last word on this issue, however one gets the general impression that the significant commercial implications on the home loans market mean that there will be much discussion on this issue, but very little actual change.

⁹⁵ *Scott* (n 4) [115].

⁹⁶ *ibid.* [122].

⁹⁷ *Cann* (n 2).

⁹⁸ *Scott* (n 4) [24].

Prisoners' votes, referendums, and common law rights:

A case note on *Moohan v Lord Advocate*

Paul Fradley*

I. THE ISSUES

M*oohan*¹ concerned an application for judicial review of section 2 of the Scottish Independence Referendum (Franchise) Act 2013, an Act of the Scottish Parliament ('ASP'). The 2013 ASP based the franchise for the Scottish Independence Referendum in September 2014 on that used for local elections in the UK. The law regarding local government election franchise in the UK is contained in section 3 of the Representation of the People Act 1983 which provides that 'a convicted person during the time that he is detained in a penal institution in pursuance of his sentence... is legally incapable of voting at any parliamentary or local government election'. The appellants in *Moohan* were prisoners in Scotland who were to be denied the right to vote in the Independence Referendum by the provisions of the 2013 ASP. The appellants' initial application for judicial review was refused by Lord Glennie in the Outer House of the Court of Session,² and this refusal was partly on the grounds that Article 3 Protocol 1 ('A3P1') of the European Convention on Human Rights ('ECHR') did not protect an individual's right to vote in a referendum. The Inner House of the Court of Session also rejected the appellant's case in July 2014.³

On appeal the appellants relied on a number of grounds;⁴ firstly, that the Scottish Independence Referendum (Franchise) Act 2013 was incompatible with A3P1; secondly, that it was incompatible with Article 10 of the ECHR; thirdly, that it was incompatible with the law of the European Union; fourthly, that it contravened the requirements of the International Covenant on Civil and Political Rights 1966; fifthly, that it was incompatible with 'the basic democratic principles of the common law constitution, namely the principle of universal suffrage and the concomitant fundamental right to vote';⁵ and finally, that it contravened the requirements of the rule of law. The central grounds for review, and which the Court considered at great length, were the potential incompatibility with Convention rights and the alleged common law right to vote.

The Supreme Court held by a majority of 5:2 that the ASP was valid, dismissing the countervailing common law and Convention-based arguments. Much of the discussion in the case concerned the content of ECHR jurisprudence and how this should be applied in a domestic context. On the Convention rights issue the result reached by the minority would have been preferable and the majority were confined by their interpretation of A3P1 and European Court of Human Rights ('ECtHR') jurisprudence. The case also represents a movement in the case law on section 2 of the Human Rights Act 1998 ('HRA') away from the so-called 'mirror principle', which is poignant given the concerns that have been expressed about the ECHR and the role of Strasbourg. On the common law rights point, the decision not to find a common law right in this case is interesting in the context of a recent resurgence in common law rights⁶ and the prospect of repeal of the HRA. However, the Supreme Court was arguably wrong not to consider more deeply the question of a common law right to vote even given the existence of clear legislative intent.

*Brasenose College, Oxford

¹ The Supreme Court gave its judgment in *Moohan* in December 2014. However, the arguments, deliberation, and verdict took place on the 24th July due to the time constraints of the Scottish Independence Referendum, which took place in September 2014.

² [2013] CSOH 199.

³ [2014] CSIH 56.

⁴ [2014] UKSC 67 [5] (Lord Hodge).

⁵ *ibid.*

⁶ See decisions in *Osborn* [2013] UKSC 61, *Kennedy v The Charity Commission* [2014] UKSC 20 and *A v BBC* [2014] UKSC 25.

II. CONVENTION RIGHTS

A. CONVENTION RIGHTS: THE ISSUE

The central question before the Court in *Moohan* was whether the Scottish Independence Referendum came within the provisions of A3P1. A3P1 requires signatory parties to ‘undertake to hold free elections at reasonable intervals by secret ballot, under conditions which will ensure the free expression of the opinion of the people in the choice of the legislature’. The decision in *Moohan* hung on the words: ‘the choice of the legislature’. All of their Lordships accepted that A3P1 of the ECHR prohibited a blanket ban on prisoners voting in elections, and that a vote in a UK General Election would clearly be captured by the phrase. This much is clear from the decisions of the Strasbourg Court in *Hirst v United Kingdom (No 2)*⁷ and, in the domestic context, the Supreme Court itself in *Chester*.⁸ *Moohan* therefore arose in the context of a highly-charged political issue, namely prisoners’ voting, an area in which UK law has been held incompatible with the requirements of the Convention and in which Parliament has recently rejected proposals to make UK law Convention-compliant.

B. CONVENTION RIGHTS: THE APPROACH OF THE MAJORITY

Lord Hodge, delivering the judgment of the majority of the Court, argued that A3P1 applied only to ‘periodic elections to a democratically elected legislature’ and that this was ‘the ordinary meaning of the words’.⁹ Much emphasis was placed by Lord Neuberger on the phrases ‘at regular intervals’¹⁰ and ‘elections’¹¹ which he argued precluded A3P1 from applying to non-binding referendums and restricted it only to ‘directly effective elections’.¹² Moreover, Lord Neuberger added that an entitlement to vote in the choice of the legislature did not automatically include an entitlement to vote in the choice of which legislature should govern.¹³

The majority relied heavily on the case law of the ECtHR in order to support their view that A3P1 did not extend to referendums. Lord Hodge quoted¹⁴ the decision in *X v United Kingdom*,¹⁵ in which the ECtHR held that A3P1 did not apply to the 1975 referendum on the UK’s membership of the European Community, along with the decisions in *Ž v Latvia*¹⁶ and *Niedźwiedz v Poland*,¹⁷ in which the ECtHR clearly restricted A3P1 to choices of the legislature. For Lord Hodge, the case law of Strasbourg was ‘unequivocal’ on the matter.¹⁸

⁷ (2005) 42 EHRR 849.

⁸ [2013] UKSC 63.

⁹ *ibid* [8] (Lord Hodge).

¹⁰ *ibid* [45] (Lord Neuberger).

¹¹ *ibid*.

¹² *Chester* (n 8) (Lord Neuberger)

¹³ *ibid* [46] (Lord Neuberger).

¹⁴ *ibid* [10] (Lord Hodge).

¹⁵ (1975) 3 DR 165.

¹⁶ (Application No 14755/03, 16 March 2006) (unreported).

¹⁷ (2008) 47 EHRR SE6.

¹⁸ *Moohan* (n 4) [14] (Lord Hodge).

C. CONVENTION RIGHTS: THE APPROACH OF THE MINORITY

However, Lord Kerr, dissenting, disagreed with this reasoning. For him the present case met the requirements that the Commission felt were absent in *X*. In that case the referendum was purely of a consultative character and there was no legal compulsion to hold the vote, while in *Moohan* both the Westminster and Holyrood Governments had agreed the result would be binding and there was legal compulsion to hold the vote through section 1(1) of the Scottish Independence Referendum Act 2013. While acknowledging it was ‘strictly unnecessary to go further’, he also argued there were serious deficiencies in the reasoning in *X*.¹⁹ Despite this, there was nothing in the Strasbourg case law that clearly excluded A3P1 from applying to referendums which ‘in effect determine the choice of legislature for a country’s people’.²⁰ Lord Wilson, also dissenting, agreed with this conclusion, arguing that there was no decision of the Strasbourg Court directly on the issue.²¹ Lord Wilson, in particular, drew a distinction between different forms of referendum, suggesting that there was a fundamental difference between referendums designed to curtail some powers of the legislature, and referendums intended to end the power of the legislature.²² Thus, Lord Wilson would have been able to distinguish *Moohan* from the decisions of the ECtHR on the basis that they concerned accession to the European Community rather than secession. This was an argument rejected outright by Lord Hodge, who saw ‘no material difference’ between the two.²³

Furthermore, the minority adopted a much more purposive approach to the interpretation of A3P1 than the majority. They argued that the purpose of A3P1 was to ‘ensure the free expression of the opinion of the people in the choice of the legislature’²⁴ and that in the independence referendum voters were choosing not merely legislators but the legislature itself.²⁵ Moreover, the minority emphasised the importance of the vote in question as the ‘most fundamental of votes’²⁶ and stated that to deny individuals the vote in such a case ‘would strike at the root of the values which A3P1 are designed to protect’.²⁷ Therefore it was not conclusive that the framers of A3P1 had failed to provide for secession referendums.²⁸ Both Lord Kerr and Lord Wilson bemoaned the emphasis placed by the majority on the phrase ‘at reasonable intervals’ in A3P1, arguing that it must be subordinate to the ‘primary aim of the provision’²⁹ and that if the phrase was allowed to detract from the overall objective of A3P1 then ‘the tail would be wagging the dog’.³⁰ The minority thus argued that A3P1 must cover cases like the Scottish Independence Referendum in order to give it fair effect in light of its objective.

¹⁹ *Moohan* (n 4) [73] (Lord Kerr).

²⁰ *ibid* [71] (Lord Kerr).

²¹ *ibid* [103] (Lord Wilson).

²² *ibid* [102] (Lord Wilson).

²³ *ibid* [18] (Lord Hodge).

²⁴ *ibid* [93] (Lord Wilson), quoting Article 3 of Protocol No 1.

²⁵ *ibid*.

²⁶ *ibid* [68] (Lord Kerr).

²⁷ *ibid* [75] (Lord Kerr).

²⁸ *ibid* [68] (Lord Kerr).

²⁹ *ibid* [69] (Lord Kerr).

³⁰ *ibid* [98] (Lord Wilson).

D. CONVENTION RIGHTS: EVALUATING THE APPROACHES

The majority of the Supreme Court in *Moohan* took a line of interpretation which is quite literal in nature. They relied on the wording of A3P1 and sought to give a narrow meaning to the provision as a result. The minority, by contrast, looked to the aim or purpose of A3P1 and sought to give it a wider meaning in doing so. It has to be appreciated that narrower interpretations are generally more appropriate in the context of international obligations. However, in the context of protection of fundamental human rights, it is of some concern that the court sought to give such a narrow interpretation to human rights in this case. Given the constitutional importance of the vote, it is questionable whether such a restrictive understanding of A3P1 was really appropriate. As the minority emphasised, giving fair effect to the purpose of A3P1 requires its extension to the independence referendum; as Lord Kerr argued, the ‘primary aim of the provision [is] to ensure that citizens should have a full participative role in the selection of those who will govern them’.³¹ The decision in *Moohan* would, it seems, fail to achieve this.

The purpose of A3P1 is presented as ensuring that citizens have the ability to participate in the governance of their community. Lord Wilson argued that ‘had [the drafters of A3P1] had it in mind, they would have expressly provided that a right to vote in [the Scottish Independence Referendum] fell within [A3P1’s] ambit’.³² In short he argued that, if one asked the drafters of the provision whether they believed a vote of the magnitude of the Independence Referendum should be included, alongside for instance a UK General Election (which is included in the right in A3P1),³³ they would most likely have said ‘yes, of course’.

To criticise the majority on this point is, arguably, going too far. Firstly, it is far from clear that popular referendums on international treaty change or other major constitutional changes should be the primary method of decision-making and that the legislative process should be subordinate. Secondly, courts are tasked, in the context of interpretation of treaties or statutes, with applying the words so as to give best effect to the intentions of the drafters. A court may recognise, as Lord Wilson did in *Moohan*,³⁴ that the drafters would probably have written something different with the benefit of hindsight. But, as the maxim goes, hindsight is a wonderful thing. The phrase the ‘choice of the legislature’ in A3P1 does, on its natural interpretation seem to exclude the ambit of AP31 from a referendum, such as the Scottish Independence Referendum.

Moreover, the restrictive interpretation is based on the interpretation given to the provisions by the ECtHR itself. Indeed, Lord Hodge notes that the limited ‘object and purpose of A3P1’ is ‘confirmed by the consistent case law of the European Commission on Human Rights and the Strasbourg court’.³⁵ Lord Kerr accepts that if the interpretation of A3P1 by Strasbourg is not doubtful (as he believed it to be) then consideration of other matters is inappropriate.³⁶ If there is a problem in the narrowness of the interpretation given to the right in A3P1 by the majority, the fault is more properly with the Strasbourg Court, or in how the courts interpret their obligations under section 2 of the HRA, and not with the Supreme Court.

³¹ *Moohan* (n 4) [69] (Lord Kerr).

³² *ibid* [98] (Lord Wilson).

³³ *Hirst* (n 7) and *Chester* (n 8).

³⁴ *Moohan* (n 4) [69] (Lord Kerr).

³⁵ *ibid* [9] (Lord Hodge).

³⁶ *ibid* [66] (Lord Kerr).

E. SECTION 2 AND THE MIRROR PRINCIPLE

Section 2 of the HRA requires the UK courts to ‘take into account’ the jurisprudence of Strasbourg. Quite what ‘take into account’ is supposed to mean has been the subject of much controversy, both in case law and in academic discussion. In *Ullah*³⁷ Lord Bingham articulated the so-called ‘mirror principle’; namely, that UK courts should mirror Strasbourg’s jurisprudence ‘no more but certainly no less’.³⁸ Given the political debate surrounding repeal of the HRA and the role of the ECtHR, whether the present circumstances warrant a revised approach to interpreting the requirements of section 2 is a matter of considerable interest. What appears to be a simple rule has in fact turned out to be anything but. In *AF (No 3)*³⁹ Lord Rodger famously stated that ‘Strasbourg has spoken, the case is closed’. In *Horncastle*,⁴⁰ the Supreme Court was willing to depart from a decision of the ECtHR at first instance in *Al-Khawaja*,⁴¹ a decision later confirmed to be correct by the Grand Chamber.⁴²

F. SECTION 2 IN MOOHAN

Lord Wilson in *Moohan* provides a useful summary of recent case law which has departed from the principle,⁴³ and argues that ‘protracted consideration over the last six years has led this court to substantially modify the *Ullah* principle’.⁴⁴ Lord Hodge makes it clear that ‘there is room for disagreement and dialogue between domestic courts and the Strasbourg court’,⁴⁵ but goes on to quote Lord Neuberger in *Pinnock*⁴⁶ to the effect that the courts should follow ‘a clear and constant line of decisions’ which is not inconsistent with some fundamental aspect of our law and does not overlook some important argument.⁴⁷ Lord Hodge concludes that there is no reason to go further in this case and that doing so would not ‘reach a conclusion which flows naturally from Strasbourg’s existing case law’.⁴⁸ This is similar to the comments of Lord Sumption in *Chester* who argued that unless Strasbourg has ‘misunderstood or overlooked some significant feature of English law or practice’⁴⁹ then it ‘would be neither wise nor legally defensible’⁵⁰ to define Convention rights differently to Strasbourg.

Lord Wilson’s comments on section 2 are animated by his conclusion, as outlined earlier, that there is no Strasbourg case law directly relevant to the matter at hand in *Moohan*.⁵¹ He argues that ‘where there is no directly relevant decision of the ECtHR with which it would be possible (even if appropriate) to keep pace, we can and must do more’.⁵² These comments echo those of Lord Kerr

³⁷ [2004] UKHL 26.

³⁸ *ibid* 350 (Lord Bingham).

³⁹ [2009] UKHL 28, [98] (Lord Rodger).

⁴⁰ [2009] UKSC 14.

⁴¹ (2009) 49 E.H.R.R. 1.

⁴² (2012) 54 E.H.R.R. 23.

⁴³ *Moohan* (n 4) [104] (Lord Wilson).

⁴⁴ *ibid* [105] (Lord Wilson).

⁴⁵ *ibid* [13] (Lord Hodge).

⁴⁶ [2010] UKSC 45, [2011] 2 AC 104.

⁴⁷ *Moohan* (n 4) [13] (Lord Hodge).

⁴⁸ *ibid*.

⁴⁹ *ibid* [121].

⁵⁰ *ibid* [138].

⁵¹ *ibid* [102] (Lord Wilson).

⁵² *ibid* [105] (Lord Wilson).

(dissenting) in *Ambrose v Harris (Procurator Fiscal)*,⁵³ who argued that '[i]f the much vaunted dialogue between national courts and Strasbourg is to mean anything' then courts must define Convention rights in the absence of Strasbourg decisions.⁵⁴ It therefore appears that increasingly 'domestic courts should regard the ECtHR case law as a floor, but not a ceiling'.⁵⁵

Moohan provides a good example of a case where judicial dialogue, between the Supreme Court and Strasbourg, would have been entirely appropriate. Judicial dialogue can entail both defining Convention rights in the absence of a definition from Strasbourg and refining a Convention right as defined by Strasbourg. The decision in *Moohan* meant that prisoners were unable to vote in the Scottish Independence Referendum, and thus were deprived of an important freedom available to those not incarcerated. A different decision in *Moohan* would certainly have given greater protection to the prisoners. *Moohan* appears to accept that the principle laid down in *Ullah* is no longer, if it ever was, an accurate reflection of the requirements of section 2. It feeds into a wider debate about what section 2 means. It is highly arguable that *Moohan* was a perfect opportunity to go beyond Strasbourg.

III. COMMON LAW RIGHTS

A. COMMON LAW RIGHTS: THE RECENT CASE LAW

The existence of fundamental common law rights is not a new idea. In *Ex parte Simms*⁵⁶ Lord Hoffmann recognised the existence of such common law rights, noting that '[f]undamental rights cannot be overridden by general or ambiguous words'.⁵⁷ Moreover, in *Daly*,⁵⁸ Lord Cooke famously said that 'some rights are inherent and fundamental to democratic civilised society. Conventions, constitutions, bills of rights and the like respond by recognising rather than creating them'.⁵⁹ In *Moohan* their Lordships were asked to consider the possible existence of a common law right to vote, an argument which they rejected. The rejection of this is important given that, following *AXA General Insurance Ltd v Lord Advocate*,⁶⁰ the Scottish Parliament is subject to judicial review at common law on the ground of 'fundamental rights or the rule of law'.⁶¹ The existence of such a common law right could, therefore, have been a ground for successful judicial review of the 2013 ASP.

The decision in *Moohan* follows a stream of cases in which the courts have recently reasserted the existence of common law rights in the context of openness and transparency in justice. For instance, in *Kennedy v The Charity Commission*,⁶² Lord Toulson bemoaned the 'baleful and unnecessary tendency to overlook the common law' in preference to convention rights,⁶³ and Lord Mance argued that 'the natural starting point in any dispute is to start with domestic law [sic]'.⁶⁴ In a speech of

⁵³ [2011] 1 WLR 2435.

⁵⁴ *ibid* [130] (Lord Kerr).

⁵⁵ Richard Clayton, 'Smoke and mirrors: the Human Rights Act and the impact of Strasbourg case law' [2012] PL 639.

⁵⁶ [2000] 2 AC 115 (HL).

⁵⁷ *ibid* 131 (Lord Hoffmann).

⁵⁸ [2001] UKHL 26.

⁵⁹ *ibid* [30] (Lord Cooke).

⁶⁰ [2011] UKSC 46.

⁶¹ *ibid* [149] (Lord Reed).

⁶² [2014] UKSC 20.

⁶³ *ibid* [133] (Lord Toulson).

⁶⁴ *Kennedy* (n 62) [46] (Lord Mance).

2014,⁶⁵ Lord Neuberger argued that the judges had seen the HRA as a ‘new toy’ and that the courts were increasingly taking the ‘old toy’ of common law constitutional rights back out of its box. Baroness Hale, in 2014, captured the recent developments when she said that ‘there is emerging a renewed emphasis on the common law and distinctively UK constitutional principles as a source of legal inspiration’.⁶⁶ In the face of possible repeal of the HRA, this development has taken on added impetus.

B. A COMMON LAW RIGHT TO VOTE? MOOHAN

Lord Hodge and Baroness Hale in *Moohan* both refused to recognise the existence of a common law right to vote. Lord Hodge had no problem recognising the right to vote as a ‘basic or constitutional right’,⁶⁷ and Baroness Hale suggested ‘it would be wonderful if the common law had recognised a right to universal suffrage’.⁶⁸ However, they both refused to hold that the common law had recognised such a right to vote, from which ‘express language or necessary implication to the contrary’⁶⁹ were required for any kind of derogation. The franchise is regulated in a statutory scheme and has been extended in statute, and as such it ‘would not be appropriate to seek to develop the common law in order to supplement or override the statutory rules which determine our democratic franchise’.⁷⁰

The opinions of Lord Hodge and Baroness Hale seem to suggest that the existence of the statutory scheme precludes the existence of a common law right in this area. Their opinions laid great emphasis on the historical development of democracy in the UK constitutional structure and the lack of judicial role in this. Baroness Hale argued that it would be ‘absurd’ to say that women had a common law right to vote, since full franchise for women was only granted in 1928.⁷¹ Lord Hodge argued that ‘it has [...] been our constitutional history that for centuries the right to vote has been derived from statute’.⁷² Emphasis should be placed on the express provision that prisoners should be deprived of the ability to vote contained in section 3(1) of the Representation of the Peoples Act 1983.

C. EVALUATING THE COMMON LAW RIGHT TO VOTE

However, the existence of the 1983 Act should not preclude the existence of the common law right as it was not this statutory scheme but the Scottish Independence Referendum (Franchise) Act 2013 that was at issue. In this context the 1983 Act is relevant only in so much as the 2013 ASP is based upon it. In this context there was no Act of the Westminster Parliament which stands normatively above the common law. The views expressed in the 1983 Act certainly inform the debate, but they are not conclusive. The Courts may be cautious given that the Westminster Parliament would seem to have dealt with the issue conclusively. However, where the Act of Parliament is not directly relevant the Courts should use it only to inform their broader discussion. There was in *Moohan* a lack of emphasis on whether the common law may have developed to recognise a right to vote, separate from the

⁶⁵ Lord Neuberger, ‘The role of judges in human rights jurisprudence: a comparison of the Australian and UK experience’ (Supreme Court of Victoria conference, Melbourne, August 2014).

⁶⁶ Lady Hale, ‘UK Constitutionalism on the March?’ (Constitutional and Administrative Law Bar Association conference, July 2014).

⁶⁷ *Moohan* (n 4) [33] (Lord Hodge).

⁶⁸ *ibid* [56] (Baroness Hale).

⁶⁹ *Ex p Simms* (n 56) 131 (Lord Hoffmann).

⁷⁰ *Moohan* (n 4) [34] (Lord Hodge).

⁷¹ *ibid* [56] (Baroness Hale).

