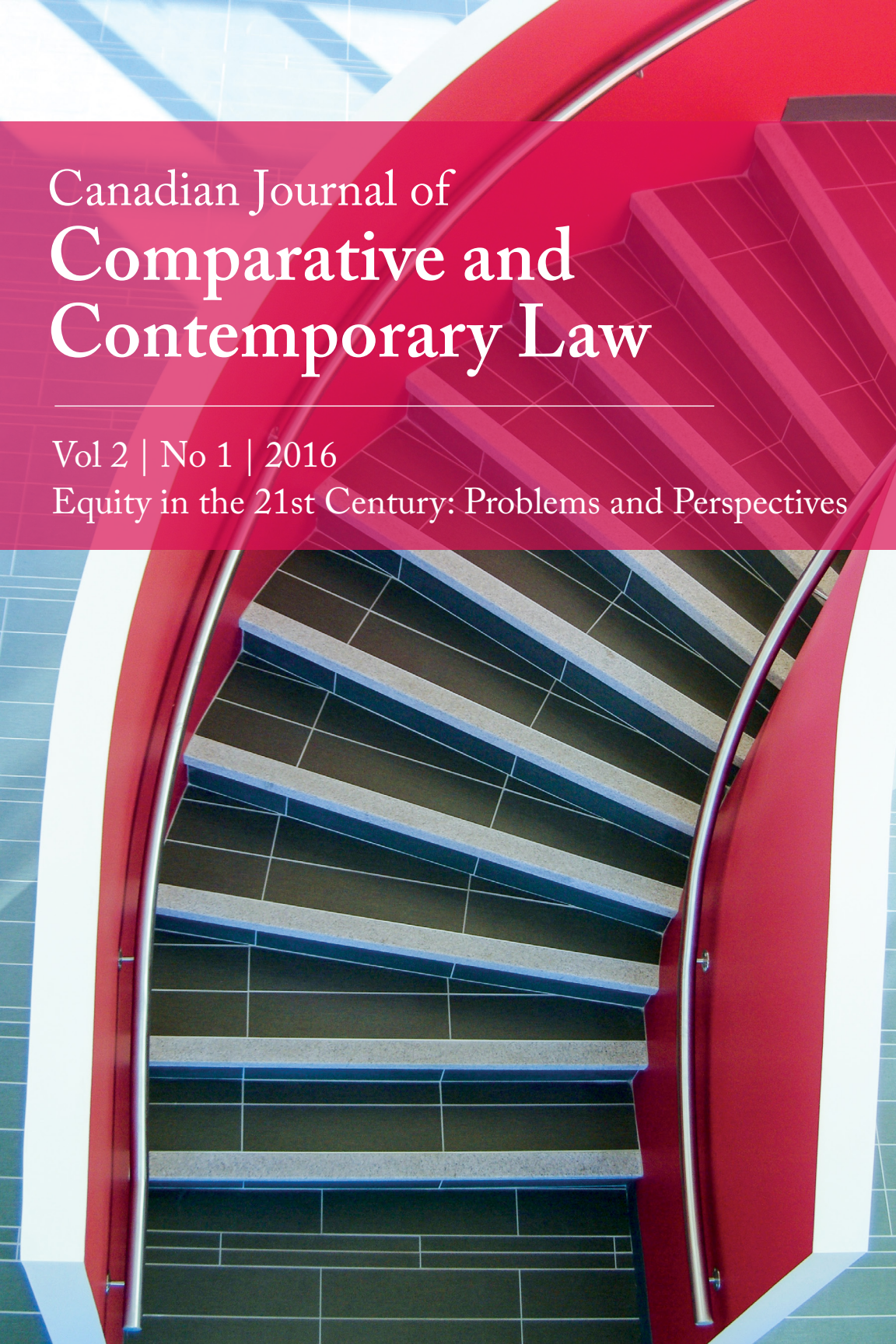

Canadian Journal of Comparative and Contemporary Law

Vol 2 | No 1 | 2016

Equity in the 21st Century: Problems and Perspectives

Canadian Journal of

Comparative and Contemporary Law

Vol 2 | No 1 | 2016

Equity in the 21st Century: Problems and Perspectives

The CJCCL is published by
The Canadian Association of Comparative and Contemporary Law
at Thompson Rivers University
Kamloops, BC

Canadian Journal of Comparative and Contemporary Law

Publication
The Canadian Journal of Comparative and Contemporary Law (CJCCL) is an open-access publication that is available online at http://www.cjccl.ca. Hardcopies can be ordered on request. Each issue focuses on a particular theme or area of law. The CJCCL encourages contributors to take a comparative approach in their scholarship.

Editorial Policy
All submissions are subject to a peer review process.

Submissions
The Journal accepts the following types of manuscripts:
(i) Articles between 8,000 to 15,000 words in length;
(ii) Case Comments between 3,000 to 6,000 words in length; and
(iii) Book Reviews less than 3,000 words in length.
Please visit our website for more details.

Contact Information
Canadian Journal of Comparative and Contemporary Law
Thompson Rivers University
Faculty of Law
900 McGill Road
Kamloops, BC, Canada V2C 0C8

E-mail: editor@cjccl.ca
Web: http://www.cjccl.ca

Cover Photo
The front cover depicts the main stairwell that leads to the atrium of Thompson Rivers University, Faculty of Law. The back cover depicts the distinct exterior of the Faculty of Law. The curved design of the roof was inspired by the natural beauty of the mountains visible from the building.

ISSN 2368-4046 (Online)
ISSN 2368-4038 (Print)

This Issue should be cited as (2016) 2(1) CJCCL

Canadian Journal of Comparative and Contemporary Law

List of Contributors

Robert Chambers, Professor of Private Law, King's College London.

Kathryn Chan, Assistant Professor, Faculty of Law, University of Victoria.

Paul S Davies, Associate Professor, Faculty of Law, University of Oxford; Fellow, St Catherine's College.

Mark Gillen, Professor, Faculty of Law, University of Victoria.

Margaret Isabel Hall, Associate Professor, Faculty of Law, Thompson Rivers University.

Matthew Harding, Professor of Law, University of Melbourne.

Alastair Hudson, Professor of Equity & Finance Law, University of Exeter National Teaching Fellow.

Craig Jones, Professor, Faculty of Law, Thompson Rivers University.

Hila Keren, Professor of Law, Southwestern Law School.

Mark Leeming, Judge of Appeal, Supreme Court of New South Wales.

Mitchell McInnes, Faculty of Law, University of Alberta.

Richard C Nolan, Anniversary Professor of Law, University of York.

Leonard I Rotman, Purdy Crawford Chair in Business Law and Professor, Schulich School of Law, Dalhousie University.

Irit Samet, Reader in Private Law, Dickson Poon School of Law, King's College, London.

Graham Virgo, Faculty of Law, University of Cambrige.

Stephen Watterson, University Lecturer in Law, University of Cambridge, and John Collier Fellow in Law, Trinity Hall, Cambridge.

David Wiseman, Assistant Professor, Faculty of Common Law, University of Ottawa.

Sarah Worthington, QC (Hon) FBA, Downing Professor of the Laws of England and Fellow of Trinity College, Cambridge.

Canadian Journal of Comparative and Contemporary Law

Volume 2 | Number 1 | 2016
Equity in the 21st Century: Problems and Perspectives

Canadian Journal of Comparative and Contemporary Law

Volume 2 | Number 2 | 2016
Equity in the 21st Century: Problems and Perspectives

Foreword

Justice Russell Brown
Supreme Court of Canada

After last year's successful inaugural issue of the *Canadian Journal of Comparative and Contemporary Law* (CJCCL) dedicated to "Health Law and Human Rights", I am honoured to provide this brief foreword to the CJCCL's second issue, containing papers exploring the theme of "Equity in the 21st Century: Problems and Perspectives".

As a theme, "equity in the 21st century" is intriguing. At first glance, one might suppose that "equity" and "21st century" are as suitably juxtaposed as "Thomas Aquinas" and "emoji". While its origins are murky, we know that equity emerged long ago from the administrative power of the mediaeval Chancellor, to whom the King had delegated the task of hearing pleas from his subjects concerning injustices at the common law courts. As most law students also know, the Chancellor, as an ecclesiastic, was concerned with conscience (wherein lay the path to the immortal soul). It was therefore on the basis of "conscience" that he exercised this delegated remedial power by ordering respondents to act according to good conscience, notwithstanding their legal rights to do otherwise.

This account — while accurate — risks, however, descending into caricature in several respects. First, common law courts were not amoral wastelands. Still, their limited forms of action could work injustice. Clear rules were preferred over avoiding hardship. John H Baker's famous example of the paid debt that must be paid a second time (owing to the debtor's failure to ensure the debt was cancelled after it was paid the first time) illustrates the sort of problem that typically arose.[1] Secondly, while the Chancellor's jurisdiction ultimately widened from the "wide

1. John H Baker, *An Introduction to Legal History*, 3d (London: Butterworths, 1990) at 118.

but vague"[2] powers wielded during the Middle Ages, the Chancellor's conscience-based jurisdiction soon narrowed, as the stream of *ad hoc* decisions were inevitably reduced to rules or principles of equity which, by the late 18th century, were as inflexible and prone to working injustice as the common law itself. By the mid 20th century, the English Court of Appeal could unashamedly proclaim in *Re Diplock* that it lacked jurisdiction to do equity on the mere basis that "we may think that the 'justice' of the present case requires it".[3]

As the latter half of the 20th century showed, however, equity had not rolled over and died. As Leonard I Rotman argues in his article on the Supreme Court of Canada's understanding of the fusion of law and equity, in recent decades equity has brought the positive law "closer to the human condition". Longstanding devices such as resulting and constructive trusts, injunctions and estoppel were extended, and the action in unjust enrichment (assuming it can truly be understood as "equitable") was pulled from its post-*Moses v Macferlan*[4] obscurity. Courts have breathed new life into equity's concern for the conscientious exercise of legal rights in property and under terms of contracts. Concerns for "fairness" and "justice" have predominated.

A more muscular equitable doctrine to quell "unfairness" and "injustice", however, raises its own set of concerns, none of which are new. Is equity really nothing more than a body of sentimental goo to be haphazardly applied when the spirit of fairness and justice moves us? Or should the conditions calling for its intervention be stated (if they can be stated) precisely and exhaustively? The obvious criticism is that a purely "I-know-unconscionability-when-I-see-it" approach is nothing more than palm tree justice. Equity would lack intelligibility, clarity and predictability in application, thereby implicating basic norms of the rule of law.[5] Little wonder Professor Donovan Waters used to warn his trust law

2. Paul Vivian Baker & Peter St John Hevey Langan, *Snell's Equity*, 29d (London: Sweet & Maxwell,1982) at 8.
3. *Re Diplock*, [1948] Ch 465 at 481 (CA (Eng)), aff'd *sub nom Ministry of Health v Simpson*, [1951] AC 251 (HL).
4. (1760), 97 ER 676 (KB).
5. Tom Bingham, *The Rule of Law* (London: Penguin Group, 2010) at 37.

students against distorting the remedial constructive trust's conscience-based origins by taking it "too far". That seems fair. Nobody committed to equity's public repute wants to see a new maxim proclaiming equity (or the remedial constructive trust) to be the last refuge of a scoundrel. At the same time, there has been no expressed appetite for a return to the rigid systematization that made *Re Diplock* possible.

This tension is never far from the surface in this splendid collection of essays. For example, Alastair Hudson maintains that we err by treating the organizing principle of "conscience" as an entirely subjective phenomenon, as opposed to the product of objectively constituted sources of normative behaviours. And, although not directly addressed to the subjective-vs.-objective dichotomy, Sarah Worthington's paper seeks, by way of analysis of the proscriptive rules which equity imposes upon fiduciaries, to bring principle to determining who is a fiduciary (fellow Canadians, please pay attention![6]), the obligations they owe, and the remedies which flow from a breach. In contrast, Hila Keren, in lamenting "the fall of equity", strikes a more subjective note by celebrating (or, more accurately from her standpoint, commemorating) equity's "non-economic" priorities of "morality, fairness, justice or equality". The other papers implicitly presume that conscience is either an objective reference point, or that — if it has a subjective dimension — such subjectivity need not defeat clear thinking and rational rule-making in equity. Richard C Nolan's article demonstrating the importance of inherent jurisdiction to the administration of trusts celebrates the innovative judicial extension of that jurisdiction, for example, to give directions where the trustee is caught between competing groups of holders of notes issued under the terms of a trust deed. At the same time, though, he calls for greater attention to identifying a theoretical basis for deciding when a court can or cannot exercise inherent jurisdiction in this fashion. Mark Gillen espouses more radical reform of trust law by providing for enforcement of certain non-charitable purposes trusts — not, however, by way of a subjective act of judicial discretion, but by way of legislative intervention.

Nolan and Gillen's papers are also representative of a distinctly

6. *CA v Critchley* (1998), 60 BCLR (3d) 92 (CA) at para 75.

pioneering flavour to the collection — not inappropriately, given the 21th century theme. They are joined in this respect by Irit Samet's consideration of whether the law should abandon *caveat emptor* to permit rescission for unilateral mistakes in contract formation, and of whether equity is the appropriate vehicle to effect such a reform; Matthew Harding's deep reflection about discriminatory public trusts, whether judges should prefer the threshold of "public benefit" over "public interest", and the place of the value of freedom of disposition under each threshold; Kathryn Chan's argument for reinvigorating the Federal Court of Appeal's equitable jurisdiction over registered charities by invoking certain curative principles oriented towards effectuating imperfect charitable gifts; Paul Davies' consideration of whether the rules of equitable compensation should follow those available for breach of contract, for negligence, or for other torts; and David Wiseman's account for the possibility that equity might prefer a beneficiary-protection power over a beneficiary-direction rule in the context of joint bank account resulting trusts. And, appropriately enough (given this Journal's dedication to comparative law), Graham Virgo seeks to harmonize the Australian and Canadian remedial constructive trust with the institutional (substantive) constructive trust in England, while Margaret Hall considers the applicability to Canadian law of a recent English judicial innovation, rooted in the equitable doctrine of undue influence, for disrupting relationships that exploit children and mentally incapable adults. At a more general level, Justice Mark Leeming shows why equity is especially suited to comparative analysis, comprising themes which are familiar to jurists throughout the common law world.

Breaking new ground can, however, be difficult work, and the results are not always universally embraced. Robert Chambers shows how, over several decades of debate about liability for knowing receipt of assets transferred in breach of trust or fiduciary duty, various accounts for such liability have been advanced — from unjust enrichment, to failure to perform a duty to restore the misapplied trust property, to Chambers' admirably plain-spoken and persuasive argument that knowing receipt is itself a breach of trust. Any resort to "waiver of tort", once hoped to be equity's elixir for overcoming indeterminate causation in mass tort

claims,[7] must now account for Craig Jones' account of its limits. Stephen Watterson explores the residual uncertainties in the wake of the House of Lords' decision in *Banque Financière de la Cité v Park (Battersea) Ltd*,[8] following which English law has had to sort out what it means to say that subrogation to extinguished rights (usually held by a disappointed unsecured creditor or by a lender whose funds were misappropriated to discharge another's liabilities) is not only a remedy, but one that is also equitable and restitutionary. And as Mitchell McInnes' essay on beneficial services in respect of land shows, the implications of expanding the scope for equitable relief — in that particular case for relief under proprietary estoppel — are often insufficiently examined, both before and after the fact.

Our legal community, howsoever one chooses to define it (Western Canadian, Canadian, Anglo-American, common law), owes a debt of appreciation to Thompson Rivers University's Faculty of Law on this initiative — the CJCCL, with its worthwhile themes and its first-rate content. The essays contained in its second issue deserve wide circulation among practicing and academic lawyers and, of course, among judges charged with doing equity. As one who has profited from reading them, I offer my thanks and congratulations to all concerned and, in particular, to the authors and to the CJCCL's editorial team.

7. *Serhan (Estate Trustee) v Johnson & Johnson* (2004), 132 ACWS (3d) 221 (Ont Sup Ct).
8. [1999] 1 AC 221 (HL).

The End of Knowing Receipt

Robert Chambers*

This article addresses the nature of liability for knowing receipt of assets transferred in breach of trust, and argues that it is no different from liability for breach of trust. It arises because the recipient has obtained assets that are held in trust and, after becoming aware of the trust, has failed to perform the basic trust duties to preserve the trust assets and transfer them to the proper trustees. It is not a form of restitution of unjust or wrongful enrichment, so it should not matter whether the assets were received for the recipient's own benefit.

* Professor of Private Law, King's College London. This paper has a long history. A shorter version was first presented at the Higher Courts Judges Conference in Napier, New Zealand in 2011, and then at the University of Melbourne and University of Western Australia. The present version was recently presented at the National University of Singapore. I am grateful for the kind invitations to present this paper, the helpful comments and questions received, and the delightful and memorable introduction provided by Justice Sir Robert Chambers (1953-2013) when it was first presented.

I. Introduction

The law regarding personal liability for knowing receipt of assets transferred in breach of trust or fiduciary duty has received an extraordinary amount of academic and judicial attention over the past 30 years.[1] Yet despite this flurry of attention (or perhaps because of it), the law in this area remains in a muddled and unsatisfactory state. There are disagreements over the various elements of the cause of action, which stem from a lack of consensus over the basic nature of the liability: is it a form of restitution of benefits received, compensation for losses caused, or something else? Part of the problem is the language used in this area. Words and phrases, such as "the first limb of *Barnes v Addy*", "liability to account as a constructive trustee", or even "knowing receipt" itself, tend to obscure more than they reveal. While complex concepts do require specialist terminology, it is possible to speak plainly in this area and reveal more.

A frequently quoted statement of the essential elements of liability

1. The modern interest in the subject can be traced to a series of cases in the 1970s and 1980s in which assets were misappropriated from companies by their directors or officers, and perhaps the longest article ever published in the Law Quarterly Review: Charles Harpum, "The Stranger as Constructive Trustee" (1986) 102 Law Quarterly Review 114-62, 267-91.

for knowing receipt was by Lord Justice Hoffmann in *El Ajou v Dollar Land Holdings*:[2]

> This is a claim to enforce a constructive trust on the basis of knowing receipt. For this purpose the plaintiff must show, first, a disposal of his assets in breach of fiduciary duty; secondly, the beneficial receipt by the defendant of assets which are traceable as representing the assets of the plaintiff; and thirdly, knowledge on the part of the defendant that the assets he received are traceable to a breach of fiduciary duty.[3]

While succinct, each part of this statement raises questions about the nature and ambit of knowing receipt. What does it mean to "enforce a constructive trust on the basis of knowing receipt"? On what basis does liability arise for "a disposal of [the plaintiff's] assets in breach of fiduciary duty" if those assets were not held in trust? Why is "beneficial receipt by the defendant" required? What degree of "knowledge on the part of the defendant" will suffice?

I must confess that I once believed, as did the late Professor Peter Birks, that liability for knowing receipt was best understood as a form of restitution of unjust enrichment. I was first introduced to the subject as a doctoral student at a seminar at All Souls College in 1992.[4] Enthusiasm for an explanation based on unjust enrichment was running high and was persuasively promoted in the writing of Peter Birks and others at the

---

2. [1994] BCC 143 (CA (Civ)(Eng)).
3. *Ibid* at 154. Quoted in *Bank of Credit and Commerce International (Overseas) Ltd v Akindele*, [2001] Ch 437 (CA (Civ)(Eng)) at para 34 [*Akindele*]; *Caltong (Australia) Pty Ltd v Tong Tien See Construction Pte Ltd*, [2002] 3 SLR 241 (CA) at para 31; *Ultraframe (UK) Ltd v Fielding*, [2005] EWHC 1638 (Ch) at para 1478 [*Ultraframe*]; *First Energy Pte Ltd v Creanovate Pte Ltd*, [2006] SGHC 240 at para 53; *Comboni v Shankar's Emporium (Pte) Ltd*, [2007] SGHC 55 at para 49; *Zambia v Meer Care & Desai*, [2007] EWHC 952 (Ch) at para 515; *OJSC Oil Co Yugraneft v Abramovich*, [2008] EWHC 2613 (QB (Comm)) at para 248; *Zage v Rasif*, [2008] SGHC 244 at para 14; *Arthur v A-G Turks & Caicos Islands*, [2012] UKPC 30 (T&C) at para 32 [*Arthur*]; *Otkritie International Investment Management Ltd v Urumov*, [2014] EWHC 191 (QB (Comm)) at para 81.
4. See Peter Birks, ed, *The Frontiers of Liability* (Oxford: Oxford University Press, 1994) Part I.

time.[5] If based on unjust enrichment, there seemed no good reason to insist on knowledge, notice, or some element of fault on the part of the recipient as a condition of liability. Strict liability coupled with the defence of change of position then seemed both logical and inevitable. Peter later retreated from that position, accepting that liability for knowing receipt was based on fault, but with liability for restitution of unjust enrichment as an added string to the plaintiff's bow.[6]

Much of what follows has been said before, although not all in one place. The law in this area is not (or least does not have to be) as complicated as it appears. Some basic principles can be stated, and although some are controversial, they are set out below in the hope that this might help resolve some of the uncertainty and controversy in the area. Perhaps that is too much to expect, but at least it cannot hurt to state things as clearly and simply as possible, and at least hope not muddy the waters any further.

Simply stated, liability for knowing receipt is nothing other than liability for breach of trust. It arises because the recipient has obtained assets that are held in trust, and after becoming aware of the trust, has failed to perform the basic trust duties to preserve the trust assets and transfer them to either the beneficiaries or the proper trustees. This requires actual knowledge of the trust or the circumstances giving rise to it. Notice is insufficient. This is not a form of restitution of unjust or wrongful enrichment, so it should not matter whether the assets were received for the recipient's own benefit. The recipient is an actual trustee and not just being treated as if that was true. This is not a form of accessory or secondary liability. It is fundamentally different from liability for knowingly assisting a breach of trust or fiduciary duty.

Liability for knowing receipt depends upon receiving trust assets and holding them in trust. Therefore, if the recipient obtains title free of

---

5. *Ibid*. See the following essays in that collection: Peter Birks, "Gifts of Other People's Money" Ch 31; Charles Harpum, "The Basis of Equitable Liability" Ch 9 at 24-25; William Swadling, "Some Lessons from the Law of Torts" Ch 41.
6. Peter Birks, "Receipt" in Peter Birks & Arianna Pretto, eds, *Breach of Trust* (Oxford: Hart Publishing, 2002) 213 [Birks & Pretto, *Breach of Trust*].

the trust as a *bona fide* purchaser or through indefeasibility of registered title, liability for knowing receipt is not possible. Where assets were not held in trust prior to receipt, but were misappropriated from a company in breach of fiduciary duty, liability for knowing receipt is not possible unless a trust arises.

None of this precludes the possibility of a separate claim for restitution of unjust enrichment. However, there is no need to recognise a new equitable cause of action to achieve this. The recipient of misappropriated trust funds can be personally liable at common law for restitution of the value of those funds, subject to the defences of *bona fide* purchase and change of position.

II. Breach of Trust

The most important contribution to this area of law in recent years is an essay by Professor Charles Mitchell and Dr Stephen Watterson called "Remedies for Knowing Receipt".[7] They demonstrate convincingly that liability for knowing receipt cannot be explained in terms of unjust enrichment, but is the liability for failing to perform a duty to "restore the misapplied trust property".[8] Where I depart from them is in their reluctance to describe this as a breach of trust. This reluctance was not shared by Mr Simon Gardner, who described knowing receipt as "liability for breach of trust",[9] and went on to say:

> 'knowing receipt' is simply the usual liability for failure to preserve trust property, applicable to all trustees, given particular application to those who are trustees because they receive illicitly transferred trust property. The cognisance requirement in 'knowing receipt' is no more than a reminder that, before a

7. Charles Mitchell & Stephen Watterson, "Remedies for Knowing Receipt" in Charles Mitchell, ed, *Constructive and Resulting Trusts* (Oxford: Hart Publishing, 2010) at 115.
8. *Ibid* at 132. See *Arthur*, *supra* note 3 at para 37. See also Michael Bryan, "Recipient Liability under the Torrens System: Some Category Errors" in Charles Rickett & Ross Grantham, eds, *Structure and Justification in Private Law: Essays for Peter Birks* (Oxford: Hart Publishing, 2008) 340 at 342-44.
9. Simon Gardner, "Moment of Truth for Knowing Receipt?" (2009) 125:1 Law Quarterly Review 20 at 22.

> trustee who loses trust property will thereby breach his duty to preserve it, he must have been aware (or could have been aware, or whatever standard is chosen) of the need to preserve it, *i.e.* of the facts giving rise to the trust.[10]

Mitchell and Watterson base their approach on the nature of the accounting process through which trustees can become personally liable to pay for the value of misapplied trust assets even in the absence of an allegation of breach of trust:

> [b]ecause the main liability of a knowing recipient is to perform his primary duty of restoration in just the same way as an express trustee, it is a distinctive form of liability which cannot be collapsed into other forms of liability which arise in the law of wrongs or the law of unjust enrichment.[11]

Having discussed this at length with Charles Mitchell, this is a point over which we agree to disagree. This may be only tangentially related to the issues at hand since it concerns the liabilities of all trustees and not just knowing recipients. However, in pursuit of the goal of speaking plainly, the liability of the knowing recipient is most usefully explained simply as liability for breach of trust.[12]

Beneficiaries are entitled to an account from their trustees because maintaining and providing accounts are primary duties of the trustee. No allegation of breach of duty is required. The account can then be falsified (by striking out unauthorised dispositions) or surcharged (by adding assets which the trustees failed to obtain), leading to a personal liability to pay.[13] In either case, the adjustment depends on a breach of duty by commission or omission. A trustee who properly performs the trust is never personally liable to pay for losses to the trust. As Lord Justice Lindley said in *Re Chapman*,[14] "a trustee is not a surety, nor is he an insurer; he is only liable for some wrong done by himself, and loss

---

10. *Ibid* at 23.
11. *Supra* note 7 at 136.
12. *Ibid*, (Mitchell and Watterson describe it as "specific or substitutive performance of his primary duty").
13. *Ultraframe*, *supra* note 3 at para 1513-17; *Glazier v Australian Men's Health (No 2)*, [2001] NSWSC 6 (Austl) at para 38, rev'd [2002] NSWCA 22 (Austl) at para 13; Robert Chambers, "Liability" in Birks & Pretto, *Breach of Trust*, *supra* note 6, 1 at 16-20.
14. [1896] 2 Ch 763 (CA (Eng)).

of trust money is not *per se* proof of such wrong".[15] It is true that an authorised disposition might be struck out if the trustees failed to keep adequate records and are therefore unable to prove that it was authorised, but again, that liability arises from a breach of their duty to maintain adequate records.

Breach of trust, like many other breaches of duty, does not require dishonesty or neglect. Honest, well-meaning trustees may be strictly liable for unauthorised dispositions of the trust assets.[16] Having undertaken a duty to perform the trust, they can be liable for failing to do so. The office of express trustee can be onerous. However, it is a startling proposition to say that trustees who have properly performed the trust could yet become personally liable to dig into their own pockets.

This proposition becomes more startling still when applied to knowing recipients. According to Mitchell and Watterson,[17] their liability to pay is generated neither by wrongdoing nor unjust enrichment, and since they have not undertaken the office of express trustee, it cannot be explained in terms of consent. If true, we are left in search for some other explanation why people who have done no wrong, received no benefit, and made no undertaking or agreement, can or should be liable to pay. This is difficult to justify to other lawyers or judges, let alone to a lay person subjected to that liability, and creates a justifiable fear that something has gone wrong with the analysis.

The better explanation is that people who know they hold assets transferred to them in breach of trust are under trust duties to preserve those assets and restore them to the proper persons. Any other use of the assets is unauthorised and a breach of trust, which may lead to a liability to pay.

The knowing recipient's liability to pay can be generated by the

15. *Ibid* at 775.
16. *Eaves v Hickson* (1861), 30 Beav 136 (Ch (Eng)) at 141; *National Trustees Co of Australasia Ltd v General Finance Co of Australasia Ltd*, [1905] AC 373 (PC (Austl)) at 379**.**
17. *Supra* note 7.

accounting process without an allegation of breach of trust.[18] The duty to account arises on proof that the recipient received the trust assets and acquired knowledge of the beneficiaries' claim.[19] Presumably, there is no duty to account before those two conditions are satisfied, so the beneficiaries would bear the onus of proving (with the aid of the normal litigation discovery process) that the recipient still held trust assets when sufficient knowledge of the breach was acquired. At that point, the onus would shift to the recipient to account as trustee for any subsequent dealings with those assets.

III. Source of the Knowing Recipient's Duties

Explaining the liability to pay in terms of breach of duty helps, but does not provide a complete solution, because it does not explain the source of the duty breached. Knowing recipients do not consent to the office of express trustee and have not undertaken the duties associated with it. When they become aware of the trust they do not thereby assume all of the duties of an express trustee, but become subject only to the duties to preserve the trust assets and restore them to the proper persons. If the proper persons cannot be identified, the duty to preserve trust assets appears to include the duty to invest trust money in an interest-bearing bank account.[20]

While it can be risky drawing analogies with other areas of law, this is somewhat similar to the duty of care that can be imposed on a bailee of goods, even though the bailee has had no direct dealings with the owner and may even be a finder of the goods.[21] As Justice Blanchard said in *R v Ngan*,[22] "[a]t common law any person who finds an item of property and takes possession of it on behalf of the true owner as a temporary custodian is treated as a bailee of that property and is under an obligation

18. *Ibid* at 136; *Green v Weatherill*, [1929] 2 Ch 213 (KB (Eng)) at 222-23 [*Green*].
19. *Green, ibid*.
20. *Evans v European Bank Ltd*, [2004] NSWCA 82 (Austl) at para 162; *supra* note 7 at 138-40.
21. *Newman v Bourne & Hollingsworth* (1915), 31 TLR 209 (KB (Eng)).
22. [2008] 2 NZLR 48 (SC).

to keep it safe and return it to the owner (if that is possible)".[23] And in *P & O Nedlloyd BV v Utaniko Ltd*,[24] Lord Justice Mance said:

> [a]s a matter of principle and because the essence of bailment is the bailee's voluntary possession of another's goods, an owner's remedies cannot necessarily be confined to situations involving either a direct bailment or a sub-bailment. A's goods may come into the possession of B as a voluntary bailee in other circumstances. … When ascertaining the scope of bailment in contemporary legal conditions, there is general wisdom in Professor Palmer's observation that: 'The important question is not the literal meaning of bailment but the circle of relationships within which its characteristic duties will apply. For most practical purposes, any person who comes knowingly into the possession of another's goods is, prima facie, a bailee.'[25]

Similarly, the knowledge that one has obtained title to an asset that is held in trust for another carries with it the limited duties to preserve the asset and get it back where it belongs. The beneficiaries of the trust cannot enforce all their rights under the express trust against the knowing recipient. Most of those rights are rights *in personam* that can be enforced only against the express trustees who have voluntarily undertaken the corresponding duties to perform the trust with care, loyalty, etc. However, the beneficiaries' right to have the trust assets held and managed by properly appointed trustees can be enforced more generally against others.[26]

It is tempting to explain the recipient's duties to preserve and restore trust assets on the basis of consent, since those duties (or at least the liability for their breach) depend on knowledge of the trust. The comparable duties of the bailee were explained on the basis of "an assumption of responsibility" by Lord Pearson, giving the advice of the Privy Council in *Gilchrist Watt & Sanderson Pty Ltd v York Products Pty*

---

23. *Ibid* at 59.
24. [2003] EWCA Civ 83.
25. *Ibid* at para 26 (quoting Norman Palmer, *Bailment*, 2d (Sydney: The Law Book Company Limited, 1991) at 1285).
26. See Richard Nolan, "Equitable Property" (2006) 122:2 Law Quarterly Review 232 ("a beneficiary's core proprietary rights under a trust consist in the beneficiary's primary, negative, right to exclude non-beneficiaries from the enjoyment of trust assets" at 233).