⁷² *ibid* [34] (Lord Hodge).

statutory scheme created by Westminster, which is not conclusive in this case, and whether the judiciary should play a greater role in the development of democracy in the UK going forward. In *A v BBC*,⁷³ Lord Reed spoke of ‘the capacity of the common law to develop’.⁷⁴ Some limited support in *Moohan* for this more dynamic view of common law rights is found in the dissenting judgment of Lord Kerr, who suggests that ‘the common law can certainly evolve alongside statutory developments without necessarily being entirely eclipsed by the latter’.⁷⁵ However, he expresses no definitive judgment on the matter, concluding that the existence of the right is ‘at least arguable’.⁷⁶

Moohan also demonstrates the limited reach of common law rights. On the facts of *Moohan* both the common law and the ECHR fail to recognise a right of prisoners to vote in the context of a referendum. However, the unwillingness to recognise a common law right to vote suggests that the common law would not recognise the right of prisoners to vote in UK General Elections, when decisions of both the ECtHR⁷⁷ and the Supreme Court⁷⁸ have held that such a right is recognised by the ECHR and the HRA. The Supreme Court in other recent cases had been willing to recognise and apply common law rights in contexts where Convention rights were also in play.⁷⁹ *Moohan* represents a change of direction in judicial thinking away from a willingness to recognise common law rights and apply them widely towards a more hesitant approach. The case reminds us that the normative reach of common law rights is still limited compared to the ECHR and is likely to remain so.

D. COMMON LAW RIGHTS: ‘STRIKE DOWN’

The courts have recently mooted the existence of a strike-down power at common law,⁸⁰ and in fact Lord Hodge tentatively suggests as much *obiter* in *Moohan* when he argues that if Parliament ‘abusively’ sought to restrict the franchise then ‘the common law, informed by the principle of democracy and the rule of law and international norms, would be able to declare such legislation unlawful’.⁸¹ Any judicial strike-down would be an unusual and extreme event, one which could shake the constitutional and political structures of the UK to their core. By comparison, the courts protecting rights using section 4 declarations has become commonplace.⁸² For example, the Supreme Court recently issued a declaration of incompatibility in *Reilly (No 2)*⁸³ in relation to the Job Seekers (Back to Work Schemes) Act 2013 which had attempted to retrospectively validate a government scheme requiring benefit claimants to work for private companies in order to receive payments which had been introduced by the Jobseeker’s Allowance (Employment, Skills and Enterprise Scheme) Regulations 2011 which the Supreme Court in *Reilly (No 1)*⁸⁴ had declared ultra vires.

⁷³ [2014] UKSC 25.

⁷⁴ *ibid* [57] (Lord Reed).

⁷⁵ *Moohan* (n 4) [86] (Lord Kerr).

⁷⁶ *ibid* [87] (Lord Kerr).

⁷⁷ *Hirst* (n 7).

⁷⁸ *Chester* (n 8).

⁷⁹ *A v BBC* (n 73), *Osborn* [2013] UKSC 61, and *Kennedy v The Charity Commission* (n 62).

⁸⁰ See dicta in *Jackson v Attorney General* [2005] UKHL 56 [102] (Lord Steyn), [104]-[107] (Lord Hope), and [159] (Baroness Hale); *AXA General Insurance Ltd v Lord Advocate* (n 60) [51] (Lord Hope).

⁸¹ *Moohan* (n 4) [35] (Lord Hodge).

⁸² For example *Anderson* [2002] UKHL 46, *In re S (Minors) (Care Order: Implementation of Care Plan)* [2002] UKHL 10, *Bellinger v Bellinger* [2003] UKHL 21 and *A v Secretary of State for the Home Department* [2004] UKHL 56.

⁸³ [2014] EWHC 2182 (Admin).

⁸⁴ [2013] UKSC 68.

The courts would likely show extreme caution in invalidating any legislation given the issues such an event would raise for our democratic system of government and the uncertainty that judicial invalidation would create. Lord Phillips has argued that before any strike-down the courts ‘would have given an interpretation to the legislation that it, faced with it, couldn’t bear, but would have chucked the gauntlet back to Parliament’,⁸⁵ citing the decision of the Privy Council in *Anisminic*⁸⁶ as an example of such an interpretation. The extent to which a strike-down power would be of any utility in cases such as *Moohan* is therefore questionable.

The prospect of a judicial strike-down on the grounds of incompatibility with supposed common law rights would amount to a huge transfer of power from the political to the judicial sphere. Such a transfer raises concerns about the institutional competence of different branches of government to decide how rights should be implemented in practice. For example, is the composition of the electorate a decision best left to the judicial or legislative branch? This is particularly concerning given, by contrast to the political sphere, the lack of democratic control or democratic legitimacy that the legal sphere has. Lord Hailsham famously referred to the UK as an ‘elective dictatorship’,⁸⁷ might judicial invalidation on common law grounds move the UK towards an ‘unelected dictatorship’?

However, there are serious concerns with this argument. Firstly, the lack of democratic legitimacy for judges could be solved by electing judges, as many states do. It is far from clear though that this leads to better outcomes and in fact it risks undermining the independence of judges – it is doubtful that an elected judiciary would be any better. Secondly, judges need to be legitimate but this does not necessarily mean that judges need to be elected. There is arguably legitimacy in expertise, in experience; democracy and legitimacy should not be conflated – it is therefore also doubtful that an elected judiciary would be any more legitimate when legitimacy is understood more broadly. Thirdly, we must look to a more normatively rich view of democracy: not simply one predicated on a view that democracy necessarily means majoritarian structures. If democracy is to serve as a primary value within our constitutional system then it must contain within its definition an element of protecting minority rights. What is required therefore is a constitutional structure which separates power and entrusts it to the branch with the institutional competence to best use them. The legislative and executive branches are often the best place for decision making, but the judicial branch has a role to protect fundamental rights particularly those of minorities. Lord Hodge’s comments coupled with other dicta⁸⁸ demonstrate a willingness on the part of the judiciary to consider the possibility of playing a more active role within their institutional competence.

IV. CONCLUSION

On the Convention rights issue the approach of the minority is to be preferred, given the importance of the vote allowing a blanket-ban on prisoner voting in a referendum but not a General Election does appear an artificial line. However, the majority were working within the confines of how they interpreted A3P1 and the jurisprudence of the ECtHR. The comments on section 2 of the HRA are indicative of recent changes to what that section is seen to require of domestic courts; however *Moohan* can be seen as a missed opportunity to go further than Strasbourg. The summary provided by Lord

⁸⁵ Political and Constitutional Reform Committee, *Constitutional role of the judiciary if there were a codified constitution* (HC 2013-14, 802-I), quoting Lord Phillips at para 41.

⁸⁶ [1969] 2 AC 147 (HL).

⁸⁷ Lord Hailsham, ‘Elective Dictatorship’, The Richard Dimbleby Lecture (BBC, 1976).

⁸⁸ See (n 80).

Wilson⁸⁹ in particular encourages us to reflect more broadly on the relationship between the ECtHR and the domestic courts. Both the comments on section 2 and the consideration of common law rights are particularly topical given the current political climate,⁹⁰ and indeed, much of the debate around repeal of the HRA centres on the question of the role afforded to the jurisprudence of the Strasbourg Court.

Moohan represents a marked change of approach on the part of the Supreme Court towards common law rights. Recent developments in common law rights have been suggestive of a highly active approach. Here, however, the matter is dealt with much more cautiously. It is wrong to dismiss the argument that a common law right could exist in an area touched by statute if an Act of the Westminster Parliament does not expressly cover the case. Legislative intent should be evidence that guides the possible recognition of a common law right.

⁸⁹ See (n 43).

⁹⁰ Namely, the election of a majority Conservative government with a manifesto commitment to repeal of the Human Rights Act and its replacement with a 'British Bill of Rights'.

A Failure to Protect: The Shortcomings of the United Kingdom's Judicial Response to the Privatisation of the Public Realm

Owen Nanlohy¹

I. INTRODUCTION

The purpose of this article is to examine the current approach taken by the courts in the United Kingdom for determining whether a private entity is exercising functions of a public nature when under contract with a public authority. If, as is argued, private entities are performing functions of a public nature, the fundamental question is whether the near-ubiquitous practice of outsourcing of public services should be susceptible to public law protection. Answering this question is vitally important where a right under the Convention for the Protection of Human Rights and Fundamental Freedoms, scheduled in the Human Rights Act 1998 (“HRA”), is claimed to have been breached. This article falls narrowly within a much broader debate about the public administration of the United Kingdom, and the consequences of increasing levels of contractualisation for the scope of judicial control over the activities of the executive.

First, the article begins by charting the growing development of outsourcing in the provision of public services. Second, the article considers the approach that the courts have taken when faced with questions relating to the privatisation of public services. The argument put forward is that the courts have inappropriately failed to extend the remedial protection of the public law when private sector parties under contract perform services of a public nature, particularly where a Convention right is engaged. It is submitted that the public law jurisdiction, and the particular remedies available as a consequence, should be expanded to encompass the increasing provision of services of a public nature performed by the private sector on a contractual basis with the State. Third, the article considers the questions resulting from the decision of the House of Lords in *YL v Birmingham*,² a case which effectively reduced the scope of public law supervision and protection in this field. The article considers why the courts have been slow to bring the increasingly complex web of private providers of public services within the judicial protection afforded to individuals by the public law of the United Kingdom and what the test may look like in instances where the courts are called upon to control public power exercised by private providers. Fourth, this article notes the contradiction between the judicial reticence to read into section 6 of the HRA a wide scope of supervisory jurisdiction, when the courts were far more ready to accept a flexible test in earlier jurisprudence for applications made under section 31 of the Senior Courts Act. Finally, the article makes the argument that private sector providers of public services should, where the service is evidently of a public nature, be captured by public law, specifically where rights scheduled under the HRA are exercised.

¹ University College, Oxford

² [2007] UKHL 27, [2008] 1 A.C. 95.

II. A BRIEF HISTORY OF PRIVATISATION IN THE UNITED KINGDOM

Between 1945 and 1979 the overwhelming majority of public services were supplied by the administrative organs of the state: the consequence of a post-war transfer of power away from a private sector that could not meet the needs of the community, and toward the state provision or subsidisation of a diverse range of services that were desperately required by a country rebuilding itself after war. However, the last 35 years has seen a reverse of that historical process, with the contractual transfer of a plethora of services from public bodies to private and non-profit organisations. Since 1979 governments formed from Conservative, Labour, and contemporarily the Coalition – comprised of the Conservative and Liberal Democratic parties – have overseen an unprecedented transfer in the provision of public services from the State to the private sector.³ This has been nothing short of a revolution in the organisation of the modern British state. The privatisation phenomenon signalled a profound departure from the dominant organising paradigm of the post-war liberal democratic state: the provision of public services by the state using a centralised command and control system of government, framed by a commitment to equality and quality of opportunity for all citizens of the United Kingdom.

III. THE BORDER COUNTRY: THE CHANGING LANDSCAPE OF THE PUBLIC-PRIVATE DIVIDE

This article is not concerned with the broader socio-political trend of privatisation. Rather, it examines the consequences for public law, and the judicial protection of individual rights, where the original public service or function has been contracted out to the private sector. The contemporary legal framework used to distinguish between the delivery of a public service, authorised by a statutory or prerogative power, and the fulfilment by a private person, non-profit organisation, or commercial entity of an outsourced by contract public service, is to define the former as a public function and the latter as the satisfaction of a private contractual agreement. The assumption of this dichotomy is that when outsourced the subject-matter of the contract changes from one which would have been captured by public law remedies, if performed by a public body, to subject-matter susceptible only to the ordinary remedies of private law.

This definitional demarcation has had a fundamental effect on how courts draw the boundaries available for the protection of individual interests and rights. In the case of public functions provided by public authorities, unlawful decisions are susceptible to judicial review and, if appropriate, injunctive relief rather than damages.⁴ The same is not true of the private provision of those same services: public law controls stop at the contract, on the rather unsatisfactory assumption that private remedies will be available, in contract or tort, to those who have a legitimate cause of action. This systemic shift in thinking about the provision of public services, and the willingness to exclude private service providers of government contracts from the scope of public law control, has the potential to fundamentally undermine the availability of appropriate legal protections for individuals who have been unfairly or unlawfully treated by private providers, whether or not they had any agency in the decision to outsource those public functions which they rely upon.

³ For further commentary on the development of privatization in the United Kingdom see Mark Freedland, 'Government by Contract and Private Law' [1994] PL 86.

⁴ See for example *Anisimic v Foreign Compensation Commission* [1969] 2 AC 147 (HL).

The courts, in recent case law, have failed to protect those rights and interests for two reasons. First, the dominance of private law in the English legal psyche has led to a resolute rear-guard action seeking to shield private providers of public services when they should actually be subject to the same judicial scrutiny exercised against public bodies and authorities. Second, the courts, perhaps because of the fundamental inclination to maintain clear borders between public and private law, have failed to adjust to the reality of a system of public service provision by outsourced private partnerships. It is submitted that a significant adjustment needs to occur in judicial thinking on this subject and that the tests thus far propounded are unsatisfactory and fail to protect individual rights caught in this changing realm of private provision of public services.

IV. THE CURRENT SCOPE OF REVIEW

Though there are many applications requesting the intervention of the court which do not engage Convention rights, brought into force in the law of the United Kingdom under the HRA, it is in respect of applications which claim the engagement of Convention rights that the procedural flaws are most evident and ripe for reform. This is because the approach adopted by the courts in this category of cases reflects the threshold test that the courts use to determine whether or not the respondent party is definable as a public authority, and whether it is therefore susceptible to the courts' public law jurisdiction. The current procedure that the courts adopt when determining, for the purposes of section 6(1) of the HRA, 'whether a public authority has acted in a way which is incompatible with a Convention right', is to first determine whether the person, authority or distinct legal entity can be defined as a public authority under the Act.⁵ For these purposes the court is concerned with section 6(3)(b), which defines a public authority as 'any person certain of whose functions are functions of a public nature'.⁶ This, arguably, authorizes the court to set a wide scope of interpretation, subject to the specific merits of the case and the power or duty that is being considered.

Parliament has certainly not confined the scope of interpretation to public authorities with powers derived from statute alone. However, as will become plainer, while there is an arguably wide scope for interpretation out of section 6(3)(b), the judicial approach would instead be more properly characterized as an exceptionally narrow interpretation when dealing with matters where the respondent is a private party providing a public service on contract. For the purposes of this article a 'public service' is defined as a service that would have generated an obligation on the public authority to provide it to those individuals who possess a right to access the service subsequently provided under private contract. This would automatically exclude a significant number of contracts made between public authorities and private parties to provide goods or services which may have an indirect effect on the provision of services of a public nature but are themselves fundamentally of a private nature.

Significantly, the current judicial approach for determining what is and is not a public authority, or what may or may not be defined as a function of a public nature, reveals an unwillingness to engage in a larger debate over the consequences for English public law of the increasing corporatisation and contractualisation of public services. Moreover, the desire to stay out of this discourse has revealed the ideological divisions playing out beneath the surface of decisions made in this arena by the appellate courts in the United Kingdom. Decisions are often subjectively decided and betray an inability to grasp the nettle of an increasingly privatised administrative state, and extend public law principles into the realm otherwise governed by private law in order to protect both the Convention rights of individuals and the public interest more broadly constituted. This unwillingness is borne out in recent authority of the Supreme Court.

⁵ Human Rights Act 1998, s 6(1).

⁶ Human Rights Act 1998, s 6(3)(b).

V. THE LEADING CASE LAW

The post-war jurisprudence of English public law has seen a general trend toward further judicial protection of individual rights and interests. Considered against the background of this trend, the unwillingness to deal with the consequences for those same rights and interests of greater levels of privatisation and the contracting out of public services appears to be something of an enigma. The decision in *YL v Birmingham City Council* follows a line of authority which set out the court's approach to the protection of fundamental rights where public duties have been discharged through the contracting out of services to the private sector.⁷ In *YL*, the case concerned an 84 year old woman suffering from Alzheimer's disease who was provided full time care and a place in a home by the first respondent local authority. The first respondent had a duty under section 21 of the National Assistance Act 1948 to make arrangements for the provision of residential accommodation for eligible persons. The first respondent contracted with the second respondent, a provider of health and social care services, to provide a place for the claimant in one of its care homes, which accommodated both privately-funded residents and those whose fees were paid by the council in full or in part. The second respondent sought to terminate the contract and remove the claimant from the home, because publicly funded places were less lucrative than the rate chargeable to privately funded places. The claimant, through the Official Solicitor, commenced proceedings in the family division of the High Court seeking, *inter alia*, declarations that it would not be in the claimant's best interests to be moved out of the home and that the company, in providing accommodation and care for the claimant, was exercising public functions within section 6(3)(b) of the HRA; further, that should the company be successful in evicting the claimant from the home, it would breach her rights under Articles 2, 3 and 8 of the European Convention on Human Rights, contrary to section 6(1) of the HRA. The judge hearing the matter held as a preliminary issue that the company was not exercising a public function and therefore did not fall within the meaning of the Act. The Court of Appeal upheld this finding. The case proceeded on appeal to the House of Lords.⁸ It was held that the second respondent was not performing functions of a public nature within the meaning of section 6(3)(b). It is submitted that this decision was incorrect and drew a far too restrictive scope for what constitutes a function of a public nature.

The core of the argument for the claimant was that the second respondent had been performing functions of a public nature within the scope of the Act. What exactly constituted "functions of a public nature" caused some considerable difficulty to the Court. Lord Neuberger expressed this problem by explaining the approach that would have to be adopted in the circumstances and how the approach was inherently susceptible to the subjective interpretations of the individual judges:

'The centrally relevant words, 'functions of a public nature', are so imprecise in their meaning that one searches for a policy as an aid to interpretation. The identification of the policy is almost inevitably governed, at least to some extent, by one's notions of what the policy should be, and the policy so identified is then used to justify one's conclusion. Further, given that the question of whether section 6(3)(b) applies may often turn on a combination of factors, the relative weight to be accorded to each factor in a particular case is inevitably a somewhat subjective decision.'⁹

⁷ *Poplar Housing and Regeneration Community Association Ltd v Donoghue* [2002] QB 48 (CA); *R (on the application of Heather) v Leonard Cheshire Foundation* [2002] EWCA Civ 366, [2002] 2 All ER 936.

⁸ Subsequently, the decision in *YL v Birmingham City Council* [2007] UKHL 27, [2008] 1 A.C. 95 has been reversed by section 1 of the Health and Social Care Act 2008, though it still maintains precedential value when considering the scope of section 6(3)(b) of the HRA.

⁹ *YL* (n 2) [128].

The remarkably honest dicta from Lord Neuberger exposes the inherently subjective approach taken when determining whether a body is to be considered a public authority for the purposes of proceedings which engage a Convention right. Ultimately, rather than being guided by a clear and transparent test, adjudication of what constitutes a function of a public nature, within the context of the Human Rights Act, requires the individual judge to consider what, in the circumstances, they believe ought to be the policy that the court adopts. Lord Neuberger is correct to argue that ‘functions of a public nature’ is an insufficiently precise phrase, and that it is therefore incredibly difficult to formulate an objective test when considering whether a respondent party is carrying out such functions. However, it is hard to accommodate such a view in the case of *YL* when an objective answer may be found on the facts. It is submitted that the court, on the facts of *YL*, could have reached the opposite conclusion by asking itself two basic questions. First, whether the function that was being carried out was being carried out in fulfilment of a statutory obligation or duty whether by a public body which is itself a creature of statute, or by a private party who has contracted with a public body to undertake certain functions which the public body is duty bound to deliver. This is inherently a question of fact. In *YL*, the first respondent had provisionally discharged its statutory duty to the claimant by providing them with appropriate accommodation. The legal form of the performance of that duty was through contract with a private provider. The function did not change, but the provision of the public service was now being carried out by the second respondent, a private entity undertaking the service for profit.

Second, the court should have asked itself whether the contracting out of that public function alters the function or service provided in such a manner as to effectively deprive it of its public nature or whether the service is unchanged and the public function flows through the private contract itself. The second limb of this test does introduce a degree of subjectivity. However, it is submitted that in the majority of cases engaging Convention rights the service cannot change significantly from that which would have been delivered by a public body because the subject-matter of the contract has to discharge the previously public obligation. Instead, when faced with the second limb of the test, the court asked itself a question which entailed an either-or response: specially, whether the public nature of the function flowed through the contract with the second respondent, or, alternatively, whether the services offered by the private contractor were in essence ‘private acts or functions’ sitting outside the scope of the Act and therefore not susceptible to public law remedies.¹⁰

It is submitted that this distinction is illogical and that the functions of a public nature must flow through outsourced private contracts. If the provision of the service is the same as would have been provided by the public authority which had a statutory duty to provide it, and if the profit made by the private contractor comes directly from a contract to provide those public services, there appears to be little reason to find that the functions have evolved sufficiently so as to be defined as functions of a private nature. It is almost impossible to think that a contract, which sufficiently discharges a public obligation to a corresponding individual right to a service, could be altered to effectively exclude the original public nature of the subject matter it was intended to discharge. Moreover, such an approach implicitly departs from the broader formulation that Parliament arguably intended, and the flexible test for a public function developed by earlier case law under section 31 of the Senior Courts Act 1981 (which will be dealt with more fully further on). This is not however the direction the courts have taken. The current judicial position suggests that the private provision of services either for profit, or by non-profit organizations, should not carry with it any susceptibility to public law remedy.

¹⁰ *YL* (n 2) [129].