Ltd:[27]

> both in an ordinary bailment and in a 'bailment by finding' the obligation arises because the taking of possession in the circumstances involves an assumption of responsibility for the safekeeping of the goods. … [A]lthough there was no contract or attornment between the plaintiffs and the defendants, the defendants by voluntarily taking possession of the plaintiffs' goods, in the circumstances assumed an obligation to take due care of them and are liable to the plaintiffs for their failure to do so … [28]

This is perhaps the best we can do, but the duties of the knowing recipient look like they were imposed by operation of law rather than having been voluntarily undertaken. They arise even if the knowing recipient was an active participant in a scheme to misappropriate assets from the trust, clearly with no intention whatsoever to undertake any trust obligations towards the beneficiaries. They can also arise when the defendant honestly receives the assets and only later discovers the breach of trust. That is different from the honest finder who chooses to take possession of a lost item, but not unlike someone who accepts goods unaware they were delivered by mistake.

Turning to the law of wrongs does not help, because the duties to preserve and restore trust assets arise even if the recipient is honest and fully intends to perform them. It is tempting to say that they arise because the recipient's conscience is affected by knowledge of the trust,[29] but that is merely a conclusion and does not explain why it is affected. If conscience requires the preservation and restoration of the trust assets, it can only be because there are duties to do so. The appeal to conscience does not help identify the source of those duties nor the precise conditions that must exist before they arise.[30]

The law of unjust enrichment might explain why the trust arises in cases where the assets were not held in trust before receipt or where the assets received are the traceable proceeds of the assets originally

---

27. [1970] 3 All ER 825 (PC (Austl)).
28. *Ibid* at 831-32. See also Robert Stevens, *Torts and Rights* (Oxford: Oxford University Press, 2007) at 10-11.
29. *Akindele*, *supra* note 3 at paras 55-70.
30. Birks, *supra* note 6 at 226.

misappropriated from the trust.[31] It could be said that the recipient has been unjustly enriched at the expense of the trust beneficiaries by receipt of those assets. However, this does not explain the superadded duty of care that arises when the recipient acquires knowledge of the trust. That duty arises because the recipient knows he or she is holding assets in trust and it does not matter whether the trust is express, resulting, or constructive, nor why it has arisen.[32]

IV. Knowledge or Notice

Many of those who have argued that liability for knowing receipt is based on fault have also said that recipients can be liable even if they did not know that the assets were transferred to them in breach of trust but had only notice of that fact.[33] Professor Charles Harpum wrote:

> [b]ecause the issue in cases of knowing receipt is essentially a proprietary one, a recipient of trust property may be liable as a constructive trustee if he failed to make the inquiries that he ought to have made, even though he acted in good faith. It is taken for granted in the cases that constructive notice of the impropriety of the transfer suffices for liability, and the emphasis is on whether the circumstances were such as to put the recipient on inquiry.[34]

---

31. *Foskett v McKeown*, [2001] 1 AC 102 (HL)(appears to rule out this possibility, at least in cases where the assets were misappropriated from a trust) [*Foskett*]; see James Penner, "Value, Property, and Unjust Enrichment: Trusts of Traceable Proceeds" in Robert Chambers, Charles Mitchell & James Penner, eds, *The Philosophical Foundations of the Law of Unjust Enrichment* (Oxford: Oxford University Press, 2009) 306 at 306. Compare Peter Birks, "Property, Unjust Enrichment, and Tracing" (2001) 54:1 Current Legal Problems 231; Andrew Burrows, "Proprietary Restitution: Unmasking Unjust Enrichment" (2001) 117:3 Law Quarterly Review 412; Robert Chambers, "Tracing and Unjust Enrichment" in Jason Neyers, Mitchell McInnes & Stephen Pitel, eds, *Understanding Unjust Enrichment* (Oxford: Hart Publishing, 2004) 263 [Chambers, "Unjust Enrichment"].
32. Gardner, *supra* note 9 at 24.
33. David Fox, "Constructive Notice and Knowing Receipt: An Economic Analysis" (1998) 57:2 Cambridge Law Journal 391 at 391.
34. Charles Harpum, "The Stranger as Constructive Trustee: Part 2" (1986) 102:2 Law Quarterly Review 267 at 273.

Justice Millett (as he then was) said in *Agip (Africa) Ltd v Jackson*:[35]

> the person who receives for his own benefit trust property transferred to him in breach of trust … is liable as a constructive trustee if he received it with notice, actual or constructive, that it was trust property and that the transfer to him was a breach of trust, or if he received it without such notice but subsequently discovered the facts. In either case he is liable to account for the property, in the first case as from the time he received the property and in the second as from the time he acquired notice.[36]

This was quoted with approval by the Supreme Court of Canada, in *Citadel General Assurance Co v Lloyds Bank Canada*,[37] where Justice La Forest went on to say:

> relief will be granted where a stranger to the trust, having received trust property for his or her own benefit and having knowledge of facts which would put a reasonable person on inquiry, actually fails to inquire as to the possible misapplication of trust property. It is this lack of inquiry that renders the recipient's enrichment unjust.[38]

Later in an *obiter dictum* in *Twinsectra Ltd v Yardley*[39] ("*Twinsectra*"), Lord Millett said:

> [l]iability for 'knowing receipt' is receipt-based. It does not depend on fault. The cause of action is restitutionary and is available only where the defendant received or applied the money in breach of trust for his own use and benefit: see *Agip (Africa) Ltd v Jackson* [1990] Ch 265, 291-292; *Royal Brunei Airlines Sdn Bhd v Tan* [1995] 2 AC 378, 386. There is no basis for requiring actual knowledge of the breach of trust, let alone dishonesty, as a condition of liability. Constructive notice is sufficient, and may not even be necessary. There is powerful academic support for the proposition that the liability of the recipient is the same as in other cases of restitution, that is to say strict but subject to a change of position defence.[40]

The most notable exception to this trend is the judgment of Vice

---

35. [1990] Ch 265 (QB (Eng)).
36. *Ibid* at 291 (there was no appeal on the issue of knowing receipt: *Agip (Africa) Ltd v Jackson,* [1990] EWCA Civ 2 at 567).
37. [1997] 3 SCR 805 at para 42 [*Citadel*].
38. *Ibid* at para 49.
39. [2002] UKHL 12 [*Twinsectra*].
40. *Ibid* at para 105. See also *Dubai Aluminium Co Ltd v Salaam*, [2002] UKHL 48 at para 87; Peter Millett, "Tracing the Proceeds of Fraud" (1991) 107:1 Law Quarterly Review 71 at 80-83 [Millet, "Tracing"].

Chancellor Megarry in *Re Montagu's Settlement Trusts*[41] ("*Re Montagu*"), in which a person had received assets in breach of trust and disposed of them with notice, but no knowledge of the breach. Megarry VC held that personal liability for knowing receipt "primarily depends on the knowledge of the recipient, and not on notice to him; and for clarity it is desirable to use the word 'knowledge' and avoid the word 'notice' in such cases".[42]

In *Bank of Credit and Commerce International (Overseas) Ltd v Akindele*,[43] Lord Justice Nourse said "[t]he recipient's state of knowledge must be such as to make it unconscionable for him to retain the benefit of the receipt".[44] The reference to knowledge might be taken as an endorsement of Megarry VC's view, and this is how it has been interpreted in at least one subsequent case.[45] However, Nourse LJ also stated that it is unnecessary to distinguish "between actual and constructive knowledge" in this context.[46] Since constructive knowledge includes "knowledge of circumstances which will put an honest and reasonable man on inquiry", it appears to extend to notice without knowledge.[47]

All agree that liability for knowingly assisting a breach of trust or fiduciary duty requires actual knowledge (perhaps with an added

41. [1987] 1 Ch 264 (HC (Eng)).
42. *Ibid* at 285.
43. [2001] Ch 437 (CA (Civ)(Eng)).
44. *Ibid* at 455.
45. In *Papamichael v National Westminster Bank*, [2003] EWHC 164 (QB), Judge Chambers QC ("in *Akindele*, the application of the precept to the facts of that case seems to leave little room for manoeuvre. The case … makes it pretty clear that the type of knowledge that is required is actual rather than constructive knowledge. Such a requirement does away with the suggestion of a balance having to be struck between the relative urgency of a transaction and the degree of notice required: if you know, you know" at para 247).
46. *Supra* note 43 at 455.
47. *Ibid* at 454.

element of dishonesty).[48] There are two different reasons (revealed in the quotations above) why the receipt of assets might justify imposing liability on the basis of notice without knowledge: one because it involves the receipt of property and the other because it involves the receipt of a benefit. However, as discussed below, neither reason justifies a reduction from knowledge to notice, and if it did, there would be no reason to stop at notice: strict liability should be the logical result.

V. *Bona Fide* Purchase

Since liability for knowing receipt depends on receiving assets held in trust, it cannot arise if recipients take the assets free of the trust as *bona fide* purchasers for value without notice. This will not change if they later acquire notice or knowledge of the breach. They are free to continue to use and enjoy the assets as they please and can sell them to others who know of the breach of trust.[49] Otherwise, the defence of *bona fide* purchase would fail to protect them adequately, and a well publicised breach of trust would destroy the market value of the assets. Knowledge only matters if the recipient is still holding the assets in trust when that knowledge is acquired.

If the purchasers have notice of the breach of trust and are therefore not protected by the defence of *bona fide* purchase, it does not necessarily mean they can be personally liable for knowing receipt. Although the trust will survive the transaction, recipients with notice may honestly be unaware of the breach, and in that case should not be personally liable for disposing of the assets in breach of trust. These are two different questions, as Megarry VC pointed out in *Re Montagu*:

> the equitable doctrine of tracing and the imposition of a constructive trust by reason of the knowing receipt of trust property are governed by different rules and must be kept distinct. Tracing is primarily a means of determining

48. *Air Canada v M & L Travel Ltd*, [1993] 3 SCR 787; *Royal Brunei Airlines Sdn Bhd v Tan*, [1995] 2 AC 378 (Brunei) [*Royal Brunei*]; *Twinsectra, supra* note 39; *Barlow Clowes International Ltd v Eurotrust International Ltd*, [2005] UKPC 37 (Isle of Man) [*Barlow*]; *Enbridge Gas Distribution Inc v Marinaccio*, 2012 ONCA 650 [*Enbridge*].

49. *Wilkes v Spooner*, [1911] 2 KB 473 (CA (Eng)) at 487.

> the rights of property, whereas the imposition of a constructive trust creates personal obligations that go beyond mere property rights.[50]

This language was criticised by Lord Millett (writing extra-judicially) as "unhelpful": "tracing is not a means of determining property rights; it is not even confined to proprietary claims: while the constructive trust does not necessarily attract personal obligations at all".[51] Nevertheless, there is a fundamental difference between the beneficiaries' proprietary interest in the trust assets (or their traceable proceeds) and their personal claim against the recipient for failing to preserve those assets and restore them to the proper parties.

It is one thing to purchase an asset and find out it is less valuable than expected. The purchaser will usually have a claim against the vendor for breach of warranty of title. However, it would be going too far to increase the purchaser's woes by adding personal liability to the beneficiaries of the trust. We are willing to enforce property rights generally against others (subject to rules protecting honest buyers) because they do not impose positive obligations against others, but only negative limitations on their use or enjoyment of things.[52] Setting the standard at notice can be regarded as consistent with this (although registration statutes usually provide greater levels of protection). To impose positive obligations normally requires consent, wrongdoing, or at least a level of knowledge that permits the recipient to make a choice whether to incur that liability or not. Notice short of knowledge does not suffice.

If, as Professor Harpum suggested (above), notice is sufficient for liability because "the issue in cases of knowing receipt is essentially a proprietary one",[53] there would be no good reason why notice should be required in cases where the recipient is a donee. Although we often refer to the defence of *bona fide* purchase as the "doctrine of notice", it should not be forgotten that notice is entirely irrelevant when assets are

---

50. *Supra* note 41 at para 58; quoted with approval in *Arthur*, *supra* note 3 at para 34.
51. Peter Millett, "Restitution and Constructive Trusts" (1998) 114:3 Law Quarterly Review 399 at 403.
52. *Rhone v Stephens,* [1994] 2 AC 310 (HL) at 318.
53. *Supra* note 34 at 273.

acquired by a donee. Taken to its logical conclusion, a justification based solely on the priority rules governing equitable interests would lead to the conclusion that liability for knowing receipt should be strict, subject to the defence of *bona fide* purchase.[54]

VI. Indefeasibility

If the *bona fide* purchaser is immune to liability for knowing receipt, the same must also be true when the asset received is registered title to land and the recipient is protected by the indefeasibility provisions of the registration statute. The increased protection provided by registration may mean that liability for knowing receipt is not possible even if the recipient has actual knowledge of the breach at the time. In a Torrens system, where indefeasibility is denied to a registered proprietor guilty of actual fraud, this depends on how the courts define fraud. It is not fraud to know that the land was held in trust,[55] nor that registration will destroy the beneficiaries' interest in the land.[56] There is no duty to inquire into the possibility that the land is being transferred in breach of trust,[57] but it had long been understood that it is fraud to know or suspect a breach of trust.[58]

Recently, Australian courts have decided that it is not fraud for the registered proprietor to obtain land knowing it was transferred in breach of trust so long as the trustees were not guilty of fraud. In *Farah Constructions Pty Ltd v Say-Dee Pty Ltd*[59] ("*Farah Constructions*"), the High Court of Australia said that Torrens fraud means "actual fraud, moral turpitude" and if the trustee is not guilty of "actual fraud", then

54. See note 55.
55. See *e.g. Land Titles Act*, RSA 2000, c L-4, s 203; *Land Transfer Act* 1952 (NZ), 1952/S2, s 182; *Land Titles Act* (2004 Rev Ed Sing), s 47; *Transfer of Land Act* 1958 (Vic (Austl)), s 43.
56. *RM Hosking Properties Pty Ltd v Barnes*, [1971] SASR 100 at 103 (SC (Austl)).
57. See Millet, "Tracing", *supra* note 40.
58. *Assets Co Ltd v Mere Roihi*, [1905] AC 176 (PC (NZ)) at 210; See also *Arthur*, *supra* note 3 at para 40.
59. [2007] HCA 22.

"the other parties are in no worse position".[60]

In *LHK Nominees Pty Ltd v Kenworthy*,[61] land was purchased at half price from a trust company by its agent, who knew this was a breach of trust. Since there was no proof he had deceived the directors of the trust company, he obtained an indefeasible title. Justice Murray said "there would be no capacity to defeat the indefeasibility of title conferred by the Act by reason merely that title to the land was acquired in circumstances in which the recipient knew that the transfer to him was in breach of trust".[62]

This is probably not the law outside Australia.[63] It is at least inconsistent with the law of knowing assistance as developed by the Privy Council. In *Royal Brunei Airlines Sdn Bhd v Tan*,[64] Lord Nicholls said that it is dishonest to participate in breach of trust, even if the trustee is acting honestly:

> [u]nless there is a very good and compelling reason, an honest person does not participate in a transaction if he knows it involves a misapplication of trust assets to the detriment of the beneficiaries. Nor does an honest person in such a case deliberately close his eyes and ears, or deliberately not ask questions, lest he learn something, he would rather not know, and then proceed regardless.[65]

There has been some debate over whether a claim for knowing receipt falls within the "*in personam* exception" to indefeasibility.[66] This may seem plausible if liability for knowing receipt is seen as a form of restitution of

---

60. *Ibid* at para 192.
61. [2002] WASCA 291 (Austl)(leave to appeal dismissed [2003], HCATrans 426).
62. *Ibid* at para 185.
63. See *Arthur*, *supra* note 3.
64. *Royal Brunei*, *supra* note 48.
65. *Ibid* at 389. See also *Barlow*, *supra* note 48.
66. This phrase was coined following a comment made by Lord Wilberforce in *Frazer v Walker*, [1967] 1 AC 569 (PC (NZ))(that the principle of immediate indefeasibility "in no way denies the right of a plaintiff to bring against a registered proprietor a claim *in personam*, founded in law or in equity, for such relief as a court acting *in personam* may grant" at 585).

unjust enrichment.[67] The debate then turns to the question whether "the important functions of land registration would be stultified if knowing receipt were allowed to operate against a registered purchaser".[68] A majority of the Victoria Court of Appeal thought that this would indeed be the consequence. In *Macquarie Bank Ltd v Sixty-Fourth Throne Pty Ltd*,[69] Justice Tadgell said that liability for knowing receipt "would introduce by the back door a means of undermining the doctrine of indefeasibility which the Torrens system establishes".[70] This was cited with approval in an *obiter dictum* of the High Court of Australia in *Farah Constructions*,[71] and so this problem seems to be resolved in Australia.

If liability for knowing receipt is understood as liability for breach of trust, then the problem created by the *in personam* exception does not arise in this context. A recipient who obtains indefeasible title free of the trust cannot be a trustee and cannot be subject to the trust duties to preserve and restore trust assets.

It should not be assumed that a claim for knowing receipt will lie against a registered proprietor who obtained title by fraud. Although the recipient holds only a defeasible title and the trust will survive registration, he or she may be unaware of the trust. For example, it is fraud to submit a document for registration knowing that it was executed improperly, and its registration will not defeat any existing trust of the land.[72] However, if the proprietor later disposes of the land and dissipates the proceeds in ignorance of the trust, there is no reason why he or she should be liable for breach of that trust.

---

67. See UK, Law Commission, *Land Registration for the Twenty-First Century* (Law Commission No 254)(HM Land Registry, 1998) at para 3.48. See also *supra* note 8.
68. Matthew Conaglen & Amy Goymour, "Knowing Receipt and Registered Land" in Charles Mitchell, ed, *Constructive and Resulting Trusts* (Oxford: Hart Publishing, 2010) 159 at 174.
69. [1998] 3 VR 133 (SC (Austl)).
70. *Ibid* at 157.
71. *Supra* note 59, at paras 193-96.
72. *Australian Guarantee Corp Ltd v De Jager*, [1984] VR 483 (SC (Austl)) at 497-98.

VII. Beneficial Receipt

If liability for knowing receipt is not restitutionary and is not based on unjust enrichment, there is no reason why it is necessary that the assets were received for the recipient's own benefit. If assets are misappropriated from a trust for A and transferred to a recipient in trust for B, this is a simple question of priorities.[73] If the recipient holds the assets subject to two inconsistent trusts, one will take priority over the other (usually, but not always, depending on which trust arose first).[74] If the recipient of assets in trust for B knows that they were previously held in trust for A and transferred in breach of that trust, then it must be a breach of the trust for A if the recipient deals with them inconsistently with it, even if the actions are taken in obedience to the trust for B.

It has been assumed that receipt for one's own benefit is a requirement for liability, and it is perhaps this assumption that leads to the conclusion that notice would be sufficient for imposing liability.[75] However, if benefit is the key, then liability should be restitutionary and limited to the actual benefit obtained by the recipient. Also, there is no good reason to stop at notice, as Lord Millett noted in *Twinsectra*.[76] Strict liability, subject to the defence of change of position is a far more sensitive means of achieving that goal. It can protect honest, well-meaning recipients who did not benefit from their receipt of the assets but nevertheless had notice of the breach of trust. Conversely, it can permit liability in cases where the assets were spent without notice of the trust, but on necessary expenditures that have left the recipient with a surviving enrichment at the beneficiaries' expense.[77]

---

73. This is essentially what happened in *Foskett*, *supra* note 31.
74. *Abigail v Lapin*, [1934] AC 491 (PC (Austl)).
75. *Citadel, supra* note 37 at para 48; *Gold v Rosenberg*, [1997] 3 SCR 767 at para 46.
76. *Twinsectra*, *supra* note 39 at para 105.
77. See *Heperu Pty Ltd v Belle*, [2009] NSWCA 252 (Austl) at para 80.

VIII. Company Assets

If liability for knowing receipt is liability for breach of the trust duties to preserve the trust assets and restore them to the proper persons, then it cannot arise unless the assets are held in trust. Knowledge or notice of a breach of fiduciary duty is insufficient in the absence of a trust. The extension of liability to cases involving the misappropriation of company assets by directors and officers makes sense only if a trust arose by operation of law when the assets were received (or at some earlier or later stage).

Where company assets have been transferred pursuant to a contract with the company, the first and most important question is whether the contract is binding on the company. This is an issue concerning the actual or ostensible authority of the company's agents, which has nothing to do with knowing receipt.[78] It is true that similar questions may be involved. Agents do not have authority to deal fraudulently with their principal's assets, and anyone who knows of the fraud cannot be relying on the agent's ostensible authority.[79] However, over the last 30 years, there has been a tendency to ignore this fundamental question of the validity of the contract and jump straight into the law of knowing receipt.

This problem was identified by the House of Lords in *Criterion Properties v Stratford UK Properties LLC*[80] ("*Criterion*"), where Lord Nicholls said:

> if a company (A) enters into an agreement with B under which B acquires benefits from A, A's ability to recover these benefits from B depends essentially on whether the agreement is binding on A. If the directors of A were acting for an improper purpose when they entered into the agreement, A's ability to have the agreement set aside depends upon the application of familiar principles of agency and company law. If, applying these principles, the agreement is found to be valid and is therefore not set aside, questions of 'knowing receipt' by B do not arise. So far as B is concerned there can be no question of A's assets having been misapplied. B acquired the assets from A, the legal and beneficial owner

78. Robert Stevens, "The Proper Scope of Knowing Receipt" (2004) Lloyd's Maritime and Commercial Law Quarterly 421.
79. Peter Watts, "Authority and Mismotivation" (2005) 121:1 Law Quarterly Review 4 at 7.
80. [2004] UKHL 28.

> of the assets, under a valid agreement made between him and A. If, however, the agreement is set aside, B will be accountable for any benefits he may have received from A under the agreement. A will have a proprietary claim, if B still has the assets. Additionally, and irrespective of whether B still has the assets in question, A will have a personal claim against B for unjust enrichment, subject always to a defence of change of position. B's personal accountability will not be dependent upon proof of fault or 'unconscionable' conduct on his part. B's accountability, in this regard, will be 'strict'.[81]

As Lord Nicholls pointed out, if the contract is voidable, the company will have a proprietary claim to any recoverable assets obtained by the other party to that contract.[82] This must also be true where the contract is void or where there is no contract and company assets are simply misappropriated and transferred to the recipient. If the recipient is holding assets in trust for the company, then (and only then) does the possibility of liability for knowing receipt arise.

If assets have been transferred pursuant to a voidable contract that has not been avoided, there is no trust but merely a power to avoid the contract and thereby create a trust. The power to recover assets through rescission is an equitable interest in the recoverable assets (sometimes called a "mere equity"), but it is not beneficial ownership under a trust.[83] The recipient is bound by the contract until it is avoided. Rescission may have retroactive effect so that the trust is deemed to have arisen at the outset, but that cannot retroactively turn the actions of the recipient at a time when there was no trust into a wrongful breach of trust. As Lord Millett said (writing extra-judicially):

---

81. *Ibid* at para 4.
82. There was no discussion of the rescission of the contract of sale in *Arthur*, *supra* note 3, but the claim was brought by the Crown as vendor seeking to recover the land sold on the basis that Her minister had acted in breach of fiduciary duty by arranging the sale at a price significantly below market value. The must be understood as a claim to rescind the transaction.
83. Peter Birks, "Property and Unjust Enrichment: Categorical Truths" (1997) New Zealand Law Review 623 at 637-48; Richard Nolan, "Dispositions Involving Fiduciaries: The Equity to Rescind and the Resulting Trust" in Peter Birks & Francis Rose, eds, *Restitution and Equity Volume One: Resulting Trusts and Equitable Compensation* (London: Routledge, 2000) 119 at 132; Chambers, "Unjust Enrichment", *supra* note 31 at 300.

> [i]n all these cases the beneficial interest passes, but the plaintiff has the right to elect whether to affirm the transaction or rescind it. If he elects to rescind it, it is usually assumed that the beneficial title revests in the plaintiff, and the authorities suggest that it does so retrospectively. But the recipient cannot anticipate his decision. Pending the plaintiff's election to rescind, the recipient is entitled, and may be bound, to treat the payment as effective. It is well settled that the plaintiff's subsequent rescission does not invalidate or render wrongful transactions which have taken place in the meantime on the faith of the receipt.... Pending rescission the transferee has the whole legal and beneficial interest in the property, but his beneficial title is defeasible. There is plainly no fiduciary relationship. The defeasible nature of the transferee's title should not inhibit his use of the property.[84]

If there is no trust pending rescission, there cannot be any duties to preserve trust assets and restore them to the company. The absence of trust would not, however, preclude an action for knowingly assisting a breach of fiduciary duty.

IX. Unjust Enrichment

In the quotation from *Criterion*, above, Lord Nicholls raised the possibility of a personal claim for unjust enrichment. It should not be assumed that he was referring to liability for knowing receipt. Writing extra-judicially, he suggested that courts should recognise an additional form of personal liability to make restitution of unjust enrichment operating concurrently with liability for knowing receipt.[85] This has been supported by Lord Walker writing extra-judicially and others.[86]

Lord Nicholls envisaged a new personal claim in equity.[87] Lord Millett suggested that the common law was not up to the task,[88] but it is not clear why that should be so. In most cases, it will be necessary to

---

84. *Supra* note 51 at 416.
85. Lord Nicholls, "Knowing Receipt: The Need for a New Landmark" in William Robert Cornish et al, eds, *Restitution Past, Present and Future: Essays in Honour of Gareth Jones* (Oxford: Hart Publishing, 1998) 231.
86. Lord Walker, "Dishonesty and Unconscionable Conduct in Commercial Life—Some Reflections on Accessory Liability and Knowing Receipt" (2005) 27:2 Sydney Law Review 187 at 202; Birks, *supra* note 6; Gardner, *supra* note 9 at 24.
87. *Supra* note 85.
88. Millett, "Tracing", *supra* note 40 at 76-80.

show that the assets transferred to the recipient are the traceable proceeds of the assets misappropriated from the trust. It has long been assumed that equity's tracing rules are superior to those of the common law, but it seems only a matter of time before the tracing rules are seen simply as rules of evidence that are the same regardless of the nature of the claim involved.[89] If equity is not needed for the tracing process, and the claim is merely for the value of the assets received, there seems no reason why this cannot be done at common law.

In most cases where funds have been misappropriated from a company, they were not held in trust prior to the misappropriation. The company will have a claim at common law for restitution of unjust enrichment, or as we still like to think of it, an action for money had and received. There is no need to turn to equity except to assert a proprietary interest in assets in the recipient's hands or to make a claim based on knowing receipt of those assets.

The same principles should apply in cases where assets have been misappropriated from a trust.[90] While the beneficiaries do not have a direct claim at common law against the recipient, the trustees can sue at common law to recover money paid by mistake or in breach of trust. That claim is a trust asset which the trustees are required to realise. The beneficiaries can compel them to do their duty, and if necessary, have them replaced or possibly even bring the action with the trustees joined

89. See *Foskett*, *supra* note 31 at 113, 128-29, per Lord Steyn and Lord Millett; *Ultraframe*, *supra* note 3 at paras 1461-64; *BMP Global Distribution Inc v Bank of Nova Scotia*, [2009] 1 SCR 504 at paras 75-85; Peter Birks, "The Necessity of a Unitary Law of Tracing" in Ross Cranston, ed, *Making Commercial Law: Essays in Honour of Roy Goode* (Oxford: Clarendon Press, 1997) 239.

90. Simon Gardner, "Knowing Assistance and Knowing Receipt: Taking Stock" (1996) 112:1 Law Quarterly Review 56 at 86 [Gardner, "Taking Stock"]; Gardner, *supra* note 9 at 24.

as defendants.[91]

In the important but difficult case of *Lipkin Gorman v Karpnale Ltd*,[92] the House of Lords recognised both the right to restitution of unjust enrichment and the defence of change of position. The facts are well known. Mr Cass, a partner in a firm of solicitors, misappropriated funds from the firm's client trust account and gambled their traceable proceeds away at the defendant's club. Although the monies paid to the defendant were undoubtedly trust funds, no equitable claims were made against it. The trust funds could no longer be traced and had been spent by the defendant honestly and in ignorance of the trust. So, the solicitors brought a common law claim for money had and received and succeeded, subject to the defendant's partial defence of change of position.

The case is difficult because the House of Lords held that Cass, as a partner with authority to draw on the trust account, had obtained legal title to the money withdrawn.[93] How then was the defendant enriched at the solicitors' expense if Cass owned the money he paid to the defendant? The case becomes easier to understand if seen as a claim by trustees to recover the value of trust assets paid to the defendant in breach of trust. While Cass was the legal owner of the money he paid to the defendant, he was a trustee of the money paid. It cannot make a difference that he paid the traceable proceeds of money withdrawn from the trust account rather than paying directly from that account. His fellow trustees had a common law right to restitution of the value of that money (and an equitable duty to realise that claim).

There are three different potential claims against a recipient of assets transferred in breach of trust: (i) an equitable property claim to the assets or their traceable proceeds (which might be regarded as restitution of

---

91. Marcus Smith, "*Locus Standi* and the Enforcement of Legal Claims by *Cestuis Que* Trust and Assignees" (2008) 22 Trust Law International 140 at 156; *Sharpe v San Paulo Railway Co* (1873), LR 8 Ch App 597 (Eng) at 609-10; *Vandepitte v Preferred Accident Insurance Corp of New York*, [1933] AC 70 (PC (Canada)) at 79.
92. [1991] 2 AC 548 (HL).
93. *Ibid* at 573.

unjust enrichment);[94] (ii) an equitable personal claim for knowing receipt (which might be regarded as the equitable equivalent to damages for breach of duty); and (iii) a common law personal claim for restitution of unjust enrichment. The first two belong to the beneficiaries and the third to the trustees. There is no need to give the beneficiaries a new equitable claim for restitution of unjust enrichment, since they can compel the trustees to assert their common law claim.

It has been suggested that the right to restitution of unjust enrichment will render the action for knowing receipt "*otiose*"[95] or "irrelevant".[96] However, they operate by different rules to achieve different goals. Knowing receipt does not require benefit to the recipient and unjust enrichment does not require wrongdoing.

X. Constructive Trusteeship

Very little has been said so far about knowing or dishonest assistance of a breach of trust or fiduciary duty. They have long been linked in our minds by Lord Selborne's famous statement in *Barnes v Addy*:[97]

> [t]hose who create a trust clothe the trustee with a legal power and control over the trust property, imposing on him a corresponding responsibility. That responsibility may no doubt be extended in equity to others who are not properly trustees, if they are found either making themselves trustees *de son tort*, or actually participating in any fraudulent conduct of the trustee to the injury of the *cestui que trust*. But, on the other hand, strangers are not to be made constructive trustees merely because they act as the agents of trustees in transactions within their legal powers, transactions, perhaps of which a Court of Equity may disapprove, unless those agents receive and become chargeable with some part of the trust property, or unless they assist with knowledge in a dishonest and fraudulent design on the part of the trustees.[98]

It is true that both forms of liability may arise in the same case, but it has been unhelpful and perhaps a source of confusion to refer to recipients

---

94. Lionel Smith, "Unjust Enrichment, Property and the Structure of Trusts" (2000) 116:3 Law Quarterly Review 412 at 435.
95. *Ibid* at 413.
96. Gardner, "Taking Stock", *supra* note 90 at 86; Gardner, *supra* note 9 at 24.
97. (1874), LR 9 Ch App 244 (Eng).
98. *Ibid* at 251-52.

and assistants both as constructive trustees or to treat their liabilities as merely two limbs of the same tree.

Knowing assistants are liable as constructive trustees. Receipt of trust assets is not a condition of liability. They are not actual trustees, but by dishonestly participating in a breach of trust, they become subject to the same personal liabilities as if they were, including the liability to give up any profits derived from their wrongdoing.[99] In this context, "constructive trustee" does not mean that knowing assistants are trustees of constructive trusts, but rather that they are only constructively trustees. As Mitchell and Watterson say:

> [a] dishonest assistant is liable for his own wrongdoing, no less than a person who commits the tort of procuring a breach of contract, but at the same time, dishonest assistance is a 'secondary' wrong in the sense that it is defined by reference to the commission of a wrong by another person.[100]

In contrast, knowing recipients are actual trustees, and they are liable for breach of their own duties as trustees. The language of constructive trusteeship is unhelpful. We are perhaps reluctant to drop the label "constructive" because the recipients have not been expressly appointed to that office. However, people can become trustees in a variety of different ways. Whether a trust is express, constructive, resulting, or statutory, its trustee is a real trustee, and on becoming aware of the trust, is expected to perform the minimum duties expected of all trustees. A trustee must preserve the trust assets, and when required, transfer them to the proper parties.