VI. THE PRE-YL SECTION 6 CASE LAW

Prior to *YL*, the courts were already travelling down the restrictive road of excluding public law remedies for the provision of services under contract. In *Aston Cantlow PCC v. Wallbank*,¹¹ Lord Nicholls' interpretation of the operation of section 6(1) in light of section 6(3)(b) has been adopted as the flexible framework through which to view cases engaging Convention rights where one or more of the respondent parties may challenge the classification of their legal definition as a 'public authority'.¹² *Aston Cantlow* developed a dual categorisation of a public authority: first, Lord Nicholls argued that there were 'core' public authorities 'whose nature is governmental in a broad sense', such as local authorities, the armed forces, or police. He argued that 'behind the instinctive classification of these organisations as bodies whose nature is governmental lie factors such as the possession of special powers, democratic accountability, public funding in whole or in part, an obligation to act only in the public interest, and a statutory constitution...'.¹³ The second category of public authorities formulated in *Aston Cantlow* comprises 'hybrid' or 'functional' public authorities, only some of whose functions are of a public nature. Under section 6(5) of the HRA, an entity which would otherwise be a hybrid public authority is nonetheless not to be treated as such in relation to an act 'the nature of [which] is private', an exception that does not apply to core public authorities. This formulation of the test had also previously been considered in *Poplar Housing and Regeneration Community Association Ltd v. Donoghue*,¹⁴ where Lord Woolf CJ said:

The fact that a body performs an activity which otherwise a public body would be under a duty to perform cannot mean that such a performance is necessarily a public function. A public body in order to perform its public duties can use the services of a private body. Section 6 should not be applied so that if a private body provides such services, the nature of the functions are inevitably public.... Section 6(3) means that hybrid bodies, who have functions of a public and private nature are public authorities, but *not* in relation to acts which are of a private nature....¹⁵

VII. THE APPLICATION OF THE TEST IN YL

This was the reasoning adopted by Lord Mance and Lord Neuberger in *YL*, when considering the application of the Act to hybrid public authorities.¹⁶ The question then seems to turn on a bipartite test of categorization: first, the court engages in a discovery of the type of public authority the respondent party is; second, the court applies a subjective test of what functions or acts that entity is undertaking to decide whether the acts in question are of a public nature and whether they are susceptible to the jurisdiction of the court in judicial review proceedings. As Lord Mance argues in *YL*, the inquiry is inherently context-specific. It is submitted that this is undoubtedly the case. Though it may not be possible to develop a universal objective test of what may or not be a function of a public nature, it does not follow that merely because a service is provided under contract it loses its public nature. Indeed, the question of whether or not the function of a public nature flows through the contract is equally context-specific.

¹¹ [2003] UKHL 37, [2004] 1 AC 546.

¹² *ibid* [7]-[9].

¹³ *ibid* [8].

¹⁴ [2002] QB 48 (CA).

¹⁵ *ibid* [58] (Lord Woolf).

¹⁶ *YL* (n 2) [105] (Lord Mance) and [110] (Lord Neuberger).

The second stage of discovery considered by Lord Nicholls in *Aston Cantlow* seems to be more in line with Lord Neuberger's discussion of how the judge's task has inherently subjective contours in reaching a conclusion. In *Aston Cantlow*, Lord Nicholls suggested that '[c]learly there is no single test of universal application'.¹⁷ Indeed, he noted that it is necessary when looking to determine 'whether a function is public for [the] purpose' to weigh the relevant factors in order to decide whether the matter falls within the public law jurisdiction of the court, and whether it is therefore capable of injunctive or remedial protection.¹⁸ In his search for what might be taken into account Lord Nicholls fell upon a list of potential, though not exhaustive, factors that might guide the judgment of the court:

[F]actors to be taken into account include the extent to which in carrying out the relevant function the body is publicly funded, or is exercising statutory powers, or is taking the place of central government or local authorities, or is providing a public service.¹⁹

On the face of it, these factors seem to be somewhat determinative; however, when given further consideration the task of the judge becomes less clear.

VIII. THE DISSENT, IDEOLOGICAL DIFFERENCES, AND MISTAKE IN YL

In *YL*, taking the test laid down in *Aston Cantlow* as a starting point, the court had to decide whether the second respondent was susceptible to the jurisdiction of the court as a hybrid authority which was providing a function of a public nature. While the court's eventual judgment was that the second respondent did not fall within section 6(3)(b) of the HRA there was significant division of opinion on this question. The dissenting opinions of Lord Bingham MR and Baroness Hale contain forceful arguments that private providers of public services of the kind of category that the second respondent fell into ought to be considered to be within the meaning of the Act. At least superficially, this is an indication of the challenge facing the courts in such instances, and the degree of disagreement in *YL* may well be indicative of the underlying ideological compasses that guide the interpretative activities of individual judges. Lord Bingham's core category of factors, which ought to influence a judgment one way or another, exposes this underlying tension. Lord Bingham noted that the presence of regulatory supervision, the existence of a statutory duty or obligation, the longstanding position of the state in reference to the provision of the service, and the funding relationship that exists all ought to be considered when deciding on the nature of the function being challenged.²⁰ In *YL*, Lord Bingham asserted, having weighed such relevant factors, that the claimant was entitled to the court's protection, not least because the courts themselves have a duty to protect human rights under the Act. Moreover, it was no answer to say that the claimant may have rights under private law for breach of contract. Lord Bingham made the perceptive and telling *obiter* observation that it was irrelevant if the body in question would not ordinarily be susceptible to judicial review, and further that there may be multiple parties with whom the claimant had a justifiable claim for breach of a Convention right:

Certain factors are in my opinion likely to be wholly or largely irrelevant to the decision whether a function is of a public nature. Thus it will not ordinarily matter whether the body in question is amenable to judicial review. Section 6(3)(b) extends the definition of public authority to cover bodies which are not public authorities but certain of whose functions are of a public nature, and it is therefore likely to include bodies which are not amenable to judicial

¹⁷ *Aston Cantlow* (n 11) [12].

¹⁸ *ibid.*

¹⁹ *ibid.*

²⁰ *YL* (n 2) [5]–[11].

review. In considering whether private body A is carrying out a function of a public nature, it is not likely to be relevant that public body B is potentially liable for breach of an individual's Convention right. The effect of the Act may be that both A and B are liable. It will in my opinion be irrelevant whether an act complained of as a breach of a Convention right is likely to be criminal or tortious: the most gross breaches of the Convention the improper taking of life, inhumane treatment, unjustified deprivation of liberty will ordinarily be both criminal and tortious.²¹

Lord Bingham exposed the weakness of the argument thus far propounded by the courts: if there is a remedy in private law, where a Convention right is also engaged because the offending party was a private entity exercising functions of a public nature, then the court will not subject the offending party to the jurisdiction of public law. This is evidently nonsensical. The presence of a private remedy points to the presence of a private remedy, it says nothing about whether it would also be appropriate and proportionate to extend the protection of public law relief to the matter under consideration. It may be the case, as was evident in *YL*, that the remedy available in private law would be wholly unsatisfactory to the circumstances of the case. Though the claimant in *YL* may have had a strong cause of action for breach of contract, and may well have received a remedy in private law, if the private provider was not held to fall within the scope of the Act, then there was no public law protection that could have been afforded. The claimant would not have been able to retain her place in the care home, something that would have been available through injunctive relief offered by the court in public law and of greater benefit to the claimant than any damages awarded in private law. Moreover, this failure to extend the protection of public law remedies existed where not doing so was likely to actively abrogate the claimant's Convention rights. There is certainly a case to be made that the private law rights of the respondent should not effectively trump, as a matter of policy, the individual's rights and fundamental freedoms guaranteed under the Convention.

This judicial reluctance to adopt a more purposive approach toward statutory interpretation has seen the development of an alternative argument aimed at narrowing the parameters of interpretation. Lord Scott, who was in the majority in *YL*, argued that to bring the second respondent within the meaning of section 6(3)(b) would have a floodgates effect on other private providers. His Lordship argued that it was 'absurd to suggest that the private contractor, in earning its commercial fee for its business services, is publicly funded or is carrying on a function of a public nature.'²² However, this argument fundamentally ignores a central point: a private contractor does not need to tender for government services. In creating a contract with a public authority to carry out the agreed services the private party is performing what would have otherwise have been provided by the public body or another private provider. Nor does Lord Scott's analysis accept, as both Lord Neuberger and Lord Mance's judgments do, that there will be instances where there will be contracts that do provide core public services and some contracts that do not. There seems no logical reason for saying that the provision of certain public services – like social and health care – sit outside the scope of the HRA, and are therefore not subject to control by the court simply because the service is provided for profit. If anything, profit motive, with its attendant qualities of cutting certain elements of service for the sake of efficiencies and heightened profit margins seem to make it riper for greater levels of judicial supervision. Lord Scott's reasoning is in keeping with a traditional English common law emphasis on the protection of private law rights. In Lord Scott's view the remedy available in tort or contract is deemed to be sufficient.

This argument is susceptible to the complaint that private law remedies do not have the same character of the relief provided by public law. For an 84 year old woman suffering from Alzheimer's

²¹ *YL* (n 2) [12].

²² *ibid* [27].

disease, the public law protections are of greater significance: the ability to stay in the home which has been provided, with all the attendant qualities of that stability for someone in need of care and dependent on the State is far more powerful an argument for public law protection than the availability of remedies for breach of contract, though these would also be available in a separate but parallel action. It is submitted that if it is manifestly inadequate to provide a private law remedy alone, and if in the interests of protecting a right possessed by an individual, the only remedy that appears appropriate to the court is a public law remedy, then it would seem logical to extend the scope of public law protection to privately outsourced contracts which provide services that would otherwise be provided by public authorities whether core or hybrid.²³

IX. WHY HAS THIS JUDICIAL APPROACH BEEN ADOPTED?

Why then has the judicial approach been to slow use flexible tests in public law to account for the systemic delivery of services and functions of a public nature by private entities? Certainly the judicial respect for Parliamentary sovereignty goes some way to explaining the current state of affairs. Governments must be allowed to conduct the effective administration of the State. Parliament legislates for the provision of certain services to be provided by public bodies overseen by a Minister of the Crown, who is accountable for all their actions to Parliament. If Parliament has legislated for the Minister or ancillary regulatory bodies administering a certain area of public policy or service provision to contract with a private provider, then the court may deem it entirely proper to allow the executive to retain unfettered discretion as to how those statutory duties will be fulfilled. In principle, there is nothing unsatisfactory about the provision of public services through contractual agreements with private parties. The issue is that the traditional *laissez-faire* attitude to letting the contract run, and seeking remedial action where the contract has been breached, ignores the vitally important status of the individual who benefits from the public service provided under the contract. Those individuals eligible for certain public services are unlikely to have any rights in respect of the discretionary power used to create the contract for the delivery of those services which are depended on. When deference to the executive is coupled with a judicial approach that prefers private law remedies, the courts' disinclination to widen the scope of public law is set in a broader socio-legal context. This stance causes a problematic conflict, in particular when the court is called upon to interpret section (6)(3)(b) of the HRA.

Moreover, an argument that runs along deference lines fails to take sufficient account of the Parliamentary intent behind section 6(1) of the HRA. By 1998, the contractualisation of government services and the formation of hybrid public service provision was substantially entrenched; to think that Parliament did not have in mind the susceptibility of a private company that was running a prison, or a hospital, and who therefore might have some regulatory powers or duties in the exercise of those functions, seems a particularly fanciful reading of the Act. It would perhaps be more appropriate, given the background of public service provision since 1979, to interpret the Act purposively and read into it Parliament's intent that hybrid public authorities, or private entities carrying out public functions under contract, are susceptible to the public law jurisdiction of the court.

²³ Paul Craig, 'Contracting Out, the Human Rights Act and the Scope of Judicial Review' (2002) 118 LQR 551.

The decision of the majority in *YL* does however reflect a reasonable concern that must be proportionately considered against the background of the individual rights of the claimant. Specifically, by not drawing the second respondent into the scope of the public law, there is greater certainty as to what liability private parties will be susceptible to when entering into contracts with public bodies. The central complaint with this argument is that the individuals for whom the service is provided are forgotten, or at least deemed to be of less importance than the private parties right to be excluded from the scope of public law interference. Arguably, there needs to be certainty for parties entering into contracts providing public services, but there appears to be no good reason, on the ground of protecting private parties, simply because they are private parties, from the exercise of the courts' public law jurisdiction when that private party fails to provide the function which they were contracted to provide. The consequences for an individual's rights is an unfortunate and potentially unintended consequence of the state's increasing reliance and belief in the private provision of services.

X. THE BLURRED BOUNDARY BETWEEN THE PROVISION OF PUBLIC SERVICES AND THE PRIVATE FUNCTIONS OF GOVERNMENT

Many government contracts may exist which will never engage a Convention right, create grounds for judicial review, or become the subject of an application for judicial protection. For example, it is difficult to foresee the success of an application relating to the purchasing of stationery by a public authority. Given the administrative scale of the modern nation state it is appropriate that there are some contractual arrangements which the courts will decline to interfere with, on the grounds that to interfere will open the floodgates for unreasonable or irrelevant claims in all sorts of fields where there is interaction between the administrative state and the private sector. There may also be reasonable disputes as to whether a particular contractual arrangement amounts to a public service, or whether it is simply a function of the ordinary running of government. This dispute could be resolved through a fairly straightforward consideration of the public benefit derived from the function. A private contract to provide cleaning services in Ministerial departments is unlikely to affect the lives of individuals on such a level as to require the court to exercise its public law jurisdiction if a dispute arises. Disputes arising in such a situation would ordinarily involve straightforward principles of contract or tort law.

Conversely, in situations such as the provision of hospitals or care homes, education, pensions or benefits, the public benefit is evident and there is a weaker argument for holding that this should not be susceptible to the public law jurisdiction of the courts. These two examples offer an obvious contrast where it is easier to distinguish between a service benefiting the public, and a function of government with no fundamental benefit to the public. Though there may be policy disagreement as to what the government should and should not do, it is not necessary for the courts to delve too deeply into this inquiry. In the majority of applications where a respondent is a private person or entity, the contract that they hold with a public body will have some statutory duty, obligation or power attached to it. The court will be able to clearly distinguish between a contract made to facilitate a statutory duty to provide a public service, and a contract to facilitate internal functions of the government itself. The court's ability to trace the duty as part of the test to establish the public nature of the function ought to be enough to negate any floodgates argument that all contracts between private parties and the State would fall inside the jurisdictional scope of public law.

XI. JUDICIAL REVIEW AND THE ADAPTABLE SCOPE OF JURISDICTION IN PUBLIC LAW JURISPRUDENCE FOLLOWING DATAFIN

While *YL* offers a more recent consideration of the boundaries of public law, it is confined to the court's statutory interpretation of the provisions within HRA. In order to appreciate the difficulty posed by the decision in *YL*, it is helpful to note that the courts have previously seen fit to extend the jurisdiction of public law to other areas of quasi-public interaction. The governing procedural mechanism for such matters is now section 31 of the Senior Courts Act 1981. Notably, the Court of Appeal's decision to extend the scope of judicial review in *Datafin*,²⁴ consequently bringing into public law jurisdiction a quasi-judicial panel without any statutory or regulatory foundation, emphasized the adaptability of the judiciary's public law jurisdiction. In *Datafin*, the applicants were in bidding competition with the defendants to take over another company. They complained to the Panel of Takeovers and Mergers that the defendant had acted in concert with other parties in breach of the City Code on takeovers and mergers. The Panel dismissed the applicant's complaint. The applicant applied to the High Court seeking, *inter alia*, certiorari to quash the Panel's decision and mandamus to compel the Panel to reconsider the complaint. The High Court judge refused permission on the ground that the Panel's decision was not susceptible to judicial review. The Court of Appeal granted leave in order to consider the question of jurisdiction and the substantive issues that would have been considered on review. Sir Donaldson MR reversed the decision of the High Court judge finding that the panel were susceptible to judicial review however refused the application on the substantive merits.

The case was important because it clearly demonstrated the flexibility of a public law framework that would adapt the jurisdiction of the court based on the presence of certain factors, though these factors need not be exhaustive or all present to attract the court's jurisdiction. First, the Panel, though self-regulated, was 'supported and sustained by a periphery of statutory powers and penalties'. Second, if the decisions of the Panel were deemed to be *ultra vires* then public law remedies would apply.²⁵ Finally, determining whether the body in question was providing a public function, which would affect the citizenry, was crucial in exercising the supervisory jurisdiction of the court under section 31. Secondary matters were also considered by the court, including the fact that though the panel was not founded on a statutory instrument, its lack of a statutory foundation was a 'complete anomaly, judged by the experience of other comparable markets world-wide'.²⁶ Most importantly, Sir Donaldson MR suggested, obiter, that possibly the only crucial factor that had to be present to exercise the public law jurisdiction of the court was that the decision questioned had to have a 'public element'.²⁷ What exactly constituted a public element was kept purposefully vague. It is submitted that there was good reason to introduce such flexibility into the definition. Requiring the fulfilment of an exhaustive, or even a non-exhaustive list of factors which would exercise the jurisdiction of the court's supervisory powers only limited the possibility that future applications would be refused, whether meritorious or not, based on a test that would ultimately fall behind the institutional management of matters affecting the public. Why exactly this flexibility has been departed from in respect of section 6 claims is unclear. It is possible that there are rule of law concerns, particularly in respect of the uncertainty that a vague test of public element introduces into the provision of public services by private bodies. But it is submitted that the degree of uncertainty in this area is to be tolerated where the alternative is the abrogation of individual rights, without recourse to appropriate remedy, simply because the breaching party is a private entity.

²⁴ *R v Panel on Takeovers and Mergers, ex parte Datafin PLC and another* [1986] 1 QB 815 (CA).

²⁵ *ibid* 835 (Sir Donaldson MR).

²⁶ *ibid*.

²⁷ *ibid* 838.

In *YL*, Lord Neuberger and Lord Mance both acknowledged that there was a connection between section 6 cases and the earlier cases concerning section 31 of the Senior Courts Act. Specifically, Lord Mance considered the connection between '[t]he existence and source of any special powers or duties', which was considered of fundamental importance in *Datafin*, and the view a court takes when 'considering whether state responsibility is engaged in Strasbourg or whether section 6(3)(b) applies domestically.'²⁸ Similarly, Lord Neuberger accepted that though *Aston Cantlow* had not deemed the existence of statutory power as being sufficient to bring the action complained of automatically into the jurisdiction of the court, if the overarching framework was one that could be categorized as a 'relatively wide-ranging and intrusive set of statutory powers in favour of the entity carrying out the function in question', then such a state of affairs would be 'a very powerful factor in favour of the function falling within section 6(3)(b). Indeed, it may well be determinative in many cases, because such powers are very powerfully indicative of a public institution or service.'²⁹ Though neither Lord Neuberger nor Lord Mance felt that such a framework existed on the facts, it is interesting to note that they were willing to accept that such a statutory arrangement may well be determinative of a situation falling within the scope of the Act. However, what is perhaps more interesting is that neither judgments considered the dicta of Sir Donaldson MR on the question of the public element. It is submitted that this element of the decision in *Datafin* ought to have been considered because it provides the earlier broader scope test for determining whether an action is a public function or not. Without explicitly departing from this earlier possible test of scope there is still inherent uncertainty in the law in respect of the matters considered in this article.

XII. CONCLUSION

This article has considered the inadequacy of the judicial response to the increasing contractualisation of government and the implications such an approach to service delivery has on the individual rights of those who are entitled to government support and protection. It must be hoped that the slide toward the privatization of public services through outsourcing does not constitute an increasing restriction of access to public law protections and remedies for individuals whose human rights are arguably engaged by providers who take the benefit of public contracts without bearing the responsibilities to act like public authorities. To arrest this slide, far more flexible tests are required when considering what constitutes a public function under section 6.

It is submitted that the appropriate test should consider the following non-exhaustive factors when deciding whether the exercise of a power or the provision of a service is public in nature: first, whether the provision of the service engages a Convention right; second, whether the service is one which the recipient is entitled to by statutory provision; third, whether the only adequate remedy is one provided by public law supervision. Finally, it is submitted that the majority judgment in *YL* significantly departed from the pre-HRA case law and set out an erroneous case that private parties are to be protected above the individuals who depend on the services provided by private contract. It is worrying that such judicial reticence has failed to grasp the necessity of extending public law protections in an era where outsourcing and privatization is the preferred method of service delivery. It is hoped that such a narrow approach gives way in time to a more flexible and adaptable scope of interpretation and review.

²⁸ *YL* (n 2) [102].

²⁹ *ibid* [167].

Something for Nothing: Explaining Single-Sided Contract Variations

Andrew Hill*

INTRODUCTION

Explaining single-sided contract variations is troublesome. This article addresses the situation where one party varies its obligations under a contract, either by increasing its rights or reducing its duties, whilst the other party maintains the same obligations. The modern case law holds that, in certain circumstances, such variations are valid, but that in others, they are not. Focusing on these cases, this article has two objectives. One is to present a negative argument, that the prevailing theories cannot explain the patterns of validity and invalidity which the case law has developed, and that those theories are, in any case, problematic. The other is to offer a positive argument, that such variations can be understood through an implied terms analysis.

I will call the situation which this article considers a ‘single-sided variation’ to a contract. There are two variants. First, one party to a contract may request additional performance from the other party in return for no additional performance from itself (a ‘more for the same’ scenario).¹ Second, one party may request that it should render less performance in discharge of its contractual duty in return for the same performance from the other party (a ‘less for the same’). *Stilk v Myrick*² and *Williams v Roffey Bros*³ govern ‘more for the same’ scenarios, and *Foakes v Beer*⁴ and *Re Selectmove*⁵ govern ‘less for the same’ scenarios. These authorities are discussed in Section I.

Single-sided variations are problematic because of the doctrine of consideration. Consideration is something of value (either a benefit or detriment) given or promised by the promisee in return for the promisor’s promise.⁶ Only promises backed by valid consideration from both sides are enforceable. This is critical for present purposes because, when a party requires a single-sided variation to a contract, it is promised more or it reduces its own obligation. However, it seemingly fails to offer any consideration in return for this beneficial alteration in its own rights/duties. Therefore, the contract variation seems to lack consideration from one party, and hence be invalid.

Section II will consider why simply distinguishing the relevant cases is an untenable suggestion. Sections III to V will then consider three existing theories which attempt to explain single-sided variations: Unilateral Variation Contracts, Duress and Promissory Estoppel. They will be the subject of my negative argument, that these theories cannot explain the operation of single-sided variations given the present state of the case law.

*Lady Margaret Hall, Oxford. I would like to thank Mr Niranjana Venkatesan and Professor Alexandra Braun for their comments and advice on a draft of this article, and the editorial team at The Oxford University Undergraduate Law Journal. Any errors are my own.

¹ The ‘more for the same’ and ‘less for the same’ terminology are adapted from Chen-Wishart, *Contract Law* (4th edn, OUP 2010) 125

² (1809) 2 Camp 317 (KB).

³ [1991] 1 QB 1 (CA).

⁴ (1884) 9 App Cas 605 (HL).

⁵ [1995] 1 W.L.R. 474 (CA).

⁶ *Currie v Misa* (1875) LR 10 Ex 153 (HL).