Another important difference is that it is possible to knowingly assist a breach of fiduciary duty, even in the absence of any trust.[101] In contrast, liability for knowing receipt requires the receipt of trust assets. This important difference has been long overlooked in cases involving the

---

99. *Consul Development Pty Ltd v DPC Estates Pty Ltd*, [1975] HCA 8 at para 5 [*Consul*]; *Novoship (UK) Ltd v Nikitin*, [2014] EWCA Civ 908 at para 82; Mysty Clapton, "Gain-Based Remedies for Knowing Assistance: Ensuring Assistants do Not Profit from their Wrongs" (2008) 45:4 Alberta Law Review 989 at 1015; Pauline Ridge, "Justifying the Remedies for Dishonest Assistance" (2008) 124:3 Law Quarterly Review 445 at 445.
100. *Supra* note 7 at 152.
101. *Consul*, *supra* note 99 at para 25; *Enbridge*, *supra* note 48 at para 57.

misappropriation of company assets, but as discussed above, the recipient of non-trust assets may hold them subject to a new trust arising on or after receipt, in which case a claim for knowing receipt becomes possible.

In *Williams v Central Bank of Nigeria*[102] ("*Williams*"), the UK Supreme Court rejected an argument that there was a significant difference between knowing assistance and knowing receipt for the purposes of the *Limitation Act 1980*.[103] In 2010, Williams claimed that $6 million had been misappropriated from an express trust and paid to the defendant bank in 1986, and that the defendant was a party to the fraud. The court held that the claim was barred by the six-year limit on "an action by a beneficiary to recover trust property or in respect of any breach of trust" and did not fall within the exception for "an action … in respect of any fraud or fraudulent breach of trust to which the trustee was a party or privy".[104]

Lord Sumption decided that, like knowing assistants, knowing recipients are not "true trustees":

> the essence of a liability to account on the footing of knowing receipt is that the defendant has accepted trust assets knowing that they were transferred to him in breach of trust and that he had no right to receive them. His possession is therefore at all times wrongful and adverse to the rights of both the true trustees and the beneficiaries. No trust has been reposed in him. He does not have the powers or duties of a trustee, for example with regard to investment or management. His sole obligation of any practical significance is to restore the assets immediately. … There may also, in some circumstances, be a proprietary claim. But all this is simply the measure of the remedy. It does not make him a trustee or bring him within the provisions of the Limitation Act 1980 relating to trustees.[105]

The distinction between "true trustees" and other trustees is difficult. It is easy to understand why knowing assistants are not true trustees, since they do not hold assets in trusts and are only being treated as if they

---

102. [2014] UKSC 10.
103. *Limitation Act 1980* (UK), 1980, c 58.
104. *Ibid,* s 21. Many Canadian limitation statutes use similar language: see Albert Howard Oosterhoff, Robert Chambers & Mitchell McInnes, *Oosterhoff on Trusts: Text, Commentary and Materials*, 8d (Toronto: Carswell, 2014) at para 17.3.3.
105. *Supra* note 102 at para 31.

were trustees for a limited purpose. However, knowing recipients are actual trustees with trust duties to preserve the trust assets and restore them to the proper persons. It is true that they do not have the usual powers and duties of investment and management, but that is also true of a great many express trustees. For example, in a typical conveyancing transaction, the lender will advance the mortgage proceeds to a solicitor to hold on bare trust for the lender with no power to use the money at all until certain conditions are fulfilled.[106] The typical *Quistclose* trust is a bare trust for the lender coupled with a power to use the trust money for an agreed purpose,[107] and custodian trustees might simply hold the trust assets to the order of the managing trustees. While these bare trustees have none of the usual powers or duties of many express trustees to invest and manage the trust assets, it is hard to imagine that they are not "true trustees" for limitation purposes.

Knowing assistants are liable for their involvement as an accessory to a breach of trust by the express trustee and so the cause of action against them accrues from the date of that breach. In contrast, knowing recipients are not liable for the express trustees' breach, but are liable for their own breach of trust, which occurs later when they become aware that they are holding assets under a bare trust to preserve and restore them to the proper persons and then decide to use those assets contrary to the terms of that trust. It looks like a "fraudulent breach of trust to which the trustee was a party",[108] but that is not to say that *Williams* was wrongly decided. There may be good reasons for imposing a six-year limitation period on claims against knowing recipients, but it is difficult to justify that on the basis that they are not true trustees.

XI. Conclusion

Liability for knowing receipt is liability for breach of trust, pure and simple. It has taken me a long time to discover this basic truth. Professor Lionel Smith began to see the light long before I did. He questioned the

106. See *AIB Group (UK) v Mark Redler & Co*, [2014] UKSC 58 at para 4.
107. See *Twinsectra*, *supra* note 39.
108. *Supra* note 103, s 21.

momentum in favour of an explanation based on unjust enrichment, and saw that liability for knowing receipt was based on wrongdoing. In 1998, he wrote, "[i]t appears that the best view of knowing receipt is that it is equity's analogue to the common law's claim in conversion".[109]

I found the analogy to conversion difficult, because it did not explain why knowledge or notice should be required for liability in equity. Liability based on notice seemed the worst of all possible worlds. Honest, well-meaning people can be caught by notice of things they might have discovered with more care. It does not provide a sufficient reason to make them personally liable if they have received no benefit from use of the trust assets, and if they are enriched, there seems no good reason to require it.

The better analogy is to bailment. The recipient, as trustee, is a custodian of the trust assets, with the duties to preserve and restore them to the proper persons. By making this plain, Mitchell and Watterson have helped us make a great stride forward in this area of law.

So why is this the end of knowing receipt? Because there is nothing special about it. All trustees, upon learning of the trust, have duties to preserve the trust assets and account to the beneficiaries. The knowing recipient is, like every other trustee, subject to those basic trust duties. Personal liability for breach of trust depends on knowledge of the existence of the trust, but that is true of all trusts, whether express, resulting, or constructive.

In *Westdeutsche Landesbank Girozentrale v Islington London Borough Council*,[110] Lord Browne-Wilkinson suggested that there is no trust until the trustee becomes aware of it:

> [s]ince the equitable jurisdiction to enforce trusts depends upon the conscience of the holder of the legal interest being affected, he cannot be a trustee of the property if and so long as he is ignorant of the facts alleged to affect his conscience, *i.e.* until he is aware that he is intended to hold the property for the benefit of others in the case of an express or implied trust, or, in the case of a constructive trust, of the factors which are alleged to affect his conscience.[111]

---

109. Lionel Smith, "W(h)ither Knowing Receipt" (1998) 114:3 Law Quarterly Review 394 at 394. See also Smith, *supra* note 94.
110. [1996] AC 669 (HL).
111. *Ibid* at 705.

However (as I have argued elsewhere),[112] the trustee's ignorance of the trust should not prevent it from arising, but will preclude personal liability for breach of the trust. This would mean that a trust does not cease to exist just because the trust assets are misdirected to an honest recipient who is unaware of the trust. On either view, there is nothing special or different about the knowledge required for knowing receipt that sets that liability apart from the personal liability of any other trustee or connects it to the liability of the knowing assistant.

Similarly, there is nothing special about receipt in this context. All trusts require subject matter and no one becomes an actual trustee until he or she receives the trust assets. When concerned with the initial creation of an express trust, we usually refer to this as the constitution of the trust.

The generality of the concept of receipt has the potential to mislead. It appears to many to be akin to, or even an instantiation of, the enrichment required to trigger liability for unjust enrichment. However, despite frequent assertions to the contrary, there is no good reason why liability should depend on the benefit to the recipient. Also, we routinely receive assets that are not held in trust. So, the mere receipt of misdirected assets is not sufficient unless they are held in trust by the recipient when knowledge is acquired. A successful defence of *bona fide* purchase or indefeasibility of registered title negates that possibility. The receipt of assets not previously held in trust cannot trigger liability for knowing receipt unless a trust arises upon or after receipt.

A person can become liable for knowing receipt of non-trust assets that were misappropriated from a company (or other principal) by an agent acting in breach of fiduciary duty, but only if a new constructive or resulting trust arises when the assets are received or perhaps later upon rescission of the transaction. Knowledge of the breach of fiduciary duty is required, not because it triggers some special form of liability, but because it establishes knowledge of the trust or of the facts that gave rise to it. An understanding of the law of trusts is not required for someone to know

112. Robert Chambers, "Distrust: Our Fear of Trusts in the Commercial World" (2010) 63 Current Legal Problems 631 at 645-49.

that the receipt of misappropriated assets gives rise to duties to preserve and return those assets to their rightful owner. In much the same way, the duties imposed on a finder of lost goods do not depend on a working knowledge of the law of bailment.

Since liability for knowing receipt of non-trust assets transferred in breach of fiduciary duty is not some special form of accessory liability, but merely the ordinary liability for breach of the trust that arises on or after receipt, there is no good reason to confine it to cases involving a breach of fiduciary duty. The receipt of assets misappropriated by theft or fraud should also suffice, provided that a trust arises and the recipient has knowledge of the misappropriation. In *Evans v European Bank Ltd*,[113] the New South Wales Court of Appeal held that the traceable proceeds of a credit-card fraud were held by the recipient on resulting trust for the victims. This followed the *obiter dictum* of Millett J (as he then was) in *El Ajou v Dollar Land Holdings*.[114] In that case, the fraud on the plaintiff was perpetrated through the bribery of his agent. Millett J also considered the plight of the other victims:

> [t]he plaintiff's fiduciary … committed a gross breach of his fiduciary obligations to the plaintiff, and that is sufficient to enable the plaintiff to invoke the assistance of equity. Other victims, however, were less fortunate. They employed no fiduciary. They were simply swindled. No breach of any fiduciary obligation was involved. It would, of course, be an intolerable reproach to our system of jurisprudence if the plaintiff were the only victim who could trace and recover his money. Neither party before me suggested that this is the case; and I agree with them. But if the other victims of the fraud can trace their money in equity it must be because, having been induced to purchase the shares by false and fraudulent misrepresentations, they are entitled to rescind the transaction and revest the equitable title to the purchase money in themselves … There is thus no distinction between their case and the plaintiff's. They can rescind the purchases for fraud, and he for the bribery of his agent; and each can then invoke the assistance of equity to follow property of which he is the equitable owner. But, if this is correct, as I think it is, then the trust which is operating in these cases is not some new model remedial constructive trust, but an old-fashioned institutional resulting trust. This may be of relevance in relation to the degree of knowledge required on the part of a subsequent recipient to make him liable.[115]

113. *Supra* note 20.
114. [1993] 3 All ER 717 (Ch).
115. *Ibid* at 712-13.

The receipt of the traceable proceeds of theft or fraud should give rise to a trust for the victims, even in the absence of a breach of fiduciary duty.[116] The recipient's knowledge of that theft or fraud should be sufficient to trigger personal liability for breach of her or his duties as trustee to preserve the assets and restore them to the victims.

Lawyers and judges will continue to use the familiar language of knowing receipt. Old habits are hard to break and not all habits are bad. Specialist terminology is useful and efficient, but only if those who use it both understand and agree on its meaning. If we do continue to use that language, we need to understand that we are simply asking whether the defendant is personally liable for breach of trust. The answer to that question depends on whether the defendant: (a) held assets in trust; (b) had knowledge of the trust or the circumstances giving rise to it; and then, (c) failed to perform the duties to preserve the trust assets and transfer them to the beneficiaries or to the trustees who were properly appointed, willing, and able to perform the trust.

116. Robert Chambers, "Trust and Theft" in Elise Bant & Matthew Harding, eds, *Exploring Private Law* (Cambridge: Cambridge University Press, 2010) 223 .

The Function (or Malfunction) of Equity in the Charity Law of Canada's Federal Courts

Kathryn Chan*

This essay explores what, if anything, it means for the Federal Court of Appeal to be a "court of equity" in the exercise of its jurisdiction over matters related to charitable registration under the Income Tax Act. The equitable jurisdiction over charities encompasses a number of curative principles, which the Court of Chancery traditionally invoked to save indefinite or otherwise defective charitable gifts. The author identifies some of these equitable principles and contemplates how their invocation might have altered the course of certain unsuccessful charitable registration appeals. She then considers the principal arguments for and against the Federal Court of Appeal applying these equitable principles when adjudicating matters related to registered charity status.

* Assistant Professor, Faculty of Law, University of Victoria. I thank Peter Broder, Mark Gillen and Blake Bromley for their helpful comments, and Ratib Islam, Roark Lewis and Spencer Chang for their excellent research assistance.

I. Introduction

The doctrines and remedies that make up the common law charities tradition belong to the realm of equity. Inherited from the ecclesiastical courts, they were, by the fifteenth century, being applied by the Lord Chancellor of England, who granted to charitable gifts the same privileges that canon law had anciently awarded to legacies *ad pias causas*.[1] Throughout the sixteenth, seventeenth, and eighteenth centuries, the English Court of Chancery developed the law of charities, determining the boundaries of the legal concept of charity, confirming its privileges, and expanding the tools for its enforcement. By the early nineteenth century, when the English law of equity was being introduced into the colonies of British North America[2], there was in place an extensive set of equitable doctrines relating to charitable trusts, charitable corporations and charitable gifts.

Today, in Canada, it is arguable that the most significant heir of this equitable jurisdiction is the appellate court of "law, equity, and

1. Gareth H Jones, *History of the Law of Charity, 1532-1827* (London: Cambridge University Press, 1969) at 4-9.
2. JE Cote, 'The Reception of English Law' (1977) 15 Alberta Law Review 29 at 57-59.

admiralty" created by the *Federal Courts Act*[3] ("*FCA*"). Pursuant to section 172(3) of the *Income Tax Act*[4] ("*ITA*"), a decision of the Minister of National Revenue to exclude an organization from registered charity status is appealed directly to the Federal Court of Appeal. The Federal Court of Appeal is therefore the first line judicial arbiter of whether an applicant is "constituted and operated exclusively for charitable purposes" (*constituée ou administrée exclusivement à des fins de bienfaisance*), or is an "organization…which devotes all of its resources to charitable activities" (*œuvre…dont la totalité des ressources est consacrée à des activités de bienfaisance*) so as to qualify it for charitable registration.[5] Due to the constraints on appealing civil matters to the Supreme Court of Canada, the Federal Court of Appeal also generally has the last word on whether a purpose or activity is "charitable". For almost twenty years, that last word has consistently been "no".[6]

It is the Canadian judiciary's longstanding approach to the interpretation of the registered charity provisions that links the Federal Court of Appeal to the charity law tradition of the Chancery Courts. The *ITA* does not define the word "charitable" (*de bienfaisance*). In this situation, the Federal Court of Appeal has always relied on the common law to give meaning to the statutory term. The Supreme Court of Canada confirmed the propriety of this approach in *Vancouver Society of*

3. RSC 1970, c 10 [*FCA*].
4. RSC 1985, c 1 (5th Supp) [*ITA*].
5. *Ibid* s 149.1(1), "charitable foundation" (*fondation de bienfaisance*); "charitable organisation" (*oeuvre de bienfaisance*). The Department of Finance has faced considerable pressure to introduce amendments that would allow charity registration and revocation appeals to be taken to the Tax Court of Canada. So far, however, the only response has been the extension of the internal objection review process of the Minister of National Revenue: see William Innes & Patrick J Boyle, *Charities, Non-Profits and Philanthropy Under the Income Tax Act* (Toronto: CCH Canadian, 2006) at 13.
6. The last time a charity successfully appealed a charitable registration decision was in 1996: see *Vancouver Regional FreeNet Assn v Minister of National Revenue*, [1996] 3 FC 880 (CA) [*Vancouver Regional FreeNet*]. Since 1999, there have been at least 17 unsuccessful appeals by charities to the Federal Court of Appeal.

Immigrant and Visible Minority Women v Minister of National Revenue[7] ("*Vancouver Society*"), stating that the *ITA*:

> … appears clearly to envisage a resort to the common law for a definition of 'charity' in its legal sense as well as for the principles that should guide us in applying that definition …[8]

and that:

> … whether a purpose would or may operate for the public benefit is to be answered by the court on the basis of the record before it and in exercise of its equitable jurisdiction in matters of charity.[9]

In the years before and since the *Vancouver Society* decision, there has been extensive debate on the merits of resorting to the common law definition of charity for purposes of interpreting federal tax legislation. However, considerably less thought has been given to the principles that should guide the Federal Court in applying that definition, or to the nature of the equitable jurisdiction that a court must exercise in the context of the registered charity regime.