Instead, the positive argument of this article is that a solution can be found in the modern law of implication of terms. I will argue that implying a Variation Term, under the approach to implication most recently stated by Lord Hoffmann in *Attorney General of Belize v Belize Telecom*,⁷ facilitates single-sided variations and explains the case law. This argument is set out in Section VI.

I. THE CASE LAW

First, we will consider the relevant case law which provides the foundation to the theories. The ‘more for the same’ and ‘less for the same’ scenarios are considered in turn.

First, however, it is useful to define the scope of the single-sided variation situation. Many contract variations are unquestionably valid. If both sides offer fresh consideration, then the contract is valid. The two parties could agree to end their previous contract and form a new one. A deed would also suffice to make a gratuitous promise (to give more/accept less in satisfaction of the contract). Therefore, these situations do not conflict with the doctrine of consideration, and so are beyond the reach of this article.

A. ‘MORE FOR THE SAME’ SCENARIOS’

Stilk v Myrick

Stilk is the foundational case for the modern law on single-sided contract variations. *Stilk* was one of eleven crew members on a ship serving under Myrick. His contract said that he would be paid £5 per month in return for doing everything that was needed in the voyage. Midway through the voyage, two of the crew deserted. Myrick therefore promised the remaining crew that, if he could not find two more crewmen (which he could not), he would divide the two deserters’ wages amongst the rest of the crew so that they should take on the roles of the deserters as well as their own for the return voyage. The nine remaining crew members sailed the ship home. Myrick then only paid them their original wages. The crew members brought an action on the contract for the additional wages they had been promised, being one-ninth shares of the two deserters’ wages. The court found for Myrick, denying the extra payments.

Campbell’s report⁸ says that the varied contract was invalid because of lack of consideration⁹. Lord Ellenborough explains that the sailors had already undertaken to do everything necessary, which included, if need be, taking on the roles of any deserters. Therefore, they had given nothing in return for the promise to increase their wages, so the agreement was void for want of consideration.¹⁰ To state this conclusion in a different manner, the promise of more rights (wages) in return for the same performance was held to be invalid.

⁷ [2009] UKPC 10, [2009] 1 WLR 1988.

⁸ *Stilk* (n 2).

⁹ The basis of the decision is evidentially problematic. *Espinasse*, 6 Exp 129, reports that the decision was based on grounds of public policy, following *Harris v Watson* Peak. Cas. 72, that it was undesirable to allow sailors to demand pay rises in return for seeing the ship home. *Espinasse*’s reported reasons therefore directly contradict those given by Campbell. The weight of modern authority, such as exists, rests behind Campbell’s report. *North Ocean Shipping v Hyundai Construction* [1979] QB 705 (HC) approved Campbell’s report for its better reputation and the Court of Appeal in *Williams* accepted Campbell’s report.

¹⁰ *Stilk* (n 2) 319.

Williams v Roffey Bros

The second 'more for the same' case is *Williams*. Shepherds Bush Housing Association contracted with Roffey to refurbish 27 flats. Roffey sub-contracted carpentry work to Williams, agreeing to pay them £20,000 in instalments. When Williams had one task still to complete in 18 of the flats, he informed Roffey that he was in financial difficulty. Therefore, he might be unable to complete the remaining work unless he received more money. At that time, Roffey had paid Williams £16,200. Roffey was subject to penalties under the main contract with Shepherds Bush Housing Association, so it needed prompt completion of the services. Therefore, Roffey offered to pay Williams £10,300 in addition to the original £20,000, at £575 for each of the remaining 18 flats. However, Roffey only paid another £1,500 over the next two months. This caused Williams to cease work on the flats. The work in eight of the 18 flats had been substantially completed.

Williams brought an action on the contract claiming the original sum owed plus the £10,300 under the variation. The Court of Appeal dismissed Roffey's appeal. The damages awarded to Williams at first instance, £4,600 (to represent the completion of eight flats minus costs, plus a reasonable sum due under the original contract), were upheld. Glidewell LJ summarised the law thus:

...the present state of the law on this subject can be expressed in the following proposition: (i) if A has entered into a contract with B to do work for, or to supply goods or services to, B in return for payment by B; and (ii) at some stage before A has completely performed his obligations under the contract B has reason to doubt whether A will, or will be able to, complete his side of the bargain; and (iii) B thereupon promises A an additional payment in return for A's promise to perform his contractual obligations on time; and (iv) *as a result of giving his promise, B obtains in practice a benefit, or obviates a disbenefit*; and (v) B's promise is not given as a result of economic duress or fraud on the part of A; then (vi) *the benefit to B is capable of being consideration for B's promise, so that the promise will be legally binding*.¹¹ (emphasis added)

Thus, the consideration which Williams had given under the variation contract was providing Roffey with practical benefits, or obviating practical disbenefits. The Court identified four such benefits: (1) continued performance; (2) avoiding the difficulties of finding an alternative carpenter; (3) avoiding liability under the penalty clauses in the main contract; and (4) performing the work in a more orderly manner. Thus, both sides to the variation contract had provided consideration, and so the variation contract was valid.

The Court of Appeal recognised the potential conflict between *Stilk* and the decision that they were making. The following passage from Russell LJ encapsulates the general tenor of the distinction drawn:

...I do not base my judgment upon any reservation as to the correctness of the law long ago enunciated in *Stilk v. Myrick*. A gratuitous promise, pure and simple, remains unenforceable unless given under seal. But where, as in this case, a party undertakes to make a payment because by so doing it will gain an advantage arising out of the continuing relationship with the promisee the new bargain will not fail for want of consideration.¹²

¹¹ *Stilk* (n 2) 15-16.

¹² *ibid* 19.

The principle that may be drawn from these findings is that the existence of a practical benefit is good consideration.¹³ When the legal obligations under the contract are not altered, a party still provides adequate consideration by offering a practical benefit, or obviating a practical disbenefit, in fact. By contrast, the principle in *Stilk v Myrick* is that a single-sided variation, where the other party received no benefit, is invalid for want of consideration. While this may have increased the economic efficiency of the transaction on the facts, Section II will explain, in light of the analysis of the ‘less for the same’ scenarios, why the acceptance of ‘practical benefits’ is unconvincing.

B. ‘LESS FOR THE SAME’ SCENARIOS

Foakes v Beer

The reverse of the ‘more for the same’ scenario is the ‘less for the same’ scenario. Here one party will *reduce*, rather than increase, their rights or duties. *Foakes*¹⁴ is the principal case. Dr Foakes owed Mrs Beer a sum of money after a court judgment. When Foakes complained that he could not afford full payments immediately, he and Beer made a written agreement that he could pay in instalments. The agreement made no mention of the need to pay any interest. Foakes paid the original sum as agreed, but no interest. Beer brought an action on the contract to claim the interest payments. Foakes argued that the variation meant that she could not enforce the sum. Beer argued that the single-sided variation was invalid because Foakes had not provided any consideration.

The House of Lords found for Beer. Their Lordships held that an agreement whereby one party agrees to accept less performance in return for the other party accepting the same performance was invalid for want of consideration.¹⁵ As Lord Fitzgerald observed, a bare agreement arises when one party fails to give consideration, and it is a rule that a bare agreement does not give rise to a legal obligation.¹⁶

In re Selectmove

The second ‘less for the same’ case is *Re Selectmove*. Selectmove owed outstanding tax and national insurance payments. The company director agreed with the Inland Revenue that Selectmove would only have to pay future tax and national insurance contributions, and that these could be made at an agreed rate. The director had pointed out to the collector that the company was in financial difficulties, so it was better to implement a reduced payment plan than to issue a winding up order on Selectmove, which would mean that none of the outstanding payments would be recovered. Later, having received insufficient payment from Selectmove, the Inland Revenue brought an action for all the outstanding payments owed.

Selectmove argued that the agreement had reduced the debt which it owed. The main issue argued by counsel for the Inland Revenue was the seeming want of consideration. The court faced two apparently contradictory authorities: *Foakes* said that an agreement to accept part-payment of a debt in discharge of a whole was invalid for want of consideration. However, *Williams* said that obtaining a practical benefit was good consideration. Selectmove argued that the agreement entailed a practical benefit because the reduced rate made it feasible for the company to make payments.

¹³ *Adam Opel v Mitras Automotive* [2008] EWHC 3205, [2008] CILL 2561.

¹⁴ *Foakes* (n 4)

¹⁵ See Earl of Selborne LC at 611, 613 – 614 and Lord Blackburn at 615–616, 621–623. Lord Watson dissented on the grounds of the construction of the document in question, and did not consider the issue of consideration. There were other reservations, notably expressed by Lord Blackburn, but all of their Lordships ultimately agreed on the issue of law relating to consideration raised by the case.

¹⁶ *ibid* 629–630.

The Court of Appeal found for the Inland Revenue. The variation, the agreement to accept part-payment of the debt in discharge of the whole, was invalid for want of consideration. The court took *Foakes v Beer* as their binding authority, and rejected the practical benefit argument from *Williams*, distinguishing it. Peter Gibson LJ explained that, if they were to accept *Williams*, it would leave *Foakes* without any application, so *Williams* could not be extended.¹⁷ Thus, the Court of Appeal rejected the application of practical benefits in the context of ‘less for the same’ variations.

II. STATE OF THE PRESENT LAW

A. SUMMARY

One way of explaining the present law, based on the practical benefit view, is that these four cases might be distinguished from one another. This would cause us to state the ‘law’ in the following terms. In the context of ‘more for the same’ variations, for a variation to be valid, fresh and valid consideration must be given by both parties, not just the promisor (*Stilk*). However, a practical benefit is sufficient consideration in such circumstances (*Williams*). In the context of ‘less for the same’ variations, for the variation to be valid, fresh and valid consideration must be given by both parties, not just the promisor (*Foakes*). A practical benefit does not amount to valid consideration; some new legal benefit must be promised (*Re Selectmove*).

B. THE LOGICAL PROBLEM

There is a logical problem in this explanation. The distinction is arbitrary. Agreeing to increase one’s duties in return for nothing (‘more for the same’) or agreeing to reduce one’s rights in return for nothing (‘less for the same’) are simply the reverse of each other. In both cases, one party suffers a legal detriment either in losing rights or gaining duties, and the other party thereby receives a benefit in either losing duties or gaining rights. Therefore, there is no principled reason why practical benefits should only be applicable in one of these scenarios, because they can be given equally in return for a loss of rights (for which they are currently not allowed) or an increase in duties (for which they currently are allowed). Thus, in principle, practical benefits should apply to both or to neither.¹⁸

C. THE PRECEDENTIAL PROBLEM

It was, as a matter of potential interpretation, not even open to the Court of Appeal to conclude that a practical benefit is adequate consideration in *Williams*. We can observe this by re-examining *Stilk v Myrick*. We have already established that *Stilk* stands for the proposition that, when only one side receives a benefit or undertakes a detriment under a contract-variation, that variation is invalid. As a matter of interpretation *Stilk* seems to stand for a second proposition, that practical benefits are not valid consideration. To observe this, we must run the concept of practical benefits back through *Stilk*.

¹⁷ *Re Selectmove* (n 5) 481.

¹⁸ It is noted that the same distinction is drawn by the law of promissory estoppel. However, consideration of the distinction in that context is beyond the scope of this article. It suffices to note, for present purposes, that it seems unsatisfactory to justify an arbitrary distinction in one area of the law by pointing to an equally arbitrary (and contentious) distinction drawn elsewhere, and thus an analogy with promissory estoppel would be unhelpful.

In *Stilk*, there was no legal benefit to the ship's master, because his rights remained the same. There was a legal benefit to the crew, because their rights were increased by the promise of additional pay. The court, therefore, declined to recognise the variation contract, because only one side received a new legal benefit, and neither side undertook a new legal detriment. However, there were two practical benefits to the ship's master, because he (1) got the ship home and (2) did not have to expend effort into finding replacement crew members. Indeed, these mirror the first two practical benefits in *Williams* ((1) continued performance and (2) avoiding the difficulty of finding a replacement contractor). There was also practical detriment to the crew, because they had to work proportionately harder to make up for the two missing sailors. Therefore, the crew members each received a legal benefit – increased pay. The ship's master received a practical benefit – he got the ship home safely – and obviated a practical disbenefit – he did not have to find replacement crew. Therefore, under the rule from *Williams*, adequate consideration was provided by both sides in *Stilk*. Hence, accepting *Williams*, *Stilk* would have been decided differently today.¹⁹

Before reflecting on what this says of *Williams*, we must reflect on what we learn of *Stilk*. The court in *Stilk* refused to recognise any valid consideration in the practical benefits obtained and disbenefits obviated. Therefore, *Stilk* can be seen to stand for a second proposition, that only a *legal* benefit/detriment is sufficient consideration; a practical benefit will not suffice.

The revised reading of *Stilk* seemingly places it in direct conflict with *Williams*. Looking once again at *Foakes*, we see that it also is in conflict, for it also rejected a practical benefit as insufficient. Although Beer received a practical benefit in actually securing payments from Foakes, the court denied that Foakes had given any consideration and, thereby, rejected the adequacy of practical benefits once again. Indeed, their Lordships were acutely aware of the commercial benefit obtained by single-sided variations in certain situations, yet this could not overcome the fact that there was no variation in Foakes' obligations so as to amount to valid consideration.²⁰ Given that *Re Selectmove* was also decided after *Williams*, so the 'more for the same'/'less for the same' distinction had yet to be made, the Court of Appeal in *Williams* was not only following *Stilk* but also the House of Lords in *Foakes*. While referring to *Stilk*, the judgments in *Williams* make no mention of *Foakes*, and it was not cited by council. Thus, the court in *Williams* appears to have been bound to reject practical benefits, and thus fell into error by failing to do so.

D. CONCLUSION ON THE DISTINCTION

Therefore, given the logical problem and the precedential problem, we must reject the suggestion that our difficulties may be resolved by distinguishing 'more for the same' and 'less for the same' cases, and applying practical benefits in the former but not the latter.

III. UNILATERAL VARIATION CONTRACTS

Thus far, we have examined the legal propositions which the cases establish, and have seen that simply drawing a distinction between them is an unsatisfactory solution. With this established, this article's negative argument can be made. Thus, in this section and the following two, three prominent explanations offered for the cases – Unilateral Variation Contracts, Duress and Promissory Estoppel – are examined and rejected.

¹⁹ Treitel, *The Law of Contract*, (13th edn Sweet & Maxwell, 2011) 99.

²⁰ See, for example, Lord Blackburn at 622-623

A. THE THEORY

Chen-Wishart has proposed that a double contract analysis is necessary.²¹ The original contract is an ordinary bilateral contract. The variation forms a second, unilateral contract. Consideration in the original contract consists of the *promise* to perform the service, in return for the promise of payment. Consideration in the unilateral contract consists of the *performance* of the services, in return for the increase in payment.²² In *Williams*, the bilateral contract was to perform three services on the 27 flats for £20,000. *Williams* *promised* those carpentry services, in return for Roffey's *promise* to pay. The unilateral variation contract was Roffey's *promise* to pay the additional £10,300 at £575 for each completed flat, in return for *Williams* *actually completing* the remaining 18 flats. This view explains *Williams* in terms of legal benefits alone, because legally different consideration is provided in both contracts; the first is the promise to perform (the necessary consideration in a bilateral contract), and the second is actual performance (the necessary consideration in a unilateral contract). Furthermore, there would be no issue of conflict between the two contracts, because terms of a valid unilateral contract formed later would prevail over contradictory ones in an earlier bilateral contract.²³

This approach also explains the damages award in *Williams*. Recall that the damages awarded were not the full varied price which Roffey had offered, but rather a sum which roughly reflected the fact that eight flats had been completed, but another 10 remained outstanding. Chen-Wishart proposes two alternative ways of conceptualising the variation contract.²⁴ First, the variation actually involves 18 separate unilateral contracts, one per flat. Acceptance and consideration in unilateral contracts constitutes full performance.²⁵ Thus, when *Williams* only serviced eight flats, he only accepted and gave consideration for eight contracts, and hence could only claim the sums due under those contracts, £575 each. However, it may have well surprised the parties in *Williams* to know that they potentially had 19 operative contracts, not just one.

B. PROBLEMS WITH THE SINGLE UNILATERAL CONTRACT ALTERNATIVE

Chen-Wishart's second proposal consists of only a single unilateral contract. Acceptance and consideration would be completing performance on the 18 flats. However, when the dispute re-arose and performance ceased, *Williams* had not completed their acceptance of the contract. Therefore, the unilateral contract had not been fully accepted, and hence had yet to form.

As a result, the law on prevention of revocations after partly-performed acceptance in unilateral contracts must be applied. To prevent a revocation of the offer once the acceptor's performance has begun, the part performance can constitute the acceptance so long as it is later followed by full completion of performance.²⁶ The effect of this rule is to retrospectively render the unilateral contract valid once performance has commenced, subject to due completion of performance at a later time, when the validity of the contract is challenged by an attempted revocation. However, this would not work on the facts of *Williams v Roffey*, as *Williams* never completed performance, and hence the unilateral contract between *Williams* and Roffey would never have been valid.

²¹ Chen-Wishart, 'A Bird in the Hand: Consideration and Promissory Estoppel' in A Burrows and E Peel (eds), *Contract Formation and Parties*, (OUP 2010) 89-113, at 92-102.

²² Service and payment here are used by analogy with *Williams v Roffey*, but of course the principle is not limited to services and payments in this order, or indeed to cases with mixture of services and payments (see, e.g., *Vanbergen v St Edmund Properties* [1933] 2 KB 223 (CA)).

²³ *City and Westminster Properties (1934) Ltd v Mudd* [1959] Ch 129 (HC).

²⁴ Chen-Wishart (n 21).

²⁵ *Errington v Errington and Woods* [1952] 1 KB 290 (CA), 295.

²⁶ *Errington* (n 25); *Soulsbury v Soulsbury* [2007] EWCA Civ 969, [2008] Fam 1.

Instead, dictum by Goff LJ in *Daulia v Four Millbank Nominees*²⁷ must be relied on. He tells us that there is an implied obligation, arising when performance commences, on the offeror not to prevent the performance in acceptance of the contract. This dictum is susceptible to two separate analyses. First, it could be said that the implied obligation is actually an implied condition of the offer. So, A offers to pay £X to B if B walks from London to York, or performs this task to the extent he is not prevented by A from performing. Applying this to *Williams*, one would say that Roffey offered to pay the additional £10,300 in return for Williams completing the remaining 18 flats, or completing them to the extent that he was not prevented from doing so by Roffey. This cannot work. Roffey did not *prevent* Williams from completing the flats. Williams voluntarily ceased performance when Roffey did not pay adequate sums. He ceased to accept, and thus impliedly rejected, Roffey's offer. Only then did Roffey expel Williams from the site. Furthermore, since the unilateral contract had yet to form, Roffey's obligation to pay had not yet crystallised, and thus Roffey had done everything which he was obliged to do, so he had not 'prevented' performance by failing to fulfil his own legal obligations. Thus, *Daulia* cannot be successfully analysed as implying a condition into the offer on the facts of *Williams*.

The alternative is to hold that the implied obligation requires a two-contract analysis of the unilateral contract. Under this analysis, the dictum maintains the rule that a unilateral contract is only formed once full performance has been rendered. Instead, the implied obligation to allow performance without hindrance must arise from a collateral contract which is validly formed by the commencement of performance by the promisee in the unilateral contract.²⁸ Therefore, Roffey was subject to an implied duty not to prevent Williams' acceptance by performance. This analysis, once again, collides with the difficulty that it was Williams who ceased performance, and only then did Roffey make performance impossible by expelling him from the site. Therefore, Roffey did not appear to breach the collateral contract. It might be objected that the collateral contract should take no account of Williams' desistance from performance, and it only imposed an absolute obligation on Roffey not to prevent performance. This is incredibly artificial, because it verges on the irrational to impose an absolute obligation on one party to permit performance, regardless of the actions of the other. However, it is perhaps theoretically possible. Thus, we would say, by expelling Williams from the site, Roffey breached this collateral contract. However, the main unilateral contract still never came into existence for want of Williams' full performance. Therefore, the damages arose not through breach of the variation contract, but through breach of the collateral contract to the proposed variation contract.

The standard rule for damages is the performance, often called the expectation, measure.²⁹ What is the performance measure in the collateral contract? The answer will be the answer to the question "Where would Williams have been had the collateral contract been performed by Roffey?" However, this hypothetical contains a crucial uncertainty. If the collateral contract had been performed, Williams would have been able to perform. However, this was not the end. If Williams *had then gone on to perform*, he could also have expected the additional payments under the unilateral contract. However, Williams might have *refused to perform*, in whole or in part. This would mean that he had no expectation under the unilateral contract, and indeed that he may be liable for breach of the main bilateral contract. In between these two extremes, there are all manner of potential part-performance outcomes. Therefore, we cannot say what the performance measure was, because we cannot know where exactly, 'but for' Roffey's breach, the situation would have ended.

²⁷ [1978] Ch 231 (CA), 239.

²⁸ Examining the apparent issues of consideration and communication of acceptance for this collateral contract are beyond the scope of this article. We will assume this collateral contract is validly formed.

²⁹ Friedmann, 'The Performance Interest in Contract Damages' (1995) 111 LQR 628. See also *Robinson v Harman* (1848) 1 Ex Rep 850, especially 855.

Such hypothetical uncertainty is not, however, unknown to the courts. In the Court of Appeal in *Walford v Miles*, Bingham LJ tells us that, when dealing with such uncertainty of hypothetical outcome, the minimum expectation formed is that which is most favourable to the innocent party (i.e. the party who is not in breach).³⁰ The most favourable outcome for Williams was, therefore, that he fully performed all the services, because that would have entitled him to payment under the unilateral variation contract, and not left him in breach of the bilateral contract. The performance measure under the unilateral contract was £10,300. Therefore, that was also the performance measure under the collateral contract. However, Williams was not awarded the performance measure, but rather seemingly a reliance measure. Therefore, unless we conclude that the court was in error in awarding damages, *Daulia* cannot explain *Williams* by a collateral contract analysis either. Thus, both of *Daulia*'s explanations have failed on the facts of *Williams*, so we have no workable explanation of how the variation in *Williams* can be conceptualised as a single unilateral contract.