This essay explores what, if anything, it means for the Federal Court of Appeal to be a "court of equity" in the exercise of its jurisdiction over matters related to registered charity status under the *ITA*. Its premise is that the equitable jurisdiction over charities encompasses a number of curative principles which traditionally functioned to effectuate indefinite or otherwise defective gifts and which might usefully be invoked in an appropriate case to lift the law of (registered) charities out of its present state of inertia. The essay begins by describing the equitable jurisdiction that is conferred on the Federal Court of Appeal by the terms of the *FCA* (Part II). It then identifies several key principles of the equitable jurisdiction over charities, and explores what impact those principles might have had on recent charitable registration appeals, had they been invoked (Part III). Finally, having established how equitable principles have traditionally influenced judicial decision-making regarding the charitable nature of purposes and activities, the essay turns to examine the chief arguments for and against their application by the Federal

7. [1999] 1 SCR 10 [*Vancouver Society*].
8. *Ibid* at para 28.
9. *Ibid* at para 175.

Court of Appeal (Part IV).

II. The Equitable Jurisdiction of the Federal Court of Appeal

The Federal Court of Appeal is the successor of the Appeal Division of the Federal Court of Canada, which was created in 1971 to be the successor to the Exchequer Court of Canada. Like our other federal courts, it is a court created by the Parliament of Canada under the authority of section 101 of the *Constitution Act, 1867*.[10] The jurisdiction of the Federal Court of Appeal is often described as "exceptional".[11] In contrast to the provincial superior courts, which have a general and inherent jurisdiction, it has no jurisdiction except that conferred by statute.[12] In a famous series of cases dating from 1976, the Supreme Court of Canada also limited the exercise of Canadian federal court jurisdiction to *substantive* federal law, holding that there must be an "existing body of federal law which is essential to the disposition of the case and which nourishes the statutory grant of jurisdiction".[13] However, the general trend that has characterized recent federal court jurisprudence is the adoption of an increasingly expansive approach to federal court jurisdiction.[14] It is uncontroversial that the federal courts may decide incidental questions of provincial law

10. (UK) 30 & 31 Victoria, c 3, reprinted in RSC 1985, App II, No 5.
11. Brian J Saunders, Justice Donald J Rennie & Graham Garton QC, *Federal Courts Practice 2015* (Toronto: Carswell, 2014) at 1-2.
12. *Ordon Estate v Grail*, [1998] 3 SCR 437. For a more detailed discussion, see *ibid* at 1-5.
13. *ITO-Int Terminal Operators Ltd v Miida Electronics Inc*, [1986] 1 SCR 752 at 766, also citing *Quebec North Shore Paper Co v Canadian Pacific Ltd*, [1977] 2 SCR 1054 and *McNamara Construction (Western) Ltd v The Queen*, [1977] 2 SCR 654.
14. See *e.g. Minister of National Revenue v RBC Life Insurance Co,* 2013 FCA 50, (the federal courts enjoy "plenary powers" analogous to the inherent powers of the provincial superior courts); *Canadian Transit Co v Windsor (City),* 2015 FCA 88 [*Canadian Transit*] (the federal courts have the jurisdiction to consider the constitutional doctrines of paramountcy and inter-jurisdictional immunity where they otherwise have jurisdiction).

in adjudicating income tax appeals.[15]

Section 3 of the *FCA* constitutes the Federal Court of Appeal as a court of "law, equity, and admiralty in and for Canada". Equity, it has been held, refers to those principles of law that were administered before 1873 by the Courts of Equity (and mainly the Court of Chancery), rather than any general notion of what is just or fair.[16] Thus, like the Federal Court, and notwithstanding that it is a creature of statute, the Federal Court of Appeal has the power to apply the doctrines of the Court of Chancery where the subject-matter of the dispute is otherwise within its jurisdiction and where equitable principles are applicable to the issue.[17] In *Algonquin Mercantile Corp v Dart Industries Canada Ltd*, the Federal Court Trial Division described this equitable jurisdiction in the following terms:

> [o]nce it has jurisdiction and subject only to any specific statutory provision to the contrary, the Federal Court of Canada may, in determining the issues before it, exercise all of the powers and enforce all of the remedies available to both courts of law and courts of equity. In other words, to dispose of any case before it, it may exercise the same powers and apply the same laws and principles as the Superior Court of the province where the cause of action lies.[18]

In *Apotex Inc v Abbott Laboratories Ltd*,[19] the Ontario Superior Court confirmed that it shared this expansive view of the Federal Court's jurisdiction, rejecting the argument that it could award equitable relief to a party in a patent-based action in circumstances where the Federal Court had denied other patentees exactly the same remedy. [20]

There is a relatively small body of case law shedding light on when the Federal Courts will, and when they will not, apply the rules of equity

---

15. *Canadian Transit, supra* note 14 at para 38.
16. *Maplesden v Minister of National Revenue*, [1998] 2 CTC 162 (FCA) at para 27 [*Maplesden*].
17. *Teledyne Industries Inc v Lido Industrial Products Ltd* (1982), 68 CPR (2d) 204 (FC (TD)) at 227 [*Teledyne*]; *Garford Pty v Dywidag Systems International Canada Ltd,* 2010 FC 997 (TD)[*Garford*]. The Exchequer Court of Canada also had an equitable jurisdiction.
18. [1987] 2 FC 373 (TD) [*Algonquin*] at para 78.
19. 2013 ONSC 356.
20. *Ibid* at para 81 [*Apotex*].

to a matter brought before them.[21] In perhaps the earliest case on point, *Teledyne Industries Inc v Lido Industrial Products Ltd*[22] ("*Teledyne*"), the Federal Court considered a claim that it should award prejudgment interest on equitable principles in an accounting of profits for patent infringement under the federal *Patent Act*.[23] Section 40 of the *FCA*, which governed the claim, provided: "[u]nless otherwise ordered by the Court, a judgment … bears interest from the time of giving the judgment". The Federal Court held that this section neither granted nor limited any authority that the court might otherwise possess regarding pre-judgment interest and awarded the equitable relief claimed. According to Justice Addy, the fact that the fundamental right being protected was a statutory right did not remove or vary the "normal requirement" to apply "all relevant equitable principles in determining the nature and the extent of the relief to which the aggrieved party was entitled".[24]

Since *Teledyne*, the equitable jurisdiction of the Federal Courts has manifested (or not manifested) itself in a variety of different ways. The Federal Court of Appeal has invoked its equitable jurisdiction to enforce a party's undertaking to pay damages[25] and to grant a bill of discovery against the Minister of National Revenue in a patent dispute between two corporate parties.[26] In other cases, it has refused to exercise its equitable jurisdiction to hear claims based on unjust enrichment[27] or to vacate a tax assessment on equitable grounds.[28] If there is a common theme in these decisions, it is the court's central concern to determine the

---

21. I speak here of the Federal Court and the Federal Court of Appeal. There are, of course, other "small f" federal courts, the extent of whose "equitable jurisdiction" would require separate analysis: see *e.g. Pliskow v Canada*, 2013 TCC 283.
22. *Teledyne, supra* note 17.
23. RSC 1985, c P-4.
24. *Teledyne*, *supra* note 17 at 230.
25. *Algonquin*, *supra* note 18. See also *Beloit Canada Ltée c Valmet-Dominion Inc,* [1997] 3 FC 497 (CA).
26. *Glaxo Wellcome v Canada (Minister of National Revenue)*, [1998] 4 FC 439 (CA) [*Glaxo Wellcome*].
27. *Bédard v Kellogg*, 2007 FC 516 (TD). See also *Garford, supra* note 17.
28. *Maplesden*, *supra* note 16.

interplay between the rules of equity being invoked and the legislative provisions at issue.[29] The Federal Court of Appeal has indicated that it will not shrink from applying an equitable principle simply because of its novelty in Canadian jurisprudence. [30] In cases where statutory and equitable tools overlap, however, the Court is tasked with determining "whether the legislation displaces or precludes resort to the common law or whether, conversely, the common law applies in addition to or in spite of the law set out in the legislation".[31] The legislature is generally presumed not to depart from prevailing law, unless it clearly expresses its intention to do so. However, this general presumption may be rebutted if it is clear the legislature intended to modify equitable rights by enacting a "comprehensive regulation of the matter at issue".[32] "Careful attention" must therefore be paid to the particular law under which jurisdiction is claimed.[33]

The Federal Courts' understanding of when a statutory regime precludes the exercise of their equitable jurisdiction can be illustrated by contrasting two decisions that came to opposite conclusions on the use of equitable tools. In *Garford Pty Ltd v Dywidag Systems International Canada Ltd*[34] ("*Garford*"), a plaintiff sought leave to amend a statement of claim based on section 36 of the federal *Competition Act*,[35] in order to raise ancillary claims sounding in unjust enrichment.[36] Section 36 provided that any person who had suffered loss due to breaches of the *Competition Act* was entitled to sue for "an amount equal to the loss or damage proved to have been suffered by him, together with any additional amount that the court may allow not exceeding [the cost of any investigation and the proceedings]". Finding that section 36 provided a civil remedy that

29. *Glaxo Wellcome*, *supra* note 26.
30. *Ibid* at para 33.
31. *Ibid* at para 36 citing Ruth Sullivan, *Driedger on the Construction of Statutes*, 3d (Toronto: Butterworths, 1994) at 298.
32. *Glaxo Wellcome, supra* note 26 at para 36.
33. *Garford*, *supra* note 17 at para 9.
34. *Ibid*.
35. RSC 1985, c C-34.
36. *Garford*, *supra* note 17 at para 3.

limited plaintiffs to the recovery of actual loss or damage, the Federal Court refused to grant the proposed amendment. The court reasoned that either the plaintiff was putting forth a new cause of action, in which case the matter was within provincial jurisdiction, or it was seeking a new remedy, in which case that remedy was outside the terms of section 36.[37] Section 3 of the *FCA* did not assist the plaintiff as that section did not give the Court a "general jurisdiction" to consider equitable claims and remedies in a civil case based on a statutory cause of action.[38]

Garford may be contrasted with *Glaxo Wellcome v Canada (Minister of National Revenue)*[39] ("*Glaxo Wellcome*"), a 1998 decision of the Federal Court of Appeal. In *Glaxo Wellcome*, the appellant sought to obtain from the Minister of National Revenue the names of unlicensed importers of a patented drug that the appellant believed was violating its patent rights. The federal *Customs Act*,[40] which governed the appellant's request, contained three relevant provisions:

- Section 107, which prohibited the Minister from disclosing information collected under the Act "except as authorized under section 108";
- Subsection 108(1), which permitted the Minister to disclose information gathered pursuant to the Act to the Department of National Revenue or "any person" authorized by the Minister;
- Subsection 108(2), which authorized a court of record to order a customs officer to disclose information gathered pursuant to the Act.

When the Minister refused to release the names of the importers to the appellant under subsection 108(1), the appellant petitioned the Federal Court to grant it an equitable remedy "of ancient origin", which would permit the court to order discovery of persons against whom the appellant

37. *Ibid* at para 12.
38. *Ibid* at para 8.
39. *Glaxo Wellcome*, *supra* note 26.
40. RSC 1985, c 1 (2d Supp).

had no cause of action.[41] On appeal, the Federal Court of Appeal agreed to exercise its equitable jurisdiction to provide the remedy sought. The Court reasoned that although Parliament had specifically designated the Minister as the decision-making authority regarding which persons should have access to information collected under the Act, the inclusion of section 108(2) indicated that Parliament did not intend to make the Minister the "sole arbiter of disclosure".[42] The *Customs Act*, in other words, was not "an exhaustive statutory code" which directly or by necessary implication prohibited the courts from incorporating equitable principles or remedies.[43] Based on this reading of the *Act*, the Court exercised its jurisdiction to order an equitable bill of discovery, effectively requiring the Minister to disclose the names of the importers of the patented drug.

It is arguable that, at least in some earlier cases, the Federal Courts exhibited a particular reluctance to exercise their equitable jurisdiction in matters related to tax. In *Maplesden v Minister of National Revenue*[44] ("*Maplesden*"), for example, a taxpayer asked the (then) Federal Court Trial Division to vacate a tax assessment on equitable grounds, arguing that it was Revenue Canada's own erroneous actions that had caused the tax liability to arise. The Court held that it had no authority to grant the relief sought, distinguishing *Teledyne* on the basis of the statutory grant of authority conferred on the Court in that case, and noting that "tax laws were never part" of the regime administered by the English Court of Chancery.[45] One year later,[46] the same court relied on *Maplesden* in refusing to use its equitable jurisdiction to grant restitution to taxpayers in the face of limitations imposed by the *Excise Tax Act*.[47] However, in *Neles Controls Ltd v Canada*,[48] the Federal Court of Appeal brought its position on the exercise of equitable jurisdiction in tax matters in line

41. *Glaxo Wellcome*, *supra* note 26 at para 20.
42. *Ibid* at para 42.
43. *Ibid*.
44. *Maplesden*, *supra* note 16.
45. *Ibid* at paras 27-28.
46. *Federated Co-operatives Ltd v Canada* (1999), 165 FTR 135.
47. RSC 1985, c E-15.
48. 2002 FCA 107.

with *Glaxo Wellcome*, holding that the determining factor in a claim for equitable relief of a tax overpayment was whether it was "precluded by a comprehensive statutory scheme".[49] Even in the tax context, the Court suggested, an exercise of its equitable jurisdiction could be warranted "in situations where the legislation is silent, where it cannot apply, or where a gap in relief is apparent".[50]

III. The Role of Equitable Principles in the Common Law of Charities

If we turn from the general nature of the Federal Courts' equitable jurisdiction to the specific nature of the jurisdiction that Chancery traditionally exercised over charities, we see that the latter jurisdiction had a few central features. On the one hand, equity subjected charitable trusts and corporations to a standard of control and scrutiny even more rigorous than that applied to private trusts.[51] On the other hand, equity treated charities as entitled to "extraordinary favor".[52] Charities were largely exempt from the rule against perpetuities, for example, and the Chancellor would not allow a statute of limitations to bar an action to enforce a charitable use.[53] In recognition of the public interest in charity property, the Court of Chancery also developed a number of curative principles that oriented the court towards the effectuation of charitable gifts. The following section identifies some of these curative principles, explains how they functioned, and points to various places in Canada's

49. *Ibid* at para 15.
50. *Ibid*.
51. Hubert Picarda, *The Law and Practice Relating to Charities*, 4d (West Sussex: Bloomsbury, 2010) at 528 [Picarda], citing *A-G v Gleg* (1738), Amb 584 (Ch (Eng)); *A-G v The Governors of Sherborne Grammar School* (1854), 18 Beav 256 (Ch (Eng)). See also *A-G v Governors of Harrow School* (1754), 28 ER 351 (Ch), and *Re Devlin's Estate* (1889), 23 LR Ir 516 (Ch)(where the Chancery Court required a charity trustee to swear an affidavit stating how he would exercise his discretion before charity funds were paid out of court).
52. *Jackson v Phillips* (1867), 96 Mass (14 Allen) 539 (USSC) at 550 [*Jackson*].
53. Jones, *supra* note 1 at 18.

registered charity jurisprudence where their exercise might have altered the course of a charitable registration appeal.