There is a further oddity with the single unilateral contract explanation too. The Court awarded damages to Williams because they held that Roffey was in breach. Roffey was in breach because they declined to pay the variation payments in addition to the normal instalments which they owed. To owe the variation payments, Roffey must be bound by the variation contract. Although they were bound to allow the variation contract to be formed, it does not follow that the variation contract was already valid. Because the unilateral contract was only valid once Williams' rendered full performance, Roffey was not yet subject to the duty to pay the variation payments. Therefore, Roffey was not in breach of any existing contractual obligation when they failed to pay the variation payments. Thus, the court should have rejected Williams' case for want of breach. Therefore, the single unilateral contract proposal seems to fail on the facts, so *Williams* must be construed as consisting of 18 separate unilateral variation contracts.

C. PROBLEMS WITH APPLICATION TO OTHER CASES

Whichever approach is taken in respect of *Williams* itself - whether one variation contract or 18 - the unilateral variation contract view is flawed because it is incapable of wider application. Quite simply, the application of the principle to *Stilk*, *Foakes* and *Re Selectmove* would have led to the opposite conclusion in those cases.

In *Stilk*, the bilateral contract was the original employment contract between Stilk and Myrick. The unilateral variation was the promise to pay more in return for *actually* working harder to fill in for the missing men, and hence getting the ship home. The performance was given in full (because the remaining crew put in the additional work for the rest of the voyage), so acceptance and consideration for the unilateral contract was given, making it valid. Therefore, the Unilateral Variation Contract theory says the variation in *Stilk* should have been legally valid.

The same issue arises in *Foakes*. The original 'bilateral contract' was the court ordered payment. The unilateral contract was the promise to accept payment in instalments without interest in consideration for *actually* receiving payment. Foakes duly paid the amount owed under the variation in full, so he gave valid acceptance and consideration, making the unilateral contract validly formed. Hence, again, the Unilateral Variation Contract theory suggests the variation should have been valid, where the House of Lords said it was not.

³⁰ See, e.g. Bingham LJ in *Walford v Miles* (1991) 62 P. & C.R. 410 (CA) at 422 – 423. Bingham LJ was forced to dissent for other reasons (specifically that he held that a non-time-limited lock-out agreement could be enforceable). The House of Lords ([1992] 2 AC 128) decided the case on the same basis as the majority, but did not consider this damages point.

Finally, *Re Selectmove* would also have been differently decided. The original ‘bilateral contract’ was to make the contributions as required by law. The unilateral variation contract was the promise to set a new, later date from when contributions would be outstanding and to reduce the monthly payments in consideration for *actually* attaining payment. Selectmove had begun, but not completed, payment when the Inland Revenue served the action and issued the winding up order. Therefore, although the unilateral contract itself was not valid for want of full acceptance and consideration, following the *Daulia* reasoning which also had to be applied to *Williams*, there was an implied condition or an implied collateral duty upon the Inland Revenue to allow Selectmove to make payments in accordance with the new agreement. Therefore, it should have been the Inland Revenue, not Selectmove, who were in breach.

D. PROBLEMS WITH FRUSTRATION

There is a further problem with the Unilateral Variation Contract theory, which arises if the performance in acceptance of the variation is frustrated. Although *Errington* and *Daulia* prevent the promisee from denying the variation by preventing performance, they do not apply when a frustrating event prevents performance. *Errington* assumes that the unilateral contract is not valid unless the performance is eventually completed. *Daulia* assumes that it is not valid until either performance is completed or performance is completed to the extent which the promisor permits. A frustrating event, not caused by either of the parties, may intervene, making full performance impossible. That will rule out the *Errington* rule, and the first alternative from *Daulia*. Since the promisor did not cause the frustration, the second ground of *Daulia* is also unavailable. Therefore, there is no valid unilateral contract, and no assistance from either *Errington* or *Daulia*. Hence, when, in applying the Law Reform (Frustrated Contracts) Act 1943, the court asks what was frustrated, the answer can only be the original bilateral contract, because that is the only valid contract between the parties. Therefore, valuations will be based on the original contract, taking no account of the fact that it was necessary to vary the price. This may lead to parties detrimentally relying on the offer of the unilateral variation, only to receive less than the expected remuneration in the event of frustration. To avoid such losses being incurred through detrimental reliance, we ought to reject the unilateral variation contract theory.

E. CONCLUSION ON THE UNILATERAL CONTRACT VARIATION THEORY

The unilateral variation contract theory cannot rescue *Williams*. The single unilateral contract analysis fails to explain *Williams* itself, and both potential analyses fail to explain the other main cases, and are problematic in instances of frustration. We should, therefore, reject it.

IV. DURESS

A. THE THEORY

A second theory sets out to distinguish the cases based on duress. The New Zealand Court of Appeal, in *Antons Trawling v Smith*,³¹ stated:

³¹ [2003] 2 N.Z.L.R. 23 (CA (NZ)).

We are satisfied that *Stilk v Myrick* can no longer be taken to control such cases as *Roffey Bros* ... where there is no element of duress or other policy factor suggesting that an agreement, duly performed, should not attract the legal consequences that each party must reasonably be taken to have expected.³²

The theory is that the absence of variation in *Stilk* can be explained because the variation was vitiated by duress, whereas the existence of variation in *Williams* occurs through the absence of duress. Such thinking has been acknowledged, though not supported, academically.³³

The theory looks to economic duress. *Universe Tankships of Monrovia v International Transport Workers Federation*³⁴ has set out the basic principle, which the subsequent cases have refined.³⁵ It requires an illegitimate threat (normally being satisfied by the unlawful action of threatening to breach a contract),³⁶ which is a 'but for' cause of the 'victim' entering the agreement, where the victim had no practicable alternative but to enter the agreement.³⁷

B. PROBLEMS WITH APPLICATION TO THE CASES

The duress theory cannot explain the result in *Stilk v Myrick*. A finding of duress would not be made in *Stilk* under the modern definition. There was no threat by Stilk, or anyone else. The increased wages arose from Myrick's own suggestion and voluntary undertaking. The other crew do not appear to have threatened to breach their contracts with Myrick. Any pressure arose circumstantially. The 'duress' identified in *Stilk*, presuming we look to Espinasse rather than Campbell (which, of course, is not associated with a duress-based explanation), is really a policy consideration, designed to deter extortion through the *potential* for economic duress being used by crews against captains during voyages, which was created by Lord Kenyon in *Harris v Watson*.³⁸ It is not a true case of duress, as counsel for the sailors pointed out that 'it was made under no coercion, from the apprehension of danger, nor extorted from the captain; but a voluntary offer on his part for extraordinary service.'³⁹ Lord Ellenborough did not seek to dissent from this. The rule applied is described as one of policy. Any references which might invoke duress are phrased as referring to the hypothetical potential for duress, not any actual accusation on the facts. Therefore, while the result in *Stilk* can be attributed to the application of a specialist policy, it cannot be attributed to duress either under the law at the time or which we recognise today.

³² *Antons Trawling* (n 31) 93.

³³ See, e.g., O'Sullivan, 'In Defence of *Foakes v Beer*' [1996] CLJ 219, 227-28, or Coote, 'Consideration and variations, a different solution' [2004] LQR 19, 21.

³⁴ [1983] 1 AC 366 (HL).

³⁵ See Lord Diplock at 383-384, and also the similar explanation in the dissenting judgment of Lord Scarman at 400.

³⁶ A threat to breach a contract is *generally* illegitimate (see *Kolmar Group v Traxpo Enterprises* [2010] EWHC 113 (Comm), [2011] 1 All ER (Comm) 46, [92]), as it is a threat to commit unlawful action as required by Lord Diplock in *Universe Tankships*. In addition to this being accepted in *Atlas Express v Kafko* [1989] QB 833 (HC) and *Adam Opel*, this was also found in *B & S Contracts and Design Ltd v Victor Green Publications* [1984] ICR 419 (CA). As Kerr LJ explains (at 428), the primary limiting factor on economic duress was that the 'victim' had no practicable alternative to enter the contract, not any narrowing of the illegitimacy requirement.

³⁷ *Huyton v Cremer* [1999] 1 Lloyd's Rep 620 (HC).

³⁸ (1791) Peake 102.

³⁹ *Stilk* (n 2), 318-19.

It is conceded that a duress theory explains the result in *Williams*. Regardless of attempting to draw a line to decide whether *Williams*' actions constituted legitimate 'commercial pressure' or an illegitimate threat, it was Roffey, not *Williams*, who suggested the increase in pay. Hence, Roffey voluntarily accepted the rise, rather than being threatened into it. Thus, *Williams* mirrors *Stilk* in that it was the offerens of the variation who was also the potential victim of the 'duress'.

Nonetheless, we are still left with an unworkable mismatch between theory and practice. The duress theory says that both *Stilk* and *Williams* ought to be valid. It is only by resort to referencing policy considerations, which were raised in a report which is not even accepted as accurately describing the basis of the decision, that we get an explanation for the result in *Stilk*. Therefore, the duress theory in and of itself fails to explain the cases.

C. PROBLEMS WITH OFFERENS BEING DETERMINATIVE

The duress theory would create arbitrary distinctions based on who the offerens of the variation is. Accepting what is said about *Williams* above, the variation was valid because it was suggested by the 'victim' of the alleged duress, not the 'oppressor'. This contrasts to the pre-*Williams* case of *Atlas Express v Kafco*⁴⁰ and the post-*Williams* case of *Adam Opel v Mitras Automotive*.⁴¹ In both these cases, a higher payment than originally agreed was demanded against a party who was subject to circumstantial pressure to ensure that the contract was maintained, and who thus agreed to increase their payments under the contract in a single-sided variation. Under these conditions, economic duress was found in both cases.⁴²

What is important to note here is that, in both cases, the 'oppressor' had demanded the variation from the 'victim' of the duress. This contrasts to *Williams*, where the 'victim' of the duress was offerens of the variation. Thus, findings of duress are heavily influenced by which party is offerens. If the 'oppressor' requires a one-sided variation in his favour, the contract should be voidable for duress, as in *Atlas Express* and *Adam Opel*. However, if the 'victim' suggests the single-sided variation, the variation is valid, as in *Williams*. Ordinarily, such a distinction would seem valid, because, one might legitimately conclude, duress is designed to protect the 'victim' from being forced to accept terms which he would otherwise not agree to. But to see the special issue in these cases, one must consider the specific fact pattern. Take a basic example. Adam is performing some work on contract for Bob, which Bob is under pressure to see completed. Adam comes to Bob and says, "I am running out of money, I will not be able to finish the work." Now, here is the distinction. In Situation 1, Bob replies, "okay, how about I pay you an additional £X for it." Alternatively, in Situation 2, Bob replies, "okay, how can we resolve this?" Adam responds, "I could finish it for an extra £X." Bob says, "alright, I will pay you £X more as you suggest." Situation 1 is a *Williams* situation: the offerens is the 'victim'. Situation 2 matches *Atlas Express* and *Adam Opel*: the offerens is the 'oppressor'. Subject to finding Adam's threat to breach the contract to be illegitimate, duress would be found in Situation 2, but not Situation 1, based on the pattern of the conversation. If anything, Bob, by being more proactive in Situation 1, has denied himself the chance to have the variation made voidable by duress. Such a distinction, therefore, appears arbitrary. Hence, such arbitrariness means duress is a poor way to regulate single-sided variations.

⁴⁰ [1989] QB 833 (HC).

⁴¹ *Adam Opel* (n 13)

⁴² Unlike Tucker J in *Atlas Express* (1989), David Donaldson QC in *Adam Opel* (2007) found himself bound by *Williams* (1990) on the issue of consideration: although Mitras had given nothing more than they were entitled to give, Opel obtained a 'practical benefit' of actually being able to maintain production. Therefore, if it were not for the duress issue deciding the case, the results in *Atlas Express* and *Adam Opel* would have been different, despite their materially similar facts, because of *Williams*.

D. LACK OF CONSIDERATION

Critically, the duress theory is invalid in law. The logic is fallacious. It rests on two premises: (1) duress makes a contract voidable and (2) there was no duress. From these, it reaches the conclusion that the contract was therefore valid. This is erroneous reasoning. Just because an agreement is not defeasible does not necessarily mean that it is a valid contract. Even in the absence of duress, the contract may have failed to validly form, and hence be invalid. That, indeed, is the present issue. The duress theory fails to provide any explanation about the issue of consideration in *Williams*, or any other single-sided variation. It holds that, as long as there is no duress, the variation is valid. Legally, that is necessary but insufficient reasoning. In addition to the absence of duress, there must be consideration (amongst other elements). The duress theory, therefore, simply does not work as a matter of law.

E. CONCLUSION ON THE DURESS THEORY

We must, therefore, reject the duress theory too. It is unworkable in law, struggles by itself to explain the cases which it claims to reconcile, and tends to an undesirable distinction.

V. PROMISSORY ESTOPPEL

A. THE THEORY

Some theorists⁴³ have allowed equity to intervene where the common law has apparently fallen down. The equitable doctrine of promissory estoppel has co-existed alongside the common law of contract formation in rendering a select category⁴⁴ of modifications which are not backed by consideration enforceable. Thus, it seems only natural that attempts have been made to explain an issue centred on an apparent absence of consideration by invoking promissory estoppel.

Promissory estoppel operates as follows. Where there has been a clear and unequivocal promise, express or implied,⁴⁵ of the promisor's intention to vary his strict legal rights against the promisee,⁴⁶ and the promisee has acted in (detrimental) reliance upon this,⁴⁷ and where it would be inequitable for the promisor to resile from his promise,⁴⁸ the court may invoke promissory estoppel to suspend⁴⁹ the promisor's strict legal rights under the contract in favour of the variation until such time as the inequity from resiling has ceased, so long as the modification consisted of the variation of existing rights rather than the creations of new ones.⁵⁰ This is said subject to the Court of Appeal's decision in *Collier v Wright Holdings*,⁵¹ which has denied the need for detriment when assessing reliance,⁵² and suggests promissory estoppel can be extinctive rather than suspensory.

⁴³ This theory has been put forward by a number of academics. It is Chen-Wishart's exposition, from 'A Bird in the Hand: Consideration and Promissory Estoppel' in A Burrows and E Peel (eds), *Contract Formation and Parties*, (OUP 2010) 89-113, at 102-105 which is being used here, as there are material differences between the different expositions

⁴⁴ See *Combe v Combe* [1951] 2 KB 215 (CA).

⁴⁵ E.g. *Hughes v Metropolitan Railway* (1877) 2 App Cas 439 (HL).

⁴⁶ *Woodhouse AC Ltd v Nigerian Produce Lrd* [1972] AC 741 (HL).

⁴⁷ See, e.g., *Hughes* (n 45); *Societe Italo-Belge v Palm Oils, (The Post Chaser)* [1982] 1 All ER 19 (HC); *Collier v Wright Holdings* [2007] EWCA Civ 1329, [2008] 1 WLR 643 (CA).

⁴⁸ See, e.g. *The Post Chaser* (n 47); *D&C Builders v Rees* [1965] 3 All ER 837 (CA).

⁴⁹ *Tool Metal v Tungsten Electric* [1955] 1 WLR 761 (HL); cf *Ajayi v RT Briscoe (Nigeria)* [1964] 1 WLR 1326 (PC).

⁵⁰ *Combe* (n 44), *Baird Textile Holdings Ltd v Marks and Spencer Plc* [2002] 1 All ER (Comm) 737 (CA).

⁵¹ *Collier* (n 47).

⁵² Though insistence on detrimental reliance had never been too strict. Detriment is hard to locate, for instance, in the foundational case of *Central London Property v High Trees House* [1947] KB 130 (HC), and Lord Hodson in *Ajayi v Briscoe* (n 48) tells us that reliance is sufficiently satisfied by a change of position

B. PROBLEM WITH APPLICATION TO THE OTHER CASES

Bearing the operation of promissory estoppel in mind, we must apply it to the cases. It can explain the result in *Re Selectmove*. The Court of Appeal, indeed, briefly considered and declined an estoppel argument. There was a clear promise to modify existing rights. As to reliance, Selectmove made some payments, albeit not full. Making some payments is, nonetheless, a change of position, so that should be sufficient to satisfy the reliance element. However, the court rejected estoppel for two reasons. First, as a matter of offer and acceptance, the tax inspector had no authority to make promises on behalf of the Revenue. Second, the court felt that Selectmove's failure to make full payments in line with the new promise ousted any issue of inequity. Therefore, Selectmove could not seek to rely on estoppel, and thus estoppel could not support the variation, consistent with the final decision.

However, promissory estoppel circumvents the rule from *Foakes v Beer*. It was not argued in *Foakes*, despite *Hughes v Metropolitan Railway*⁵³ having arisen less than a decade before, and the Earl of Selborne LC and Lord Blackburn having presided over both cases. *Foakes* agreed to pay only part of what he owed. *Beer* accepted that voluntarily. *Foakes* paid in reliance on the promise, which is sufficient to constitute reliance under the understanding adopted in *Collier v Wright*. Therefore, *Beer's* right to the interest under the original contract should have been suspended. Thus, the promissory estoppel theory, rather than explaining this foundational case, overhauls it. It might be objected that estoppel would not be granted on the facts of *Foakes* because there was no inequity in resiling. However, as noted, change of position has normally been sufficient to satisfy the inequity requirement. Therefore, estoppel would still seem to reverse the finding in *Foakes*.

The problem becomes greater when applying promissory estoppel to the 'more for the same' cases. *Combe v Combe*⁵⁴ (confirmed in *Baird Textile Holdings v Marks and Spencer*)⁵⁵ tells us that promissory estoppel can only be invoked in variation (i.e. reduction) of existing rights, and not creation of new ones. This is because '[t]he doctrine of consideration is too firmly fixed to be overthrown by a side-wind.'⁵⁶ The sentiment is that estoppel may not be used in place of consideration. This is not strictly the case, because it can be used (as in *Collier v Wright*) to reduce a promisor's existing rights without the promisee giving anything in return. This is the idea of promissory estoppel as a shield: it can be used to defend a promisee from the promisor seeking to assert his strict legal entitlement to more rights. However, the distinction still remains that promissory estoppel may not be invoked where the promisor has promised new rights to the promisee. This is the idea of promissory estoppel not being used as a sword: it cannot be used by the promisee to enforce more rights on the promisor than his strict legal entitlement. Thus, promissory estoppel cannot apply to the 'more for the same' cases, which involve the creation of additional rights.⁵⁷ In *Stilk*, the promisee, *Stilk*, gained a new right to the extra pay. In *Williams*, the promisee, *Williams*, gained a new right to the extra pay. In both instances, therefore, the rule from *Combe* prevents estoppel being invoked, and therefore the original contract would have to have stood. While this explains the result in *Stilk*, it conflicts with the result in *Williams*.

On reflection, therefore, the promissory estoppel explanation is unsatisfactory. It only succeeds in applying to *Stilk* and *Re Selectmove*. That result is all the more surprising when it is recalled that, of the four principal cases, *Stilk* alone predates the Judicature Acts and was argued before the common law courts, so its 'explanation' by a non-applicable body of law is questionable. Therefore, promissory estoppel is not a satisfactory explanation for the law in single-sided contract variations by reason, once again, of conflict with significant authorities.

⁵³ *Hughes* (n 45).

⁵⁴ *Combe* (n 44).

⁵⁵ *Baird* (n 50).

⁵⁶ *Combe* (n 44), 220.

⁵⁷ *Syros Shipping Co SA v Elaghill Trading Co, The Proodos C* [1981] 3 All ER 189 (HC).

VI. A NEW THEORY: IMPLIED VARIATION TERMS

The negative argument is now complete. We have considered the Unilateral Variation Contracts, Duress and Promissory Estoppel theories and we have examined the reasons why they are unsatisfactory. Now commences this article's second purpose, to establish a positive argument which offers a solution to the problem of single-sided variations.

A. MY ARGUMENT

The rules of contractual interpretation, specifically the law of the implication of terms, offer a solution. The argument may be outlined in these terms. The court may imply a term from the existing contract which allows for necessary variations of the performance, normally the remuneration. To work within the law governing implied terms, the implication of such a term will be subject to narrow limits, and guiding principles are suggested to ensure these limits adhered to, which will be expanded upon below.

This will solve the apparent conflict with the doctrine of consideration which made single-sided variations problematic in the first place. Since the variation will be construed as nothing more than an option implied into the original contract, the original consideration is all that is necessary. This approach also reconciles the present cases so as to explain the existing pattern of validity or invalidity. This will be demonstrated below.

B. THE LAW OF IMPLICATION OF TERMS

Implication of terms is a rule of construction. Prima facie, this claim may seem odd, because interpretation would seem to imply that there are already some express terms to interpret. Indeed, certain judges, such as Sir Thomas Bingham MR,⁵⁸ have envisaged a strict separation between the two. However, when we recall that what we are construing is the contract *as a whole*,⁵⁹ it should become clear that construction covers both the express and the implied fields, because contracts can include both express and implied terms. Thus, the rules for implication of a term are most easily regarded as a sub-doctrine of construction, specially adapted for dealing with the constructive process of reading words and terms into the pre-existing express terms.

Therefore, it is first useful to understand the general power, and limits of power, of the process of construction. It is an error to think that interpretation is limited to a set of available meanings which can be read from the express terms.⁶⁰ Such restraint was famously rejected by the House of Lords in *Investors Compensation Scheme v West Bromwich Building Society*.⁶¹ Lord Hoffmann, in his five point summary of the rules of interpretation, reminds us that interpretation aims to ascertain the meaning 'which the document would convey to a reasonable person having all the background knowledge which would reasonably have been available to the parties in the situation in which they were at the time of the contract'. The reasonable man, looking at all the surrounding facts, may not

⁵⁸ [1995] EMLR 472 (CA), 481. "The courts' usual role in contractual interpretation is, by resolving ambiguities or reconciling apparent inconsistencies, to attribute the true meaning to the language in which the parties themselves have expressed their contract. The implication of contract terms involves a different and altogether more ambitious undertaking: the interpolation of terms to deal with matters for which, *ex hypothesi*, the parties themselves have made no provision."

⁵⁹ Hooley, 'Implied terms after *Belize Telecom*' (2014) LQR 315, 334

⁶⁰ A mistake made by Leggatt LJ, amongst many examples, in the Court of Appeal in *Investors Compensation Scheme v West Bromwich Building Society* [1998] 1 B.C.L.C. 521 (CA).