A. Benignant Construction

A first equitable principle that distinguishes the law of charities from the broader law of trusts is that a court may give a benignant construction to a document in order to carry into effect a donor's charitable intent. "In regard to the construction of charitable gifts … the rule of widest application is that the court leans in favour of charity".[54] Where a gift is capable of two constructions, one of which would make it void and the other effectual, the latter construction is to be preferred.[55] There are limits to the principle of benignant construction: the court, it is sometimes said, must not "strain a will" to gain money for a charity.[56] Where the terms of a gift permit the trustees to apply the trust fund to charity *or* some other non-charitable purpose, the gift may be found too uncertain to be valid. Even in such circumstances, however, courts of equity have sometimes saved an unfortunately worded gift, construing the word 'or' conjunctively on the basis of either the *ejusdem generis* principle, benignant construction, or the testator's charitable intent.[57]

The Canadian courts have occasionally applied the principle of benignant construction to save a charitable gift. A leading authority is *Jones v T. Eaton Co*[58] ("*Jones*"), a case which concerned the validity of a bequest to the Eaton Company's executive officers "to be used by them as a trust fund for any needy or deserving Toronto members of the Eaton Quarter Century Club as the said Executive Officers in their

54. Jean Warburton et al, *Tudor on Charities* 9d (London: Sweet & Maxwell, 2003) at 175 [Warburton].
55. *Inland Revenue Commissioners v McMullen*, [1979] 1 All ER 588 (CA (Civ)) citing *Houston v Burns*, [1918] AC 337 (HL) at 341-42 and *Re Bain, Public Trustee v Ross*, [1930] 1 Ch 224 (CA (Eng)) at 230. See also *Bruce v Presbytery of Deer* (1867), LR 1 SC 96 (HL) at 97.
56. Warburton, *supra* note 54 at 176, citing *Dolan v Macdermot* (1868), LR 3 Ch 676 (Eng) at 678.
57. Picarda, *supra* note 51 at 330-31.
58. [1973] SCR 635.

absolute discretion may decide".[59] Before the Supreme Court of Canada, the argument focused on whether the word "or" must be interpreted disjunctively or conjunctively, and whether the word "deserving" was too vague to communicate a charitable intent. The Court found it unnecessary to deal with the first argument, but upheld the validity of the trust on the basis that the testator was expressing a charitable intent when he used the word "deserving".[60] Citing English authority for the principle that charitable bequests must receive a benignant construction, the Court held that in the context of the bequest, the word "deserving" must be understood to refer to a person "who although not actually poverty-stricken was nevertheless in a state of financial depression".[61] On the strength of this construction, a unanimous Court concluded that the disputed bequest was a valid charitable trust.

If the Supreme Court of Canada has invoked the principle of benignant construction in construing the terms of an allegedly charitable bequest, it has been more equivocal on the role of the equitable principle in construing allegedly charitable objects for purposes of the registered charity regime. This is the conclusion that must be drawn from *Vancouver Society*, where the Court was called upon to determine whether the purposes and activities of the Vancouver Society of Immigrant and Visible Minority Women were charitable so as to qualify the organization for charitable registration under the *ITA*. The objects of the Society included, *inter alia*:

> (a) to provide educational forums, classes, workshops and seminars to immigrant women in order that they may be able to find or obtain employment or self-employment ...; and
>
> (e) to provide services and to do all such things that are incidental or conducive to the attainment of the above stated objects, including the seeking of funds from governments and/or other sources for the implementation of the aforementioned objectives.[62]

The Supreme Court of Canada held that object (a) was charitable, a

59. *Ibid* at 637.
60. *Ibid* at 642.
61. *Ibid* at 646.
62. *Vancouver Society*, *supra* note 7 at paras 129-35.

conclusion that was only made possible by the Court's bold expansion of the educational head of charity to include more informal training initiatives aimed at teaching necessary and practical life skills.[63] However, the appellant's case fell on object (e), with a majority of the Court concluding that the inclusion of the words "or conducive" placed the Society as a whole "outside the scope of legal charity".[64]

Vancouver Society provides a dramatic illustration of the potential for the rules of equity to influence the outcome of charitable registration appeals. The Supreme Court of Canada's majority and dissenting judgments ran, together, for over 200 paragraphs, addressing a number of difficult and substantive questions regarding the scope of the registered charities regime. Nevertheless, the issue on which the majority and dissenting judges ultimately split was the proper construction of the Society's final object. Justice Gonthier, while not explicitly invoking the principle of benignant construction, appeared to rely on it in spirit. Writing in dissent, he found that the "obvious intent" of the drafter of object (e) was to provide a mechanism by which the Society's main (and charitable) purpose could be achieved.[65] Moreover, Gonthier J stated, based on the Court's decision in *Jones,* it would be erroneous to simply assume that the term "or" should be interpreted disjunctively in construing an allegedly charitable object.[66] Finally, Gonthier J noted that in an earlier case, the Supreme Court of Canada had affirmed the validity of an association constituted to do things "incidental or conducive" to the attainment of its charitable objects.[67] The conclusion that begged to be drawn was that the language raised no concerns.[68]

The majority judges evaluated the Society's final object through a far less benignant lens. Writing for the majority, Justice Iacobucci opined that while doing things "incidental" to the attainment of charitable purposes could safely be treated as a means of fulfillment of those

63. *Ibid* at paras 166-74.
64. *Ibid* at para 116.
65. *Ibid* at para 118.
66. *Ibid* at para 119.
67. *Ibid* at para 120.
68. *Ibid* at paras 117-21.

purposes, the term "conducive" implied only that the action *contributed* to that result.[69] Based on this reading of object (e), and bolstered by his view that the Society had previously carried out non-charitable activities that would be covered by the word "conducive", Iacobucci J concluded that the Society could not be classified as a charitable organization.[70] It was an odd end to the *Vancouver Society* story, particularly as no party had raised an objection to object (e) in either written or oral argument.[71] As the decision stands, however, the majority's actual construction of object (e) sits in tension with its earlier endorsement of relying on equitable principles to determine charitable status under the *ITA*. This tension goes some way in explaining why the Federal Court of Appeal has not mentioned the principle of benignant construction in the twenty-odd charitable registration appeals it has decided since.[72]

B. Presumption of Lawful Trustee Behaviour

A second equitable principle that tends towards the effectuation of charitable gifts is closely related to the first. Where the terms of an

---

69. *Ibid* at para 193.
70. *Ibid* at para 195.
71. See *ibid* at para 116, per Gonthier J. The Supreme Court of Canada has since confirmed that an appellate court may only raise new issues on its own initiative when failing to do so would risk an injustice: see *R v Mian*, 2014 SCC 54 at para 41; *Kahkewistahaw First Nation v Taypotat*, 2015 SCC 30 at para 26.
72. The registered charities jurisprudence contains a singular, passing reference to the principle of benignant construction. In *Vancouver Regional FreeNet, supra* note 6, Justice Décary in dissent described the majority of the Federal Court of Appeal as having "applied a benignant construction" in construing the Native Communications Society's "deficient constituting document" as charitable at para 42. It does seem likely that the Court implicitly relied on the equitable principle in *Native Communications*, one of the few successful charitable registration appeals. In recent years, however, the Federal Court of Appeal has neither relied upon nor rejected the equitable principle. The Supreme Court of Canada did not address benignant construction in the *AYSA Amateur Youth Soccer Association v Canada (Revenue Agency)*, 2007 SCC 42 [*AYSA*] decision, but arguably construed the appellant's objects strictly at para 41.

otherwise charitable trust allow for the possibility of improper action, equity allows the court to presume that the trustees will act in a lawful and appropriate way.[73] In *McGovern v A-G*,[74] Justice Slade applied this presumption in support of his conclusion, in *obiter dicta*, that a trust for the undertaking and dissemination of research into the maintenance and observance of human rights would be a valid charitable trust.[75] The "mere theoretical possibility" that trustees might implement these objects in a political manner would not render them non-charitable, Slade J held, as the clauses would be entitled to a benignant construction and "to the presumption … that the trustees would only act in a lawful and proper manner appropriate to the trustees of a charity and not, for example, by the propagation of tendentious political opinions".[76] In *Re Koeppler's Will Trusts*,[77] the English Court of Appeal relied on this presumption of lawful trustee behaviour in upholding a gift for the furtherance of potentially political work.[78]

Equity's presumption of lawful trustee behavior is not irrefutable; in cases where it is unclear whether an organization's objects are exclusively charitable and trustees have acted pursuant to those purposes, extrinsic evidence of those acts is admissible for purposes of clarifying the charitable nature of the enterprise.[79] However, the Federal Court of Appeal has not allowed applicants for registered charity status to benefit from the presumption of lawful trustee behaviour, even absent any

---

73. Warburton, *supra* note 54 at 175-76; Peter Luxton, *The Law of Charities* (Oxford: Oxford University Press, 2001) at 205 [Luxton].
74. *McGovern v A-G*, [1982] Ch 321 (CA)(Eng)).
75. *Ibid*.
76. *Ibid* at 353, 346 (the court should infer that trustees will act lawfully and only use means appropriate to the trustees of the charity). See also, *Southwood and Parsons v AG* [1998] EWHC Ch 297 at para 15.
77. [1986] Ch 423 (CA (Civ)(Eng)).
78. *Ibid* at 438. The gift was for the furtherance of the Wilton Park project which organized and conducted conferences on a wide variety of topics, some of which, the Court acknowledged, had a political flavour.
79. *Guaranty Trust Co of Canada v Minister of National Revenue*, [1967] SCR 133. See also, Warburton, *supra* note 54.

evidence of improper acts. In *Travel Just v Canada (Revenue Agency)*,[80] for example, the Federal Court of Appeal considered the appeal of a not-for-profit corporation that had not carried out any activities or even accrued any funds before unsuccessfully applying to the Minister for charitable registration.[81] The objects of the organization included the creation and development of "model tourism development projects that contribute to the realization of international human rights and environmental norms".[82] In rejecting this object as excessively broad and vague, the Court stated that the development of such model tourism development projects *could* include "the financing and operation of luxury holiday resorts in developing countries".[83] The Court implied that this would be an unlawful or at least inappropriate usage of charity funds, but did not address the appellant's argument that it was entitled to the presumption that its trustees would not use the funds in this way.[84] In the result, despite the absence of any evidence of improper trustee behavior, the organization's appeal was dismissed.

C. Resolution of Technical Defects

A third equitable principle that tends towards the effectuation of charitable gifts authorizes the court to perfect an otherwise charitable transfer that suffers from a technical defect.[85] As Lord Eldon stated in *Moggridge v Thackwell*[86] ("*Moggridge*"):

> if the testator has manifested a general intention to give to charity, the failure of the particular mode in which the charity is to be effectuated shall not destroy the charity: but, if the substantial intention is charity, the law will substitute

80. 2006 FCA 343 [*Travel Just*].
81. *Ibid*.
82. *Ibid* at para 9.
83. *Ibid* at para 6.
84. *Travel Just,* Factum of the Appellant at para 39. I was involved in the preparation of the appellant's factum.
85. Jones, *supra* note 1 at 60-68, 80. It was originally thought that this jurisdiction flowed from the preamble of the 1601 Statute, but by the late 17th century it was accepted that imperfect transfers could be perfected in Chancery, even if the proceedings were brought by bill or information.
86. (1803), 7 Ves 36 (Ch (Eng)).

> another mode of devoting the property to charitable purposes, though the formal intention as to the mode cannot be accomplished.[87]

A court of equity may intervene to perfect an otherwise charitable transfer where a donor's directions are indefinite, ambiguous, or insufficient. It may also intervene where the prescribed manner of carrying out the donor's general charitable intention is *illegal*.[88] In *Re Bradwell*,[89] for example, the Court of Chancery considered a contested bequest with a direction to accumulate income and apply the proceeds of the accumulation to the benefit of Wesleyan ministers. The direction was in plain violation of the governing property law statute. Nevertheless, the Court of Chancery rejected the argument that the bequest was invalid. As the testator had expressed his intention that his estate "be, in substance, applied for charitable purposes," the Court held that the fact that he had chosen an illegal mode for effecting those purposes should not undermine the bequest.[90] The function of a court of equity in such circumstances was not to declare the gift invalid, but rather to carry the testator's charitable intention into effect through an administrative or *cy-près* scheme.

Canadian case law includes several examples of superior courts exercising the curative power described in *Moggridge*. Most often it has been invoked to effectuate indefinite or ambiguous gifts, such as a bequest "to the work of the Lord".[91] In recent years, however, at least one provincial superior court has relied on its scheme-making powers to bring a charitable trust into conformity with the rules of the registered charity regime. In *Toronto Aged Men's and Women's Homes v Loyal True Blue and*

---

87. *Ibid* at 69.
88. A gift will fail if the general intention is illegal or contrary to public policy. However, if it is evident that the donor had a general charitable intention, but the prescribed manner of carrying of the intention is illegal, the court will execute the gift cy-pres; see Picarda, *supra* note 51 at 450-51.
89. [1952] Ch 575 (Eng)) [*Bradwell*].
90. *Ibid* at 579-80.
91. See *e.g. Re Brooks Estate* (1969), 4 DLR (3d) 694 (Sask QB). See also *Re Leslie*, [1940] OWN 345 (SC) and *Phelps v Lord,* [1894] 25 OR 259.

Orange Home,[92] the testatrix had created an endowment under which the trustees were to retain the capital of the residue, while distributing the income to charitable beneficiaries.[93] The trust was subsequently registered as a private foundation under the *ITA*, which placed upon the trust a requirement to disburse annually an amount equal to at least 4.5% of the fair market value of its property.[94] Since the trustees had no power to encroach on the capital of the residue, they were in a difficult position: compliance with the *ITA* would require them to breach their trust, while a failure to comply would entitle the Minister of Customs and Revenue to revoke the trust's charitable registration and render it liable to a ruinous revocation tax. In these circumstances, the court had "no real concern" with exercising its equitable jurisdiction to declare the trust "impracticable" and to order a *cy-près* scheme that expanded the investment and distribution powers of the trustees.[95]

Assuming that the Federal Court of Appeal has all the same equitable powers as the superior courts with respect to disputes within its jurisdiction, it has never followed the Ontario Superior Court in exercising a scheme-making power to bring an imperfect applicant for registered charity status into conformity with the *ITA*. One applicant for registered charity status that might have benefited from such an exercise of the Court's equitable jurisdiction is the Fuaran Foundation, a religious not-for-profit society that appealed a negative registration decision in 2004.[96] The Minister of National Revenue had refused to register the Foundation as a charitable organization on the ground that its objects permitted the distribution of resources to persons who were not "qualified donees" under the *ITA*. The constitution of the Foundation provided, in relevant part:

> [t]he objects and purposes of the Foundation are to advance the Christian Religion and to advance education by undertaking programs and projects in pursuit of its purposes as are exclusively charitable at law by … (a) providing financial assistance for the establishment and continued support of individual

92. (2003), 68 OR (3d) 777 (SupCtJ).
93. *Ibid*.
94. *Ibid* at para 14.
95. *Ibid* at para 15.
96. *Fuaran Foundation v Canada (Customs and Revenue Agency)*, 2004 FCA 181 [*Fuaran Foundation*].

> Christians and Christian organizations engaged worldwide in [Christian teaching and poverty relief.][97]

In presenting its case for charitable registration, the society submitted that its constitution "clearly and expressly" limited the discretion of its trustees to the carrying out of projects that were exclusively charitable at law.[98] The Fuaran Foundation also invoked the principle of benignant construction and offered to specifically undertake that its directors would only donate to "qualified donees" as the *ITA* required.[99] Nevertheless, the Federal Court of Appeal upheld the Minister's refusal to register the society on the ground that the language of the constitution was broad enough to allow the trustees to undertake "non-charitable activities" under the *ITA*, including the provision of financial assistance to non-qualified donees.[100] Neither the parties nor the court appeared to consider whether the Federal Court of Appeal might invoke the scheme-making powers of the Court of Chancery to affirm the validity of the charity and to cure its illegal mode.[101]

IV. Is it Appropriate to Invoke Equitable Doctrines in the Application of the Registered Charity Regime?

In Part II of this work, I argued that, subject to any specific statutory provision to the contrary, the Federal Court of Appeal has the authority to exercise all of the equitable powers of the Court of Chancery in determining matters properly before it. In Part III, I described some of the specific doctrines and powers that characterized the Chancery jurisdiction over charities. I identified places in the registered charity jurisprudence where those tools were conspicuously absent and explored how those absences contributed to the negative resolution of each charitable registration appeal. The question that remains is whether the Federal Court of Appeal

97. *Ibid* at para 6.
98. *Fuaran Foundation,* Factum of the Appellant at paras 45 and 53.
99. *Ibid* at para 10.
100. *Fuaran Foundation, supra* note 96 at para 6.
101. *Ibid* at para 11.

should actually apply these curative doctrines in determining whether an organization's purposes and activities are exclusively charitable under the *ITA*. Despite *Vancouver Society*'s suggestion that a court's determination of registered charity status is an exercise of its equitable jurisdiction in matters of charity, such application would clearly be a departure from the *status quo.* In the final part of this piece, I address some of the arguments in favour of having the Federal Court of Appeal act as a court of equity in the adjudication of charitable registration appeals, as well as some of the possible objections to that position.