⁶¹ [1998] 1 WLR 896 (HL).

only interpret ambiguity in the express terms, but might also conclude that there is an error in the words or syntax.⁶² As his Lordship subsequently said in *Chartbook v Persimmon Homes*, ‘there is not, so to speak, a limit to the amount of red ink or verbal rearrangement or correction which the court is allowed’.⁶³ Thus, the general constructive processes are not limited to understanding the express terms of the agreement, but extend to permit courts to read words and terms in and out. There is, therefore, a power to correct erroneous additions or omissions.

Implication of terms must be isolated from the general law of interpretation. If we were to impose a rough taxonomy⁶⁴ on the law of interpretation, the primary division would be between interpretation of express terms and implication.⁶⁵ The purpose of drawing this taxonomy is to understand that the power of the process of construction may well vary between the two primary branches. In the express branch, the courts may understandably be more conservative, because they are working within the confines of express words which they must accord sufficient respect to. The argument presented here looks, however, to the second category, implication, specifically implication in fact.

Implication in fact is subject to prima facie strict, if ill-defined, limits. The general principle for such implications was restated by Lord Hoffmann in *Attorney General of Belize v Belize Telecom*.⁶⁶ Three basic rules emerge from his speech. First, the implied term must be consistent with ‘the meaning that the instrument would convey to a reasonable person having all the background knowledge which would reasonably be available to the audience to whom the instrument was addressed’.⁶⁷ This reminds us, therefore, that we are undertaking a process of construction. The heart of the process is in the second rule. The proposed term must be necessary⁶⁸ ‘to spell out what the contract actually means’.⁶⁹ This has, in the past, been stated in various ways, such as a business efficacy test⁷⁰ (that the term must be necessary to give business efficacy to the transaction) and a officious bystander test⁷¹ (that, had the officious bystander asked the parties whether they meant to include the term, they would have quickly replied in the affirmative). The third rule is a slight repetition on the first, in reminding us that the proposed term must be consistent with the express terms of the contract when subject to a contextual interpretation in line with *Investors Compensation Scheme*.⁷²

From this, therefore, we must define the power and the limits of implication by fact. The primary limit is the second rule: any implied term must be necessary to, in effect, make the contract work. It is worth pausing for a moment to consider the meaning of ‘necessary’. Traditionally, it has not been used in this context to mean ‘essential’. In *The Moorcock*,⁷³ it was not essential that the wharf-owners should guarantee that they had taken reasonable care to check the safety of the mooring. The basic function of the transaction, to give a ship a place to dock, could still occur without such a promise. All the implication did was to increase the business sense of the transaction, by adding a term which would seem sensible to reasonable people on both sides, because it provided a reasonable degree

⁶² *ICS* (n 61) 912-913.

⁶³ [2009] UKHL 38, [2009] 1 AC 1101 [25].

⁶⁴ This division has been adapted from Lord Grabiner, ‘The iterative process of contractual interpretation’ (2012) LQR 41.

⁶⁵ It is recognised that there are deeper issues with the express/implied term distinction. However, space precludes an analysis, which is in any case fairly unnecessary, because this taxonomy is only a superficial means of introducing powers of interpretation, and is not relevant to the argument later put forward

⁶⁶ *Belize* (n 7)

⁶⁷ *ibid* [16].

⁶⁸ Confirmed in *Mediterranean Salvage & Towage Ltd v Seamar Trading & Commerce Inc* [2009] EWCA Civ 531, [2010] 1 All E.R. (Comm) 1.

⁶⁹ *ibid* [27].

⁷⁰ *The Moorcock* (1889) 14 PD 64 (CA); *Reigate v Union Manufacturing Co (Ramsbottom) Ltd* [1918] 1 KB 592 (CA).

⁷¹ *Shirlaw v Southern Foundries (1926) Ltd* [1939] 2 KB 206 (CA).

⁷² *ICS* (n 60).

⁷³ *The Moorcock* (n 70).

of confidence to ship-owners, who would have no practical means of inspecting the mooring themselves, and avoided imposing too great a burden on the wharf-owners, because it was only an obligation to take reasonable care to inspect the mooring. Take also *Liverpool City Council v Irwin*.⁷⁴ Their Lordships rejected Lord Denning MR's approach from the Court of Appeal that the term only needs to be reasonable. However, they reach the same conclusion on the facts. Moreover, as Atiyah points out:

It is not *necessary* to have lifts in blocks of flats 10 stories high (indeed high-rise buildings existed long before lifts were invented), though it would no doubt be exceedingly inconvenient not to have them. So "necessary" really seems to mean "reasonably necessary", and that must mean, "reasonably necessary having regard to the context and the price". So in the end there does not seem to be much difference between what is necessary and what is reasonable.⁷⁵

Thus, the term implied in *Liverpool City Council v Irwin*, while 'reasonable' (spelling out what the contract might reasonably have meant), it was not 'essential' (it did not spell out what the contract *must* have meant). *Belize* itself professes this same use of 'necessary'. It was not essential that, if no-one had the power to remove the two directors, a term must be implied to allow their removal. Article 112 stipulated workable terms of office for them. The only issue arose if it became desirable to remove the directors for a reason other than those stipulated in Article 112. Thus, while the implied term was 'reasonable', it was not 'essential'.⁷⁶ Therefore, when considering necessity, we should remember that its real meaning imports a lower standard than a literal interpretation might suggest.

Beyond demonstrating 'necessity', there is the interpretive caveat imposed by Lord Hoffmann in *Belize*:

'The court has no power to improve upon the instrument which it is called upon to construe, whether it be a contract, a statute or articles of association. It cannot introduce terms to make it fairer or more reasonable. It is concerned only to discover what the instrument means.'⁷⁷

Discovering what the instrument means does not, however, limit the court to the literal or express wording of the contractual document. In *Belize* itself, the contract said that directors held office "subject only to Article 112". However, given the outcome, "subject only" cannot be understood literally, because the court determined that they also held office subject to the term which they implied.⁷⁸ His Lordship must have had his *Chartbrook* red pen to hand. Furthermore, we cannot take Lord Hoffmann's own statement literally. Implying a term, even in the simplest case, is *necessarily* an improvement on the instrument, because it corrects some fault within it. Thus, the implication necessarily makes the instrument fairer and/or more reasonable.⁷⁹ Because the instrument is silent on the matter, it cannot be a simple process of 'discovery', as there is no content to discover. The court attributes the term to the contract. Thus, 'necessary interpretation' is understood to be somewhat wider than the words 'necessary interpretation' themselves suggest.

Two more rules must be remembered. First, there is the default position of no implied term.⁸⁰ Second, when implying a term, the courts are not limited to subjective intentions. In accordance with

⁷⁴ [1977] AC 239 (HL).

⁷⁵ Atiyah, *An Introduction to the Law of Contract* (Clarendon Press, 5th edn, 1995) 207

⁷⁶ For a potentially even wider meaning, see *SNCB Holding v UBS AG* [2012] EWHC 2044 (Comm), especially [65].

⁷⁷ *Belize* (n 7) [16].

⁷⁸ Law and Loi 'The Many "Tests" for Terms Implied in Fact: Welcome Clarity' (2009) 125 LQR 561, 564.

⁷⁹ Davies, 'Recent Developments in the Law of Implied Terms' [2010] LMCLQ 140, 144.

⁸⁰ *Belize* (n 7) [17].

the general rules of interpretation, the law of implication looks to the objective meaning of the instrument,⁸¹ though this meaning need not be ‘immediately apparent’.⁸²

Therefore, the power and limits of implication by fact may be stated thus. It will be rare for a proposed term to be legitimately implied. The default position is that no term will be added. To overcome this, the term must be necessary (viz. ‘reasonable’) for the contract to ‘work’ and result from a reasonable (though not necessarily literal or even entirely faithful) interpretation of the contract and the background facts. We do not have to be constrained by the parties’ reading of the contract and understanding of the background facts (though, of course, their opinions may be a very helpful contribution to the interpretive process). The issue which the proposed term seeks to resolve need not be expressly mentioned by the contract, or even alluded to.⁸³ It can be entirely absent. Nor need it be obvious from either the contract or the background facts. Implication in fact is not, therefore, an impotent doctrine.

C. THE PROPOSED TERM

In light of this examination of the law, the following term is proposed. It is phrased as if it were implied in *Williams v Roffey Bros*. ‘The remuneration offered to Williams under the contract will at least be sufficient to facilitate performance by ensuring that it is economically viable. The initial price recorded in this document is not conclusive, but may be subject, if necessary, to increase in accordance with an independent valuation of the performance cost’. This need not be the exact wording; this example is only meant to establish the approximate nature of the term.

How would this work in practice? The default position is that no such term will be implied, and only rarely will it succeed. The price variation must be necessary to make the transaction work, given the purpose of the transaction as understood from the contractual document and the background facts. To emphasise this, the implied term is subject to a ‘necessity’ caveat. When looking for necessity in our context, it will be highly relevant to consider whether performance has become actually financially unfeasible. This is because, if the contract is financially unfeasible, it cannot ‘work’, as performance will be unobtainable. Therefore, implying the term will make the contract ‘work’. As a matter of construction, it must also be reasonable to imply the term. Two principles are relevant here. The first flows from the ‘necessity’ criteria: the original, ‘ball-park’ price must be less than the performance cost in ‘more for the same’ scenarios or greater than the value of the consideration given by the promisor in ‘less for the same’ scenarios (as valued by an impartial, objective source – this provides a mechanism for making any performance measure objectively ascertainable, so the term is not too vague). This principle is needed to control the manner in which the work becomes financially unfeasible; it would only be reasonable to imply a term to remedy the work being financially unfeasible if the unfeasibility were attributable to the terms of the initial agreement, rather than any external factor. The second principle limits the frequency of implication, and explains why it is reasonable to imply this term: obtaining the primary performance, over and above any secondary remedies or recourse to the courts, must have been so important to the promisor that the parties must have meant to permit variations of the remuneration agreed under the contract in the conditions of unfeasibility described above. The term could not, of course, be reasonably implied if the risk of financial hardship is expressly allocated to one party, or variations are expressly prohibited. However, if the contract does no more than provide a figure for the remuneration, there may yet be room for reasonable implication

⁸¹ *ibid* [16].

⁸² *ibid* [25].

⁸³ *ibid* [18].

of the proposed term. In addition to these principles, the other rules stated by Lord Hoffmann about implied terms apply.

As a preliminary point, this Implied Variation Terms approach addresses a couple of the issues which afflicted the theories discussed in Sections II-V. It explains the damages award. The reasonable increase is objectively verifiable by reference to an independent third party (and therefore is not too vague to be enforceable).⁸⁴ Such a term, therefore, would also explain why the damages awarded to Williams were approximately equal to the estimation by the surveyor, not the total amount Williams requested in the variation. It also addresses the frustration issue. Since the implication of the term is objectively verifiable by the court, and would arise ab initio or never at all, the court can always determine whether or not the term must be borne in mind when considering cases of frustration.

D. APPLICATION TO THE CASES

The strength of this term must now be assessed on the facts of *Williams*. Was such a term necessary to make the contract work? As said above, it would be highly relevant to bear in mind financial unfeasibility, which is the case in *Williams*. Quite simply, without the variation, Williams was not going to be able to perform the carpentry, and therefore the extra remuneration through the variation was needed to make the contract work. Is it reasonable to imply a term? On its face, the document was a simple commercial agreement. Therefore, there was nothing in it to absolutely preclude implication, but nothing to strongly support it (although it may be noted that, at first instance, another term was implied into the contract, suggesting that it was not complete). The background facts are, however, crucial. Recall the two principles relevant to reasonable interpretation suggested above. First, in *Williams*, the price agreed was less than the independent surveyor's approximation of the performance cost. Thus, the financial unfeasibility, and hence the 'necessity' issue, was attributable to the terms of the agreement. Second, because Roffey was subject to penalties under the main contract, he had strong reason to desire primary performance. Therefore, it is reasonable to infer that Roffey would have meant to allow a variation in the remuneration if it turned out that performance was financially unfeasible for Williams, in order that he was able to secure Williams' continued performance. We can also factor into the process of the interpretation the fact that both parties actually consented to the variation when they turned their minds to it, and so it is more likely to be reasonable. Thus, *Williams* appears to be a (rare) case where the implication of an Implied Variation Term would succeed.

Williams having been explained, the theory needs to be applied to the three cases where no variation was found. Such a term as in *Williams* could not be implied in *Stilk v Myrick* because it was not necessary to make Stilk's performance possible. The variation was financial. Stilk's performance was not, however, financially unfeasible without the variation: he was just as able to sail the ship for his normal remuneration as for the increase. Therefore, the implied term was not necessary, and so could not be implied.

The explanation of *Foakes v Beer* and *Re Selectmove* is even simpler. The original agreements were not contracts, but were imposed by the general law – in *Foakes* as the result of the court judgment, and in *Selectmove* by statute. Therefore, there was no contract to imply a term into. While it might be pointed out that statutory construction involves the same interpretive process, this is a formal observation, and ignores the substantive difference made by the source of the obligations when undertaking the interpretive process. The factual background of a contract is filled by the interactions and desires of the specific parties concerned in the case. The background of a statute is different. The courts look to, inter alia, Parliament's intent. They do not, however, concern themselves with specific

⁸⁴ *May and Butcher v R* [1934] 2 KB 17 (HL).

cases when interpreting a general statute. Therefore, the factual matrix is different between the different sources, and hence, without the relevant background, establishing a claim for Implied Variation Terms would be implausible. Hence, quite simple explanations can be given to reconcile *Stilk*, *Foakes* and *Selectmove*, in addition to *Williams*, with an implied terms theory.

VII. POTENTIAL OBJECTIONS

It is now pertinent to dismiss a few objections which might arise against Implied Variation Terms.

A. THE THEORY ENCOURAGES RECKLESS UNDERCUTTING

A brief glance at the Implied Variation Terms theory might suggest that it would encourage parties to bid purposefully low, far below the performance cost, in order to secure a contract. They would then seek an implied term at a later date to cover their costs, which would be commercially disruptive.

Such concerns need not arise. Through the rules for the implication of terms, looking to both (a malleable) ‘necessity’ and ‘reasonableness’ requirement, the courts have a measure of discretion. Therefore, they may refuse to imply a term, or at least alter the nature of the term, if they concluded on the evidence that the bids were purposefully low, because it may be ‘unreasonable’ to allow the variation in these circumstances.

B. THE THEORY RELIEVES PARTIES FROM BAD BARGAINS

It might be objected that the theory relieves parties from bad bargains. *Williams*, by asking for insufficient remuneration, entered a bad bargain, but was reprieved by being granted the variation. In *Pink Floyd Music Ltd v EMI Records Ltd*,⁸⁵ Lord Neuberger MR cautioned that the fact that the literal meaning produces a ‘bad bargain’ for one of the parties should not, in and of itself, suggest something has gone wrong and therefore that interpretation can be used to ‘fix’ the contract.⁸⁶ Thus, implied terms cannot be used to grant such relief.

This criticism does not stand for two reasons. First, at no stage have I claimed that a bad bargain would be sufficient to trigger an implied variation term. I have laboured the point that implication will be rare and subject to narrow constraints, and I have provided guiding principles to this effect. In any case, it is worth considering whom the contract is a bad bargain for. Traditionally, ‘bad bargain’ is assessed narrowly, on purely financial grounds: a party enters a bad bargain if he stands to make an economic loss by it. However, in the fact situation where an implied variation term could arise, the detriment would not only be to the party who stands to make an economic loss, for the term will only be found if the other party had good reason to desire primary performance. Thus he, too, stood to suffer a detriment through the contract should the primary performance not be given, so it would have been a bad bargain for him as well. Therefore, even if relief is given for a ‘bad bargain’, this is not the sort contemplated by Lord Neuberger. It is a bad bargain for both parties, so relief may well be desirable.

⁸⁵ [2010] EWCA Civ 1429.

⁸⁶ *ibid* [20].

C. THE THEORY CAUSES COMMERCIAL UNCERTAINTY

Implying terms to vary the remuneration in commercial contracts would cause commercial uncertainty, says this argument, as the price would vary from that appearing on the face of the document. How could commercial parties, therefore, have any certainty that the express wording of their agreement will be adhered to?

In reply to this, one must remember that the confines for implying terms are narrow. Since they must be necessary to make the contract 'work', it follows that, if they were not implied, the contract would not have been performed anyway. They are, therefore, limited to situations in which commercial certainty would otherwise have been violated by failure of performance. Beyond these situations, the rules of construction would prevent a term being implied, and therefore parties should not be concerned that the remuneration would be unpredictably altered.

CONCLUSION

Two arguments have now been made. In my negative argument, I have sought to demonstrate that three analyses applied to single-sided variations are unsupportable. Having seen why a simple distinction drawn between the present cases could not amount to a satisfactory solution, we first examined the Unilateral Variation Contracts theory. It had to be envisaged in the form of multiple unilateral contracts to work on the facts of *Williams*, could not explain the results in *Stilk*, *Foakes* or *Selectmove* and caused difficulties when a frustrating event intervenes. The Duress theory could not even explain the cases – *Williams* and *Stilk* – which it was proposed to reconcile, and also caused the issue of arbitrariness on the basis of offerens, and was anyway invalid in law for want of consideration. The Promissory Estoppel theory fared no better, failing to explain *Foakes* and *Williams*, and being entirely inapplicable to a 'more for the same' scenario.

My positive argument then sought to establish an explanation for single-sided variations based on Implied Variation Terms. Subject to strict limits, and bearing in mind certain guiding principles, a term can be implied, when necessary, to permit a variation. Given the narrow requirements, this will be a rare occurrence, and thus such a term could not be implied in *Stilk*. *Foakes* and *Selectmove* were not in any case susceptible to an implied term. On the other hand, *Williams* was a rare example of the correct conditions manifesting, and hence a term could be implied. The Implied Variation Terms theory can, therefore, offer a suitable method of analysing single-sided contract variations without demanding fresh consideration, keeping them within the bounds permitted by our law as it presently stands.

Broad, Inflexible and Redundant?: Fixing The Anti-Avoidance Rule in Section 75A Finance Act 2003

Vincent Ooi¹

I. INTRODUCTION

Sections 75A- 75C of the Finance Act 2003 ('the Sections') were enacted with the intention of countering schemes that have the effect of reducing Stamp Duty Land Tax ('SDLT') liability.²

These sections were subjected to criticism right from the start, with practitioners noting its exceptionally broad scope and some going so far as to call it 'fundamentally deficient' and 'almost unworkable in practice'.³ The recent decisions of the First-Tier Tribunal in *Project Blue Ltd v Revenue and Customs Commissioners*⁴ ('*Project Blue FTT*') and its subsequent appeal to the Upper Tribunal in *Project Blue Ltd v Revenue and Customs Commissioners*⁵ ('*Project Blue UT*') have since provided some clarification about the interpretation of the scope of the Sections.

This article will begin with a brief analysis of the Sections and the decisions at both courts in *Project Blue*, with a focus on the inherently broad scope of the Sections created by the language of the statute. It will then go on show that the initial framework that the Sections was intended to create relied heavily on the power of discretion by Her Majesty's Revenue and Customs ('HMRC') in deciding which cases the Sections should be applied to. As the courts in *Project Blue* have held that HMRC has no such discretion, it will be argued that the decision strains the other mechanisms which restrict the scope of the Sections. The law as it stands after *Project Blue* now seems indeed almost unworkable in practice.

The article will focus on the three main flaws of the Sections, as elucidated from the interpretation of the court in *Project Blue*: 1) the lack of discretion for HMRC to determine the cases to apply the Section to; 2) their excessively broad scope; and 3) the lack of flexibility in determining the consideration paid for the land. The article will then concentrate on four potential solutions to the problem of the excessively broad scope of the Sections. It will be argued that the scope of the Sections could be narrowed by: 1) a greater exercise of the powers of the Treasury to widen the scope of exceptions;⁶ 2) restoring the power of discretion to HMRC on when to apply the Sections (preferably by amending the statute based on the current General Anti-Avoidance Rule ('GAAR'));⁷ 3) refining the requisite connection test⁸ and basing it on whether the transaction is in substance one of the disposal and acquisition of land; or 4) accepting motive as a regulatory mechanism.

¹ Trinity College, Oxford. I am very grateful to Dr Glen Loutzenhiser for all his insightful comments. I would also like to thank the editors of the Oxford University Undergraduate Law Journal for reviewing the article. All errors and omissions are my own.

² HM Revenue and Customs, 'Section 75A Finance Act 2003' <http://webarchive.nationalarchives.gov.uk/20140109143644/http://www.hmrc.gov.uk/so/advice75a.htm> accessed 1 July 2015.

³ Natalie Lee (ed), *Revenue Law: Principles and Practice* (27th edn, Bloomsbury Professional 2009) para 3.1.

⁴ [2013] UKFTT 378 (TC).

⁵ [2014] UKUT 564 (TCC).

⁶ *Project Blue* (UT) (n 5).

⁷ The GAAR is a broadly drafted provision that attempts to counter tax avoidance schemes that have 'slipped through the net' of specific anti-tax avoidance legislation. The current GAAR was enacted by Parliament through the Finance Act 2013, ss 206-215 and sch 43.

⁸ Based on the interpretation of the words 'in connection with' in s 75A(1)(b) Finance Act 2003.

It is accepted that the language of the Sections and the current interpretation adopted in *Project Blue* does restrict the ability of the courts to interpret the Sections freely. However, it will be argued that though legislative changes would definitely be welcome for an Act which ‘drafting leaves much to be desired’,⁹ in the meanwhile, judicial solutions can be applied and the courts need not be forced to come to unjust and impractical outcomes. The four solutions proposed above represent a good mix of statutory and judicial measures.

The need for a mini-GAAR in this context will then be questioned, looking at the present situation where we already have Disclosure of Tax Avoidance Schemes (‘DOTAS’) legislation¹⁰ and a broad GAAR. If the Sections are to be useful, they must be able to perform a function in addition to or better than the existing measures. It will be argued that the key to this lies in reforming the scope of application of the Sections. Of the four solutions proposed, it will be argued that the ‘requisite connection’ test¹¹ is ultimately the best solution. This is because it can provide a more refined approach to determining the scope of the Sections that goes beyond what the GAAR can currently do.

II. THE INTRODUCTION OF GENERAL ANTI-AVOIDANCE RULES

For as long as there has been the collection of tax, there have also been attempts to avoid the payment of tax. Despite the constant attempts of the tax authorities to brand tax avoidance as immoral,¹² the constant intellectual battle between the tax authorities and the taxpayer is unlikely to cease anytime soon. As our tax laws have become increasingly sophisticated, the ingenuity of the taxpayer (or more realistically, his lawyers and accountants) has kept pace, producing more and more elaborate schemes to ‘optimise’ the amount of tax payable. Understandably, this has caused the tax authorities much consternation as they are forced to constantly review their policies, plugging loopholes and ensuring that enough revenue is collected. Tax avoidance is a particularly political issue at the moment and some view curbing it as an opportunity to unlock a potentially large source of much needed revenue.

With taxpayers constantly testing the line between tax avoidance and tax planning (and a few bolder ones crossing over into tax evasion),¹³ the Government has instituted several policies intended to tilt the odds in their favour. Targeted (or Specific) Anti-Avoidance Rules (‘TAAR’) have been introduced to deal with loopholes and schemes in numerous areas of the law. DOTAS legislation has been implemented to provide HMRC with more information on potential tax avoidance schemes.

⁹ *Project Blue* (UT) (n 5) [132] (Nowlan J).

¹⁰ The Finance Act 2004, ss 306-319 introduced a duty on those involved in implementing tax avoidance schemes to disclose details to HMRC. This was reviewed and tightened in FA 2007, FA 2008 and FA 2010. For more details on disclosure, see J. Tiley and G. Loutzenhiser, *Revenue Law* (7th edn, 2012) para 5.3.2 and Freedman [2004] BTR 332, 339-42.

¹¹ Based on a particular interpretation of the words ‘in connection with’ in s 75A(1)(b).

¹² The morality of tax avoidance is a controversial area. See J. Freedman, ‘Interpreting Tax Statutes: Tax Avoidance and the Intention of Parliament’ (2007) 123 LQR 53, where she considers Honoré’s argument from T. Honoré, ‘The Dependence of Morality on Law’ (1993) 13 O.J.L.S. 1. ‘[Honoré] argues that, whilst there is a moral obligation to pay taxes, this obligation is incomplete apart from law because the law has to fix the amount or rate of tax. He might have added that law has to fix the basis on which tax is payable.’

¹³ The terms used here are in common use though there is some ambiguity as to their exact meaning. They are helpfully explained in J. Tiley and G. Loutzenhiser, *Revenue Law* (7th edn, 2012) para 5.1. Tiley and Loutzenhiser distinguish Tax avoidance and tax evasion by explaining that in the former case, no liability to tax arises, while in the latter case, a charge arises but the tax cannot be collected. They define tax planning as ‘what all sensible people do in order to reduce their tax liabilities.’ Also see Freedman (n 12) at 70.

General Anti-Avoidance Rules ('GAAR') represent the latest attempt of a Government determined to regain control of the tax process and force businesses to scale down, if not eliminate, their tax avoidance efforts.¹⁴ GAARs involve a substantial conceptual departure from the measures taken by the Government in the past. Unlike TAARs, which attempt to pre-empt and stop existing or foreseeable schemes of tax avoidance, GAARs attempt to catch tax avoidance schemes which are not caught by TAARs, whether they were foreseeable at the point of the enactment of the GAAR or not.¹⁵ The ability of the GAARs to do this lies in their very broad scope which reduces the ability of the taxpayer to structure a scheme that takes advantage of the limitations of statutory language caused by imperfect information. As GAARs have very broad scopes, there is a need for some other kind of mechanism to regulate this scope. A popular mechanism is to rely on the tax authority's discretion on when to apply the GAAR.¹⁶

While GAARs can be more effective at catching cases of tax avoidance, they are also controversial as, *prima facie*, they raise questions about certainty of legislation, the Rule of Law and the justifiability of the potential retroactive effects.¹⁷ A strong Rule of Law based criticism of GAARs would be that they hamper the ability of individuals to find out what the law is and to structure their affairs accordingly. These serious concerns hampered the smooth legislation and enactment of a GAAR in the United Kingdom. However, after much debate and a few false-starts, Parliament has since enacted a GAAR into law.¹⁸

The history of the GAARs and the related anti-avoidance measures are important for a full understanding of the Sections, which were one of the first successful attempts of the Government in getting a GAAR into law (though on a small scale and not drafted in the best way). However, the Sections differ from the GAAR itself in that it does not seem to have an express provision for a test of 'reasonableness', to be used when determining the scope of application of the anti-avoidance rule.

With this understanding of the place of the Sections in the general scheme of things, two questions are raised: 1) if the Sections are a primitive (and unrefined) attempt at a GAAR, does this mean that we should revise or construe them in accordance with the final (refined) version of the GAAR?; and 2) if the Sections were enacted before the GAAR was passed, do they still serve a useful purpose now that SDLT is one of the taxes to which the GAAR applies?¹⁹ These questions will be answered in detail later in the article. For now, suffice it to say that the answer to the first question is, yes; and the answer to the second question is, yes, but only if we use them correctly.

¹⁴ An interesting point raised by M. Thomas is that the GAAR does not have the effect of eliminating all tax avoidance schemes, but deals with 'standard schemes' and forces bespoke planning of tax avoidance. This has the effect of making it much harder, though not impossible, to avoid tax. See M. Thomas, 'Section 75A 2003: the Death of SDLT Planning?', *GITC Review* (Volume VI Number 2, June 2007) 1.

¹⁵ The UK GAAR was arguably designed only to target the most 'abusive' schemes. See J. Tiley and G. Loutzenhiser, *Revenue Law* (7th edn, 2012) para 5.52 and The Aaronson Report (Report of the GAAR Study Group, November 2011) 5.1.

¹⁶ Noted that this is subject to some oversight. For example, in Canada, there is a GAAR committee made up of Revenue, Finance, and lawyers from the Department of Justice which makes the final decision as to whether to apply the GAAR. In the UK, there is the UK GAAR Panel oversight.

¹⁷ The concerns regarding GAARs and its potential benefits are addressed in J Freedman, 'Designing a General Anti-Abuse Rule: Striking a Balance' (2014) *IBFD Asia- Pacific Tax Bulletin* 165. Freedman argues that such provisions are essential in a modern tax system. With the increased complexity of such systems, it would be impossible for specific legislation to catch every abuse.

¹⁸ Finance Act 2013, ss 206-215 and sch 43.

¹⁹ This is provided for in Finance Act 2013, s 206(3)(f).

III. PROBLEMS WITH THE ACT AND ITS APPLICATION

A. SECTIONS 75A-75C OF THE FINANCE ACT 2003²⁰

Sub-sections 75A(1) and (2)

(1) This section applies where–

- (a) one person (V) disposes of a chargeable interest and another person (P) acquires either it or a chargeable interest deriving from it,
- (b) a number of transactions (including the disposal and acquisition) are involved in connection with the disposal and acquisition ('the scheme transactions'), and
- (c) the sum of the amounts of stamp duty land tax payable in respect of the scheme transactions is less than the amount that would be payable on a notional land transaction effecting the acquisition of V's chargeable interest by P on its disposal by V.

(2) In subsection (1) 'transaction' includes, in particular–

- (a) a non-land transaction

Section 75A(1) lays out the scope and application of the anti-avoidance rule. There are three conditions which must be satisfied if the Sections are to be applied: 1) a disposal of land (or an interest deriving from it) by a person (V) and the acquisition of land by another person (P);²¹ 2) a number of transactions involved in connection with the disposal and acquisitions ('the scheme transactions'); and 3) a lower amount of SDLT is payable than if there was a notional sale from V to P. Sub-section (2) clarifies that 'transaction' includes, *inter alia*, non-land transactions.

The resulting effect is an extremely broad scope for the Sections. For any transaction where land is involved; where there are a number of connected transactions; and where less SDLT is payable than with a direct transfer; SDLT will be payable in full as if there was a direct transfer. This has the potential to cover almost any transaction with the slightest link to land which is conveyed through intermediary companies. The question then becomes whether it was possible that the Sections were indeed intended to catch all of these transactions, regardless of whether an intention to avoid tax was present.

²⁰ The Sections were first introduced by the Stamp Duty Land Tax (Variation of the Finance Act 2003) Regulations 2006 (SI No. 3237) and a new version was enacted by s 71 Finance Act 2007. The Sections were thus drafted as subsidiary legislation and later enacted as primary legislation.

²¹ It is noted that there was considerable discussion on identifying P and V in *Project Blue* at both the First-Tier Tribunals ('FTT') and Upper Tribunal ('UT') levels. The lack of clarity on this issue is certainly a problem in interpreting the Section. However, it will not be covered in detail here. For a more in-depth discussion on this point, see Patrick Cannon, 'Project Blue' (News Article, 20 January 2015)

<<http://www.patrickcannon.net/news/currentnews/article.cfm?id=81>> accessed 1 July 2015.

Sub-sections 75A (4) and (5)

(4) Where this section applies–

(a) any of the scheme transactions which is a land transaction shall be disregarded for the purposes of this Part, but

(b) there shall be a notional land transaction for the purposes of this Part effecting the acquisition of V's chargeable interest by P on its disposal by V.

(5) The chargeable consideration on the notional transaction mentioned in subsections (1)(c) and (4)(b) is the largest amount (or aggregate amount)–

(a) given by or on behalf of any one person by way of consideration for the scheme transactions, or

(b) received by or on behalf of V by way of consideration for the scheme transactions.

Sub-section (4) then provides for the effect of the anti-avoidance rule. If the conditions under sub-section (1) are met, all the intermediate (or scheme) transactions will be disregarded under this head of tax and the full amount of SDLT payable on a direct (notional) transfer from V to P will be payable instead. Sub-section (5) provides that the SDLT shall be computed based on the largest amount of consideration paid at any point of the scheme.

Sub-sections 75C(1)–(10)

Sub-sections 75C(1)–(10) qualify sub-section 75A(5), with sub-section (5) in particular allowing for the 'just and reasonable' apportionment of the consideration paid between the various chargeable interests if there are more than one in the scheme transactions.

Keeping in view sub-section 75A(5), however, It is noted that the Sections provide no power to the courts to make their own estimate of a fair quantum of the consideration paid on which to base the calculation of SDLT. It is simply to be calculated on the highest quantum of consideration paid in the scheme transactions. Sub-section 75C(5) allows for separation of land transactions (for each property) from the rest of the transaction, with SDLT being payable individually on each land transaction. This means that SDLT will not be blindly calculated on the full quantum of the consideration paid, which may include consideration given for non-land transfers. However, there is at least one type of situation where the strict application of the rule might lead to harsh consequences. This is where there are fluctuations in the price of the property as the transactions occur. Thus, SDLT would be calculated on the highest quantum of consideration given at any point of the scheme transactions, even if the market price of the property subsequently drops.

If the intention behind the statute is merely to prevent tax avoidance, then there is no justification for making the parties base the quantum of consideration on the highest possible price of the property across all the scheme transactions. Not all land transactions go through multiple stages for tax avoidance reasons. Counsel for the Appellant in *Project Blue* raised some objections on this point, with the strongest objection being the argument that the subsequent purchaser may be wholly unaware of the consideration paid in the earlier part of the scheme transactions and end up being subject to a

much larger tax bill than initially expected.²² In *Project Blue* itself, the transaction had to be structured in such a way so as to comply with the rules of Islamic Financing.²³ The impression one gets from reading subsection 75A(5) is that it was drafted to prevent the parties from transacting at an undervalue and then arguing that the tax should be calculated based on that lower value. If that was indeed the intention of Parliament (which fits well with the idea of the Sections being a mini-GAAR), then the literal wording of subsection 75A(5) seems to be unduly inflexible. It would arguably be fairer to allow the court the discretion to fix a single point when the transfer should have taken place, and base the calculation of SDLT on the consideration paid at that point.²⁴

B. THE FACTS OF PROJECT BLUE²⁵

For the purposes of the article, the facts of *Project Blue* can be quickly summarised. It will soon be apparent that the case has much broader implications. Most of the points in the judgment which are being considered here were expressed in a general manner and it does not seem that they were particularly dependent on the facts of the case. Project Blue Limited ('PBL') bought land from the Ministry of Defence ('MOD') for £959m but had to structure the transaction such that the land had to be sold to a financial institution ('MAR') at £1.25b and leased back to PBL with options providing for the eventual transfer of the property back to PBL.²⁶ PBL argued that the structuring was done in order to comply with the Sharia (Islamic) law of finance and it was claimed (though not established in court) that there was no intention of tax avoidance.²⁷ HMRC then sought to apply s 75A and claim SDLT based on the full consideration of £1.25b.²⁸ PBL contested that assessment and claimed that under s 45(3) Finance Act 2003 or s 71A Finance Act 2003, sub-sale relief applied and they were not liable to pay any SDLT.²⁹

C. PROJECT BLUE³⁰ AT THE FIRST TIER TRIBUNAL

The First Tier Tribunal interpreted the Sections with a rather literal approach, largely coming to the same conclusions as were listed above in the analysis of the Sections. As noted, this interpretation of the Sections by the courts produces two problems: 1) an excessively broad scope; and 2) a lack of flexibility in determining the consideration paid for the land. Setting aside the issues of the identification of 'P' and 'V' in the transactions and the impact of the Sections on reliefs, which are beyond the scope of this article;³¹ the most important issues which the decision of the First Tier Tribunal dealt with relate to 1) the discretion of HMRC to decide which cases to apply the Sections to; and 2) the interpretation of the terms 'in connection with' in s 75A(1)(b). The court also dealt with the issue of 'motive' and its relevance to the applicability of the Sections. These issues will be considered in detail subsequently.

²² *Project Blue* (FTT) (n 4) [121].

²³ *ibid* [123]. Noted that this was not actually proved in the case itself.

²⁴ There is one more notable concern with the way that the Sections focus on a notional transfer and disregard the other parts of the transaction. This has to do with the applicability of reliefs and the difficulty of determining whether and which reliefs are relevant to the scheme transactions. This point is covered by M. Thomas, 'Section 75A 2003: the Death of SDLT Planning?' *GITC Review* (Volume VI Number 2, June 2007) 72- 75.

²⁵ Referring to the case both at the FTT and the UT levels. *Project Blue* (FTT) (n 4), *Project Blue* (UT) (n 5).

²⁶ *Project Blue* (FTT) (n 4) [2]-[15].

²⁷ *ibid* [123].

²⁸ *ibid* [18]-[20].

²⁹ *ibid* [16].

³⁰ *ibid*.

³¹ For more in-depth coverage of these points, see (n 21) and (n 24).

D. PROJECT BLUE³² AT THE UPPER TRIBUNAL

The focus of the court at the Upper Tribunal level was mainly on the issues of identifying ‘P’ and ‘V’³³ and over the quantum of consideration on which to calculate the SDLT on (Nowlan J holding that the sum was £1.25b and Morgan J, with the casting vote, holding that the sum was £959m). We briefly note that Nowlan J felt constrained by what he termed to be the ‘numerous strict and mechanical rules ... not to be capable of purposive interpretation’.³⁴ We have previously noted the problems with the lack of power of the judges to be flexible in determining the relevant quantum of consideration. Nowlan J’s dissatisfaction of this point reflects the concerns of the judiciary with this issue, though it is conceded that Nowlan J did not actually advocate for changes to be made to this rule. It is noted that as Nowlan J was a member of the Aaronson Committee designing the GAAR he might be said to be particularly well informed on these points.

Apart from those issues, the judgment of the Upper Tribunal did not go much further than affirming the decision of the First-Tier Tribunal. However, the court did give extensive reasons for its rejection of the relevance of motive and actual tax avoidance as factors to be considered in determining the applicability of the Sections.³⁵ The issue of motive will be considered in detail subsequently.

E. THE FRAMEWORK INTENDED BY PARLIAMENT

A brief look at the wording of the Sections, coupled with an understanding of the background of the legislation and how GAARs work in general suggests that the general idea of the anti-avoidance framework here was one of a ‘catch-all with exceptions’. This would change the default position from the tax authority having to find a specific provision to cover the situation in question. Instead, HMRC would simply impose the relevant tax on any transaction which met very basic criteria. The onus would then be on the taxpayer to find the relevant exception to prove why he should not be taxed. This understanding of the framework is supported by various features of the Sections, including the very broad scope of the ‘trigger requirements’ in s 75A(1) and the fact that the Sections expressly include some specific exceptions to the application of the general rule. The court in *Project Blue* FTT argued that sub-sections 75C (11) and (12) were an indication that Parliament intended the scope of the Sections to be broad, providing for the eventuality that the provisions might ‘overshoot’.³⁶ Finally, counsel for the Respondent in *Project Blue* FTT argued that the exclusion of commercial transactions by sub-section 75A(7) was indicative of the broad scope of the provisions.³⁷

It is thus argued that Parliament intended for a ‘catch-all with exceptions’ approach. The most obvious exception can be found in sub-sections 75A(7), 75C (11) and (12). However, it cannot be the case that Parliament wished to provide only these three exceptions. The scope of sub-section 75A(7) is very specific and limited. A quick look at sub-sections 75C (11) and (12) tells us why they could not possibly have been intended to be adequate exceptions. Sub-section (11) provides that ‘the Treasury may by order provide for section 75A not to apply in specified circumstances’. Sub-section (12) provides that ‘an order under subsection (11) may include incidental, consequential or transitional provision and may make provision with retrospective effect.’

³² *Project Blue* (UT) (n 5).

³³ As defined in s 75A(1)(a). Not covered in this article. See (n 21).

³⁴ *Project Blue* (UT) (n 5) [168].

³⁵ *ibid* [50]-[58].

³⁶ *ibid* [225].

³⁷ *ibid* [144].

Thus, the Treasury has the power to retrospectively declare that certain transactions would not attract tax liability. However, the Treasury cannot reasonably be expected to go through the numerous different kinds of transactions that take place daily and pick out the ones which it believes should be exempt from SDLT. It is not practical for the Treasury to exercise this power as the main mechanism for limiting the scope of the Sections.

If the three sub-sections mentioned above were not meant to be the main regulators of the scope of the Sections, then it is likely that Parliament intended for some other regulators.³⁸ It is argued that the two best candidates for this position of ‘main regulator’ are 1) the discretion of HMRC; and 2) a ‘requisite connection’ test introduced by the terms ‘in connection with’ in s 75A(1)(b). Parliament may indeed have intended for these two regulators to work concurrently.

The literal interpretation of the scope of the Sections is broad enough to encompass all effective transfers of land. We must consider the possibility that the Sections were indeed intended to levy SDLT on all effective transfers of land, regardless of whether the transfer of land was merely incidental to the full transaction. This would mean that a sale of a manufacturing company, for instance, would attract SDLT liability on the transfer of the fee simple of the factory, even though the transaction was in substance one of a sale of the whole business and not one where the transfer of land was a major feature of the transaction.

It is inconceivable that the Sections were intended to have this effect. There was no mention of any desire to implement such a radical departure from existing practice when the Sections were being considered.³⁹ Furthermore, Guidance from HMRC suggests that the Sections were enacted with the intention of countering schemes that have the effect of reducing SDLT liability.⁴⁰ Therefore, there must clearly be some mechanism intended to control the scope of application of the Sections, and this mechanism cannot simply catch any transaction with some connection to the transfer of land.

IV. THE VARIOUS SOLUTIONS

A. SOLUTION (1): HAVE THE TREASURY EXPAND THE EXCEPTIONS UNDER SUB-SECTIONS 75C(11) AND (12)

Sub-section 75C

(11) The Treasury may by order provide for section 75A not to apply in specified circumstances.

(12) An order under subsection (11) may include incidental, consequential or transitional provision

and may make provision with retrospective effect.

³⁸ It is noted that the word ‘regulators’ in this case refers to factors which control the scope of application of the Sections. Parliament may have simply envisioned a narrower scope of application and not consciously considered some of these factors. The ‘requisite connection test’ in particular may not have been actively considered by Parliament but represents an attempt to give effect to Parliament’s general intention.

³⁹ Hansard records no debate regarding this point, which would be highly unusual if the proposed statute was intended to have the effect of widely expanding the incidence of SDLT.

⁴⁰ HM Revenue and Customs (n 2).

There is the possibility that Parliament intended that the Treasury should use sub-sections (11) and (12) to regulate the scope of the Sections generally by issuing subsidiary legislation as to the 'specified circumstances'. If this was indeed the intention, it has not been done as HMRC's current 'Whitelist' of transactions to which it will not seek to apply the test⁴¹ is mere advice and not subsidiary legislation, and thus lacks the force of law. It is argued that the complete lack of exercise of this power by the Treasury suggests that the power under sub-sections 75C (11) and (12) was meant to be exercised as a 'back-up', in case the main measures for determining the scope of the Sections missed specific instances which should not attract tax liability.

This does not mean that the sub-sections cannot be actively used to control the scope of the Sections now. If the Treasury creates exceptions on a case by case basis, then the use of the sub-s 75C(11) and (12) powers would indeed be very limited. If the exceptions are drafted to describe classes of transactions, then it may be of more use. On the whole, however, it seems impractical to expect the Treasury to constantly draft new exceptions to the rule whenever different kinds of transactions arise. This is especially so as there are so many potentially different kinds of transactions.

B. SOLUTION (2): RESTORE THE POWER OF DISCRETION TO HRMC ON WHERE TO APPLY THE SECTIONS

The First-Tier Tribunal in *Project Blue* considered the legal arguments regarding the existence of the power of discretion and eventually concluded that HMRC did not have that power of discretion. The court considered the Guidance Note⁴² issued by HMRC and its declaration that 'HMRC will not seek to apply 75A where it considers transactions have already been taxed appropriately'.⁴³ It held it to be at odds with the general obligation of HMRC to collect a tax imposed by Parliament.⁴⁴ The court cited the case of *Vestey v IRC*,⁴⁵ where HMRC had claimed the right to select which one of the beneficiaries of a discretionary trust to tax and to apportion the tax between several beneficiaries according to any method they thought fit.⁴⁶ HMRC was denied this discretion, with Lord Wilberforce laying down the principles of tax assessment.⁴⁷

The first key point of Lord Wilberforce's judgment is his decision that 'A proposition that whether a subject is to be taxed or not, or that, if he is, the amount of his liability is to be decided (even though within a limit) by an administrative body, represents a radical departure from constitutional principle'.⁴⁸ He then went on to require express enactment by Parliament before the courts would give effect to such a proposition.⁴⁹ The second key point of Lord Wilberforce's judgment is his reason for objecting to such a proposition, which was that it 'would be taxation by self-asserted administrative discretion and not by law... one should be taxed by law, and not be untaxed by concession'.⁵⁰ From these points, the court in *Project Blue* (FTT) went on to conclude that 'unless it clearly provides otherwise, section 75A should be construed as not giving HMRC a discretion whether to apply the statute nor as conferring on HMRC a discretion either whom to tax or as to the amount of tax to be levied'.⁵¹

⁴¹ HM Revenue and Customs (n 2).

⁴² *ibid*.

⁴³ *Project Blue* (FTT) (n 4) [218].

⁴⁴ *ibid* [219].

⁴⁵ [1980] AC 1148 (HL).

⁴⁶ *Project Blue* (FTT) (n 4) [220].

⁴⁷ *Vestey* (n 45) 1172-1173.

⁴⁸ *ibid* 1172.

⁴⁹ *ibid*.

⁵⁰ *ibid* 1173.

⁵¹ *Project Blue* (FTT) (n 4) [223].

From a perspective of strict judicial precedent, *Vestey* is a decision of the House of Lords that has not been overruled. It might be distinguished on its specific facts, but the rather broadly worded judgment of Lord Wilberforce makes this difficult. A bolder (and, in my opinion, better) argument would simply be to distinguish *Vestey* on the grounds that it does not apply where it potentially conflicts with GAARs. *Vestey* was decided in 1979 and there have been considerable developments in the field of tax law since then. The most important development in this regard is the introduction of the GAAR and the accompanying changes in the general approach to tax assessment.

By their nature, GAARs have to rely on a certain degree of administrative discretion in practice. The scope of the power to tax is drafted broadly, with the onus on the tax authority to exercise discretion whether to question the transaction. The courts are then relied upon to give an authoritative judgment on the issue if the matter is contested. In the vast majority of cases, however, the courts are not involved and the discretion whether to tax a particular transaction is in effect made by the tax authority. As there has not yet been any GAAR litigation in the UK at the current moment, this is based on GAAR litigation in other jurisdictions (e.g. Canada, Australia and New Zealand) and on the framework outlined by the Aaronson Report.⁵²

One can see that there is something inherently objectionable with allowing the tax authority untrammelled discretion to ‘over-tax’. However, HMRC deals with a very large volume of tax cases and arguably should be much better placed than the courts to distinguish between those cases which Parliament would wish to impose SDLT on and those which Parliament considers not within the ambit of the tax. In light of this, allowing them some discretion may not be objectionable. Furthermore, there is always the possibility of review of their decisions.⁵³ If they ‘under-tax’ in a few cases, that may be an acceptable cost of efficient administration of the tax system and not sufficient cause for removing their discretion.

That HMRC should have a discretionary power as to the cases to which the Sections apply is wholly consistent with the position that they have been taking to date; namely that they view s 75A as an anti-avoidance provision applying only where there is avoidance of tax and that they will not seek to apply it where they consider that transactions have already been taxed appropriately.⁵⁴ The idea is further supported by an argument from the law of statutory interpretation: that ‘Official statements by government departments are important because Acts are supervised by a government department which may be assumed to know what the legislative intention was. Their interpretation of the Act, as inferred from their official statements, is likely to coincide with the intended interpretation of the Act as passed by Parliament’.⁵⁵ The fact that HMRC consistently affirmed that they had the discretion to determine the kinds of cases to apply s 75A to is suggestive that this was indeed part of Parliament’s framework.

The idea is also compatible with the concept of institutional competence. Instead of attempting to draft ever more complex legislation which is unlikely to cover all the eventualities, it may be better to draft the legislation broadly and then let the experts decide which cases to apply the legislation to. In this case, the best experts would be HMRC, who undoubtedly have the most experience in distinguishing between genuine cases of tax avoidance and innocent cases.

⁵² The Aaronson Report (n 15).

⁵³ Applications for review have been brought by third parties before, even if the taxpayer and HMRC are comfortable with the arrangement to ‘under-tax’. See *R v Inland Revenue Commissioners, ex p National Federation of Self-employed and Small Businesses Ltd* [1982] AC 617, [1981] STC 260. For more recent cases, see *R (UK Uncut Legal Action Limited) v HMRC* [2013] EWHC 1283 (Admin) and *R v IRC, ex parte Wilkinson* [2005] UKHL 30, [20] – [23] (Lord Hoffmann).

⁵⁴ HM Revenue and Customs (n 2).

⁵⁵ Oliver Jones, *Bennion on Statutory Interpretation* (6th edn, LexisNexis Butterworths 2013) 231.

Following *Vestey* and *Project Blue* (FTT) strictly would mean that HMRC would have to apply the Sections to every single case where the conditions of s 75A(1) are met unless clear exceptions can be found. The taxpayers would then have to appeal each of these individual cases to the courts to determine if they have any tax liability. Given the broad scope of s 75A(1), this would be completely impractical and almost impossible to enforce.

Returning to our previous point of the framework intended by Parliament, it is most unlikely that Parliament considered that the Sections would be applied in this way. There are strong arguments for the proposition that Parliament intended for the discretion of HMRC to be one of the ‘main regulators’ of the scope of the Sections, and now that *Project Blue* (FTT) has denied HMRC this power, this is likely to create a significant strain on the other regulating mechanisms. The most straightforward way to solve this issue would be for Parliament to pass an amendment, updating the Sections to include a provision similar to that found in s 207(1) Finance Act 2013 (The GAAR). Failing that, the Court of Appeal could recognise that times have changed since *Vestey* and affirm the discretion given to HMRC when *Project Blue* comes before it again.⁵⁶ If these two measures are not taken, the only alternative would be to place more reliance on the other regulating mechanisms of the scope of the Sections.

It is worth mentioning that this measure of allowing HMRC considerable discretion was not favoured by the GAAR Committee in the Aaronson Report. The committee argued that the ‘determination of a reasonable and just result is an issue which should be justiciable before the Tax Tribunal, and not left to HMRC’s discretion’.⁵⁷ However, it is noted that this recommendation was made together with a proposal that the courts should take into account guidance from HMRC approved by the GAAR panel.⁵⁸ For the GAAR, since guidance from HMRC must be taken into account, there is no need to allow them any further discretion. This may not be the case for the Sections, where no such provision applies. While some of the objections of the GAAR Committee would admittedly still apply to this proposed solution as well, it is worth considering giving more discretion to HMRC in the absence of any other provision involving them in the framework.

C. SOLUTION (3): THE ‘REQUISITE CONNECTION’ TEST

Section 75A(1)(b) requires that ‘a number of transactions (including the disposal and acquisition) are involved *in connection with* the disposal and acquisition (‘the scheme transactions’) in order for the transactions to fall within the scope of s 75A. The First-Tier Tribunal recognised that s 75A(1)(b) was crucial to the application of the Sections, being one of the three main requirements in s 75A(1) and governing the applications of sub-sections s 75A(1)(c), (4) and (5).⁵⁹

⁵⁶ Projected date: 23 March 2015.

⁵⁷ The Aaronson Report (n 15) 5.35- 5.40.

⁵⁸ Eventually included in the statute at Finance Act 2013, s 211.

⁵⁹ *Project Blue* (FTT) (n 4) [247].

After analysing the existing case law,⁶⁰ the court held that the meaning of ‘in connection with’ required that there be a sufficient linkage between the transactions and the disposal and acquisition of the land. In the court’s words, ‘a transaction which is part of a series of transactions will not be “involved” with other transactions simply because it is part of a series or sequence of successive conveyancing transactions. The linkage must be more than merely being a party in a chain of transactions and the test must be more than a “but for” test’.⁶¹ The court went on to hold that in this case, the sub-sale and the lease were ‘involved in connection with’ the disposal of the property, but provided no further guidance on the application of the proposed test.

It is argued that the proposed test by the courts is far too broad. As noted earlier, there is a high probability that Parliament intended for the test imposed to determine the requisite level of connection to be a main regulator of the scope of the Sections. In that case, the courts’ ‘more than but for’ test is too vague and has a high likelihood of being practically completely redundant: a sieve without any mesh. Regardless of whether the use of the ‘requisite connection’ mechanism as a way to control the scope of the sections was expressly considered by Parliament, it provides an excellent way to control the scope of the otherwise excessively broad ss 75A- 75C, provided we make some necessary changes.

One way which imposing a test on the requisite level of connection can help to regulate the scope of the Sections is to conceive ‘in connection with’ as a test for whether the transaction is substantively one of transferring the land, or whether the transfer of the land is merely incidental to the rest of the transaction. Returning to our example of the sale of a manufacturing company, the transfer of the factory building would indeed technically be part of the whole transaction, but be insufficiently connected to the other parts of the transaction because the transfer of the factory would only be incidental to the true purpose of the transaction, which would be the sale of the manufacturing company. Consequently, SDLT need not be paid on the transfer of the factory, since there was never a direct transfer, and s 75A would not operate to impose SDLT on a notional transfer since there are no ‘number of transactions’ which are sufficiently connected to the transfer of the factory (thus, s 75A(1)(b) is not satisfied). This would achieve an outcome which would not significantly extend the current scope of SDLT. This would not be achieved if we were to apply *Project Blue* in determining the applicability of the Sections.

The test thus becomes one of looking at the substance of the transaction and determining if SDLT should be imposed because it is a ‘land transaction’ or whether it really is another kind of transaction where the transfer of land just happens to take place. This is consistent with the general approach to tax law to look at the substance rather than the form of the transactions.⁶²

To remove any sort of assessment as to the substance of the transaction and impose a blanket application of SDLT on any transaction with the slightest link to land is not beyond the powers of Parliament. But that would be most unlikely in the context in which the Sections were enacted and in any case, caution ought to be exercised before we decide that Parliament intended to vastly expand the current scope of SDLT through what was thought by HMRC to be a mere anti-avoidance provision.

There is a need for a test of the substantive nature of the transaction and it is currently not apparent that any other part of the Sections provides such a test. In light of this, it is argued that Parliament probably intended for s 75A(1)(b) to fulfil this function. Even if this were not the case, there is a strong argument that it should now be used as such a test.

⁶⁰ *Project Blue* (FTT) (n 4) [248]-[250].

⁶¹ *ibid* [250].

⁶² Well established in case law and famously noted in the case of *WT Ramsay Ltd v IRC* [1982] AC 300 (HL).

D. SOLUTION (4): USE MOTIVE AS A REGULATORY MECHANISM

It is noted that the ‘requisite connection’ test discussed above could technically have ‘motive’ instead of, or in addition to ‘whether the transaction is in substance, one involving land’ as the factor determining if the requisite level of connection has been met. It is argued that the substance of the transaction is a better candidate for the ‘requisite connection’ test as there is a more pressing need for a test that screens out transactions with only incidental connections to land. If motive is to be an additional factor in the ‘requisite connection’ test, then the reasoning of the courts in *Project Blue* rejecting this, as analysed below must be addressed. It will ultimately be argued that the reasoning of the courts on excluding motive as a relevant factor is unpersuasive and that motive should be considered relevant, whether as a standalone test or as part of the ‘requisite connection’ test. The former option is preferable as it serves to enhance clarity regarding the tests. This is important as there may be up to four potential tests involved.

The reasons given by the Upper Tribunal can largely be divided into two main arguments: 1) there is a lack of any express provision allowing the courts to consider motive (‘the omission argument’); and 2) the Sections define ‘avoidance’ themselves, making it impossible to imply a requirement that the transaction must have the purpose of tax avoidance in the conventional sense (‘the definition argument’). It is noted that the court did consider the implications of rejecting motive as a regulatory mechanism but considered itself bound by the wording of the statute.⁶³

The omission argument notes that ‘the enacting words of section 75A do not contain any provision which limits its scope to a case where there is a relevant purpose of tax avoidance’ and that there are other provisions (including those within the Finance Act 2003 itself) which clearly do impose a requirement that there be such a purpose.⁶⁴ As a matter of statutory interpretation, the court considered this omission to be intentional and reflective of the intentions of Parliament that showing that tax avoidance was a relevant purpose was not of utmost importance when applying the Sections.

Given that there are examples elsewhere of the express inclusion of the requirement of motive in the same Act, the court has a strong case for arguing that proof of motive should be irrelevant. As argued above by Jones,⁶⁵ at least one canon of statutory interpretation permits such a conclusion and it is not an unreasonable inference on the facts. However, it might also be reasonable to argue that the Sections were enacted solely for the purpose of preventing anti-avoidance schemes and that the courts in *Project Blue* have become unnecessarily bogged down in the technicalities of the statutory language. In that case, it would be a very positive thing for the Court of Appeal to take a highly purposive approach to interpretation and give effect to the statute as Parliament intended.⁶⁶

The definition argument notes the numerous references in the headings or side-notes of the Sections to the words ‘anti-avoidance’. It acknowledges that it may be possible to argue that a purposive approach to construction should be taken and that the scope of the Sections should cover only cases of ‘avoidance’. However, the core of the definition argument is that section 75A already defines what is meant by a case of ‘avoidance’ itself. The court held that ‘section 75A explains that a case which comes within section 75A(1)(c) is a case of “avoidance” and the sections are to operate to counter that avoidance. It is therefore neither necessary nor appropriate to read more into the side notes and to hold that the side notes are to be taken to refer to an unstated requirement that there be a purpose of tax avoidance’.⁶⁷

⁶³ *Project Blue* (UT) (n 5) [56].

⁶⁴ *ibid* [52]. The relevant examples are schedule 7, para 2(4A) and schedule 8, para 1(3).

⁶⁵ Jones (n 55).

⁶⁶ The Court of Appeal is due to hear this case on 23 March 2015.

⁶⁷ *Project Blue* (UT) (n 5) [54].

This argument is a more persuasive one. However, even if it is inappropriate to infer from the headings and side-notes in the Sections that actual tax avoidance is required, motive will still be relevant if it can be successfully argued that Parliament actually intended it to be a requirement and merely inadvertently omitted it in the drafting process. Of course, regardless of the intentions of Parliament when the Sections were drafted, Parliament can establish the requirement of motive now through a statutory amendment. This would bring the Sections in line with how the GAAR is currently functioning and it is argued that Parliament should consider this.

V. MOVING FORWARD: MECHANISMS AND SOLUTIONS

A. OTHER MECHANISMS

When the Sections were first enacted, anti-avoidance measures were largely limited to TAARs and other specific forms of legislation. However, the tax authorities have continually developed more sophisticated measures and have successfully managed to get legislation establishing the GAAR and the DOTAS passed. With the GAAR covering (some) loopholes in the TAARs and the DOTAS managing the information asymmetry between the tax authorities and the taxpayers, one cannot help but wonder if the Sections are actually adding anything useful to the system, or are merely redundant. Given the complexity, excessively broad scope and problems with the Sections, if they do not have any useful functions, it might be better for them to be altogether abolished.

There are at least three ways to move forward with the Sections: 1) change the Sections to mirror the GAAR; 2) abolish the Sections and rely on the GAAR; or 3) keep the Sections, but modify them such that they add something to the tax framework. The third option is preferable if it is at all possible to do so. The strength of a mini-GAAR is that it is more specific and hence a sharper tool for distinguishing between innocent cases and cases of tax avoidance. In the context of SDLT and land transactions, what distinguishes the Sections from the GAAR is its potential sensitivity to when it would be reasonable to impose SDLT. It is argued that the criteria for distinguishing the cases would be the nature of the transaction and whether it is in substance, one of the disposal and acquisition of land. Thus, I would argue that the way to make the Sections play a useful function would be to focus on the requisite connection test and interpret the words ‘in connection with’ in s 75A(1)(b) in the manner proposed above.

B. UNDERSTANDING THE SECTIONS IN LIGHT OF THE CURRENT GAAR FRAMEWORK

The Aaronson Report and the subsequent enactment of the GAAR after the Sections help us to shed some light on the framework required for a GAAR to work effectively. The Aaronson Report makes it clear that the GAAR is meant to be aimed at ‘obvious’ and the most ‘egregious’ cases only. Furthermore, there are other safeguards in the legislation to protect taxpayers. In Appendix I of the Aaronson Report, the GAAR Committee lists four main safeguards, including the referral of potential counteraction to the Advisory Panel (Safeguard 4).⁶⁸ HMRC sought approval from the GAAR panel before issuing guidance on the application of the GAAR. The statute itself provides that such approved guidance must be taken into account by the court.⁶⁹ This seems to be the model that Parliament intended GAARs to be based on.

⁶⁸ The Aaronson Report (n 15) 44-54.

⁶⁹ Finance Act 2013, s 211.

There are indeed two ways to look at this issue. One could argue that since the GAAR expressly provides for such a framework and the Sections do not, Parliament did not intend that the framework should be applicable to the Sections as well. However, looking at the context in which the Sections are enacted, it is possible to advance an alternative argument. It is possible that Parliament intended for the Sections, as a mini-GAAR, to function as the GAAR was eventually enacted. However, as the idea of a GAAR was still being introduced in the United Kingdom at that time, the Sections were not drafted well enough to reflect Parliament's intention and failed to accurately capture this. Based on this argument, the intention of Parliament would be that all the safeguards to protect the taxpayer would be applicable to cases covered the Sections and not just limited to those cases caught by the GAAR.

It is worth explaining this argument in detail. It is not being proposed that the Parliamentary intention has changed and that the courts should give effect to the new intention. Rather, it is argued that the Parliamentary intention all along was to have a functioning and effective GAAR mechanism and what has changed is how the statute has been drafted to achieve that same outcome. Therefore, the mere omission of express safeguards in the Sections should not be taken to mean that Parliament has intended that such safeguards should not apply in these cases.

C. SUMMARY OF THE POTENTIAL SOLUTIONS

We have thus far discussed four potential solutions to the broad scope of the Sections: 1) have the Treasury expand the exceptions under sub-s 75C(11) and (12); 2) restore the power of discretion to HMRC on when to apply the Sections; 3) refine the requisite connection test and base it on whether the transaction is in substance, one of the disposal and acquisition of land; and 4) use motive as a regulatory mechanism. The relative merits of each potential solution have been discussed above, but it is noted that the relative ease of application of the solutions differ significantly. Solution (1) is the easiest to implement as it requires absolutely no change to the law; in statute or as established by *Project Blue*. All that is required is that the Treasury take a more proactive role in exercising its powers under ss 75C(11) and (12). There would be a slight advantage in terms of the speed with which this can be implemented. It is noted that the advantage would probably not be particularly significant as Parliament passes at least one Finance Bill every year that could potentially implement this solution. It is unlikely that the powers would be enough to solve the problems of the Sections in practical terms. It is argued that ss 75C(11) and (12) were intended as 'back-up' provisions and trying to go beyond that would place excessive strain on the system.

Solutions (2) and (4) have currently been rejected by the courts in *Project Blue* and it would be more difficult to implement them. Judicial or legislative intervention would be required, with the former requiring the provision of reasons for departing from the existing legal position. Going with these two methods has a bonus in that they result in the clarification of the Sections as anti-avoidance measures and nothing else. It remains to be seen if the Court of Appeal in the pending appeal of *Project Blue* will take either of these approaches.⁷⁰ These approaches would probably be sufficient to solve the problems that we currently have with the Sections, though it is noted that solution (2) would be better executed by Parliament than the courts.⁷¹ The model provisions can already be found in the GAAR and Parliament would simply have to enact an amendment to introduce the 'reasonableness' requirement as in the GAAR to the Sections.

⁷⁰ There currently being no directly binding precedent on this issue from the higher courts, the Court of Appeal has free rein to implement solutions (2) and (4).

⁷¹ If *Vestey* is held to be good law and binding on the Court of Appeal, then it is possible to argue that solution (2) can only be implemented by Parliament.

Solution (3) technically does not go against the current state of the law. What it does require is a shift in mindset in how the Sections are to be construed. As noted above, solution (3) has the distinct benefit of being able to perform a unique function of determining which transactions are substantially land transactions and therefore should be subject to SDLT in the conventional application of that head of tax. It is argued that solution (3), with its ability to properly identify the kind of transactions which would have been subjected to SDLT if not for the scheme transactions, is the best solution to adopt. While solutions (3) and (4) are distinct, they can also be applied together. Solution (3) can be adjusted to include considerations of motive if it is decided by the courts that it should be relevant to the application of the Sections. Otherwise, it works perfectly well when considerations of motive are irrelevant.

Finally, it is noted that a combination of the proposed solutions may be used. Indeed, whilst it is argued that the 'requisite connection' test is the best solution, each solution focuses on different (though overlapping) factors and different situations might call for the use of different judicial tools. In some cases, more than one tool might be required. Much depends on the circumstances of each case.

VI. CONCLUSION

The Sections were subjected to criticism when enacted⁷² and indeed subsequent cases have shown that their application is procedurally difficult and practically almost unworkable. In the rapidly developing field of tax law, statutes must be understood in both the context in which they were enacted and the subsequent developments which have taken place since. The introduction of a GAAR has significantly changed the playing field and we must be prepared for even greater and more rapid changes as the GAAR is implemented.

The main problem of the excessively broad scope of the Sections arose because of a relatively literal and inflexible interpretation of the Sections, coupled with bad drafting and insufficient information on the intention behind their enactment. This article has sought to show that while the judges naturally felt constrained by the wording of the Sections, there was and still is considerable room for a broader construction which may be able to avoid the major problems currently faced. The ideal solution here would simply be for Parliament to intervene and replace the Sections with a much better drafted revised version. However, in the interim, the courts still do have room to manoeuvre and are not necessarily forced into making impractical and unjust decisions. Ultimately, tax law is constantly evolving and Parliament will have to keep pace with the developments by constantly updating its statutes. The GAAR may have evened the fight between the tax authorities and the taxpayer briefly but we can be sure that human ingenuity will soon require more sophisticated anti-avoidance measures.

The enactment of the GAAR also introduces a new dimension to the rationale for the TAARs or mini-GAARs. There is absolutely no point in having mirror provisions that do the same thing as the GAAR. If the TAARs are to be useful, they will have to be drafted such that they strike a balance between being specific enough to distinguish between the innocent cases and the tax avoidance cases; and being broad enough to catch cleverly designed avoidance schemes. With the GAAR in place, it is argued that the focus should now be more on the 'specific' side of things, for the GAAR's function is to deal with the 'broad' side of things. Attempting to perfectly balance both functions may lead to a situation where the legislation achieves nothing and may even be counterproductive. The GAAR has not rendered TAARs or mini-GAARs redundant, but they need to find their own niche in order to serve useful functions in the overall legal framework.

⁷² See N. Lee (n 3).