A. The "Moving" Nature of Charity Law

For those concerned about the Canadian law of charities having long since ceased to be the "moving subject" described by Lord Wilberforce,[102] there are several compelling arguments for having the Federal Court of Appeal apply the traditional doctrines of the Court of Chancery in the adjudication of charitable registration appeals. A first is that it might give applicants for registered charity status a fighting chance in challenging the decisions of the Minister. With no not-for-profit organization having won a charitable registration appeal in almost twenty years, and even *Vancouver Society* having been lost on the basis of the strict construction of an ancillary object, unsuccessful applicants for registered charity status are today generally being advised to surrender or reconstitute themselves, rather than appeal even questionable revenue decisions. This trend is evidenced by the steady decrease in the number of charitable registration appeals that have been brought in recent years.[103]

Less litigation is not, of course, problematic in itself. However, the cumulative result of the dramatic record of losses at the Federal Court of Appeal has arguably been the near eradication, in Canada, of the common law method of developing the legal definition of charity by judicial

102. *Scottish Burial Reform and Cremation Society Ltd v Glasgow Corporation*, [1968] AC 138 (HL).

103. From 1999 to 2008, 13 different organizations appealed negative registration decisions to the Federal Court – an average of more than one appeal per year. Betweem 2009 and the time of writing, there had only been four charitable registration appeals.

analogy. The recent charitable registration decisions of the Federal Court of Appeal can scarcely claim to be charity law decisions at all, so brief are their reasons and so fleeting their allusions to the appellants' disputed objects.[104] For a jurisdiction with no statutory definition of charity, the effect of this trend is to tether the concept of charity ever more rigidly to a dated body of English case law. In addition to causing frustration for not-for-profit litigants, the disappearance of the common law analogical method is at odds with the apparent design of the registered charity regime. In *Vancouver Society,* Gonthier J described that regime as a signal of Parliament's acceptance that the courts "have a continuing role to play" in keeping the definition of charity in tune with social and economic developments.[105] If we accept that analysis, we may regard the current trend not only as a loss to the sector, but as a subversion of Parliamentary intent.

B. Federal-Provincial Consistency

A second principal argument in favour of having the Federal Court of Appeal exercise the curative powers of equity in charitable registration appeals is that such exercise would better align the federal charities jurisprudence with the charity law of the provinces. Based on the comments of certain members of the Supreme Court of Canada, it is arguably an open question whether there is still a single "common law of charities" in Canada or whether the field is split between the common law of the provinces and the "common law" of the federal courts.[106] The majority decision in *Vancouver Society* supports the better view that the

104. See *Humanics Institute v Minister of National Revenue*, 2014 FCA 265; *World Job and Food Bank Inc v R*, 2013 FCA 65 and *Sagkeeng Memorial Arena Inc v Minister of National Revenue*, 2012 FCA 171.

105. *Vancouver Society*, *supra* note 7 at para 28, per Gonthier J (The Supreme Court of Canada has subsequently stated that the scheme of the *ITA* "does not support a wide expansion of the definition of charity" and that "wholesale reform" is best left to Parliament, but it has not denied this continuing role); see *AYSA*, *supra* note 72 at paras 43-44.

106. *Vancouver Society*, *supra* note 7 at para 28, per Gonthier J. See also *AYSA*, *supra* note 72.

Federal Court of Appeal draws upon and contributes to a single equitable tradition that was received by the provinces from England when it interprets the statutory term "charitable" under the *ITA.*[107] If this is the case, it follows that the federal and provincial courts should rely on the same equitable principles in developing this body of law. If, on the other hand, there are effectively two common laws of charity in Canada, it is nonetheless the case that each of these bodies of law tends to influence each other.[108] In either case, therefore, it must be considered preferable that the same body of equitable principles apply throughout.

C. Equity and Charitable Corporations

If there are compelling arguments in favour of having the Federal Court of Appeal act as a court of equity in the adjudication of charitable registration appeals, there are also predictable objections to this approach. A first is that the majority of applicants for registered charity status are corporations rather than trusts. To the extent that this objection reflects a view that corporate charities fell outside the traditional jurisdiction of the courts of equity, it can be dealt with in short order. While Chancery's jurisdiction over charities tends to be portrayed as a branch of its inherent jurisdiction over trusts,[109] and the two jurisdictions undoubtedly overlap,

---

107. For a discussion, see Kathryn Chan, "Taxing Charities/Imposer les Organismes de Bienfaisance: Harmonization and Dissonance in Canadian Charity Law" (2007) 55 Canadian Tax Journal 481.
108. There are a number of provincial charity law decisions, for instance, that have relied on *Vancouver Society*: these include *Re TLC The Land Conservancy of British Columbia Inc No S36826*, 2014 BCSC 97 at paras 221-26; *Cassano v Toronto Dominion Bank* (2009), 98 OR (3d) 543 (SC) at paras 28, 29; *Save the Heritage Simpson Covenant Society v Kelowna (City)*, 2008 BCSC 1084 at paras 112, 113, 115; *Chénier v Canada (Attorney General)* (2005), 12 CBR (5th) 173 (Ont Sup Ct J) at para 47; and *Alberta Assn for Community Living v Alberta (Municipal Government Board)*, 2000 ABQB 263 at para 29.
109. Picarda*, supra* note 51 at 729, noting that the 8th edition of Warburton suggested that the inherent jurisdiction depended exclusively on the existence of a trust. See also *Construction Industry Training Board v A-G*, [1973] Ch 173 (CA (Civ)(Eng)) at 176.

the judicial power to cure uncertain and defective charitable dispositions predates the legal form of the trust and was historically applied by the Chancellor to a variety of charitable vehicles.[110] Because of the manifest authority of the Crown, the Court of Chancery admittedly had a very limited authority to administer charitable corporations founded by statute or Royal Charter.[111] That being said, the court would intervene in the administration of a charitable corporation where there was no other person to ensure that the charitable funds were being properly managed or where it could find within the corporate structure a trust upon which its authority could be based.[112] Modern Canadian courts have confirmed that, because of the trust-like obligations of corporate charities, "the court maintains its supervisory scheme-making power whether a charity's legal form is as a charitable trust or a charitable corporation".[113] Thus, there is little merit in the argument that the Federal Court of Appeal has no equitable jurisdiction over charitable corporations *per se.*[114]

A more challenging question is whether the principle that the court leans in favour of charity extends to the construction of corporate documents or whether it applies only where a disposition of property would otherwise be void. Several English cases have considered this question but the case law does not provide a unanimous response. In *Inland Revenue Commissioners v Oldham Training and Enterprise Council*[115] ("*Oldham*"),

---

110. Jones, *supra* note 1 at 5, 59, 80. See also Marion R Fremont-Smith, *Foundations and Government: State and Federal Law and Supervision* (New York: Russell Sage Foundation, 1965) at 17.
111. Warburton, *supra* note 54 at 131, 371-73.
112. For a more detailed explanation of the court's historical jurisdiction over charitable corporations, see Maurice Cullity, "The Charitable Corporation: a 'Bastard' Legal Form Revisited" (2001) 17 The Philanthropist 17 at 19-23. See also Kathryn Chan, *The Public-Private Nature of Charity Law* (Oxford: Hart Bloomsbury, 2016) at ch 2.
113. *Re Christian Brothers of Ireland in Canada* (2000), 184 DLR (4th) 445 (Ont CA) at para 71. This is to ensure that gifts made with charitable intent will not fail even if the object of the gift is unclear or uncertain, or contains a technical defect.
114. *Ibid*.
115. [1996] STC 1218 (Ch (CA)(Eng)) [*Oldham*].

the Chancery Division of the High Court refused to bring the principle of benignant construction in aid of a not-for-profit company that sought to benefit from an exemption available to charities under the applicable UK legislation. Justice Lightman opined that so far as he could see, the principle "only applies where a provision or a gift will be held void and fail unless held charitable … I cannot see how this principle has any application where the validity of the provision or gift is not affected by the determination whether the gift is charitable or not".[116] However, in the earlier case of *Guild v Inland Revenue Commissioners*[117] ("*Guild*"), the House of Lords had no hesitation in applying the doctrine of benignant construction in circumstances where the validity of the disposition was not in issue.[118] The First-Tier Tribunal acknowledged these competing approaches in the recent *Human Dignity Trust v The Charity Commission For England and Wales*[119] decision, but found it "unnecessary in the circumstances" to comment on the availability of benignant construction to a corporate body seeking charitable status.[120] However, a benignant approach has been taken with regards to statutory corporations in other jurisdictions.[121] Given the scope of the traditional Chancery jurisdiction and the strange consequences that might arise from having different rules of construction for different types of charitable institutions, the better argument seems to be that the favourable posture towards charity should extend to all charities, regardless of form.[122]

D. Equity and Tax

Another set of objections to having the Federal Court of Appeal apply the tools of equity in adjudicating charitable registration appeals relate to the particular statutory context in which that adjudication takes

116. *Ibid* at 1235.
117. [1992] 2 AC 310 (HL) [*Guild*].
118. *Ibid* at 316.
119. [2014] UK First Tier Tribunal 2013_0013_B (General Regulatory Chamber).
120. *Ibid* at paras 18, 19, 29.
121. *CIR v Medical Council of NZ*, [1997] 2 NZLR 297 (CA) at 318.
122. Luxton, *supra* note 73 at 204.

place. Within the context of the *ITA*, the Court's determination that an appellant's purposes and activities are exclusively charitable results in the appellant being subject to both the significant tax advantages and the significant regulatory burdens of registered charity status.[123] It also results in the overturning of a decision of the Minister of National Revenue. In these circumstances, two questions must be considered. First, as a general matter, do the curative principles of equity apply where an allegedly charitable instrument is being construed for tax purposes? Second, in the specific context of the registered charity regime, has Parliament precluded resort to the rules of equity by enacting a comprehensive code?

With regard to the general question, we may once again seek guidance from the conflicting decisions of the English courts. In *Oldham,* the High Court refused to apply the principle of benignant construction to a party seeking a charitable tax exemption on the ground that the favourable rules of equity were only applicable where the gift would be otherwise void. However, the House of Lords espoused a different approach in *Guild*. Guild was the executor of the estate of one James Russell, a man who bequeathed part of his estate to a town council "for the use in connection with the sports centre in New Berwick or some similar purpose in connection with sport". The Inner House of the Court of Session approved a *cy-près* scheme in connection with the bequest, but the Inland Revenue subsequently ruled that the property had not been "given to charities" for purposes of a capital transfer tax exemption[124] under the *Finance Act 1975*.[125]

Before the House of Lords, the argument focused on whether the sports centre qualified as a charity under the *Recreational Charities Act 1958*,[126] and whether the second branch of the bequest, referring to

123. Canadian courts have tended to emphasize the benefits rather than the burdens of regulatory status: see, for example, *Vancouver Society, supra* note 7 at para 128. However, the burdens are also significant: see generally Jonathan Garton, *Public Benefit in Charity Law* (Oxford: Oxford University Press, 2013).
124. *Guild*, *supra* note 117 at 317.
125. (UK), c 7, Sch. 6, s 10.
126. (UK), c 17 6, s 1.

"some similar purpose in connection with sport", was so widely expressed as to admit of the funds being applied in a manner inconsistent with that legislation.[127] The executor urged the court to apply a benignant construction to the bequest, while the commissioners argued that a benignant construction should not be applied, "since the question was not whether the trust was valid or invalid, but whether it qualified for exemption from tax".[128] The House of Lords favoured the executor's approach, holding:

> the importation into [Scottish] law, for tax purposes of the technical English law of charities involves that a Scottish judge should approach any question of construction arising out of the language used in the relevant instrument in the same manner as would an English judge who had to consider its validity as a charitable gift. The English judge would adopt the benignant approach in setting about that task, and so the Scottish judge dealing with the tax consequences should do likewise.[129]

Adopting a benignant construction, the House of Lords concluded that the testator's intention had been that any "similar purpose" to which the town council might apply his bequest should display the characteristics that qualified the first bequest as a charity.[130] Both bequests had therefore been "given to charities" for purposes of the taxation regime.

Taken together with the Supreme Court of Canada's comments in *Vancouver Society*, *Guild* provides a persuasive argument that, as a general matter, the Federal Court of Appeal should approach the construction of purportedly charitable instruments in the same manner as would a Chancery judge considering the validity of charitable gifts, regardless of the tax consequences of the decision. Parliament retains the power to preclude or modify such an equitable approach, as we have seen, by creating a "comprehensive regulation" for charitable registration and its associated appeals.[131] The Supreme Court of Canada implied in *Vancouver Society* that Parliament had not done this, and that the *ITA* envisaged a resort to "the common law" and "the equitable jurisdiction"

---

127. *Guild, supra* note 117 at 317.
128. *Ibid* at 322.
129. *Ibid* at 323.
130. *Ibid*.
131. *Glaxo Wellcome*, *supra* note 26 at para 36.

for the principles that should guide the court in applying the definition of charity.[132] The question that remains is whether subsequent statutory changes have so transformed the registered charity regime as to prohibit the Federal Court of Appeal from invoking the curative doctrines of equity in adjudicating charitable registration appeals.

The statutory framework governing the charitable registration process has not undergone radical changes since *Vancouver Society* was decided in 1999. The Minister of National Revenue's authority to register qualified organizations continues to flow from subsection 248(1) of the *ITA*, which defines a "registered charity" (*organisme de bienfaisance enregistré*) as:

> [a] charitable organization, private foundation or public foundation, within the meanings assigned by subsection 149.1(1) [or a branch thereof], that is registered in Canada and was either created or established in Canada…that has applied to the Minister in prescribed form and that is at that time registered as a charitable organization, private foundation or public foundation.
>
> L'organisme suivant, qui a présenté au ministre une demande d'enregistrement sur formulaire prescrit et qui est enregistré, au moment considéré, comme oeuvre de bienfaisance, comme fondation privée ou comme fondation publique … au sens du paragraphe 149.1(1), qui réside au Canada et qui y a été constituée ou y est établie.

The subsection 149.1(1) definitions of a "charitable organization" and "charitable foundation" continue to rely heavily on the largely undefined concepts of "charitable purposes" (*fins de bienfaisance*) and "charitable activities" *(activités de bienfaisance*).[133] If the Minister decides that an applicant for registered charity status does not meet the criteria for registration under 149.1(1), an appeal lies from that decision to the

---

132. *Vancouver Society*, *supra* note 7 at para 28. In 2007, the SCC also rejected the argument that the Registered Canadian Amateur Athletic Associations regime was a complete code for amateur sporting activities, or that it indicated an intent to modify the meaning of charity under the *ITA*: See *AYSA*, *supra* note 72 at para 23.

133. Subsection 149.1(1) also requires that no part of the income of either entity be available for the personal benefit of any proprietor, member, shareholder, trustee or settlor*: ITA*, *supra* note 4, at s 149.1(1), "charitable organization" (*oeuvre de bienfaisance*) and "charitable foundation" (*fondation de bienfaisance*).

Federal Court of Appeal.[134]

Within this largely similar framework, one may identify two statutory changes that arguably signal a slight shifting away from the regime's traditional reliance on the charity law tradition of the Chancery court. First, Parliament has recently provided slightly more direction on what purposes and activities are charitable under the *ITA*, primarily by clarifying that it is not a "charitable purpose" to fund the political activities of another charity.[135] This amendment achieves a minor narrowing of the legislative gaps that have historically been filled in by the common law of charities, but is a long way from filling them in completely. Second, Parliament has introduced an internal appeals process for charitable registration decisions, requiring persons who want to appeal a failed application for registered charity status to first serve a written notice of objection on the Minister, and giving the Minister 90 days to respond.[136] This amendment has, to some extent, shifted decision-making authority over registered charity status away from the Federal Court of Appeal, and towards an administrative unit that has no equitable powers of its own.[137] Nevertheless, the Federal Court of Appeal continues to be vested with sole authority to hear appeals of the Minister's decision, and to review "extricable questions of law", such as the proper approach to the construction of charitable objects on the standard of correctness.[138] While both of these amendments move the registered charity regime some way

---

134. This is the combined effect of *ITA*, *supra* note 4, at ss 149.1(22), 168(4), 172(3)(a.1).
135. As per s 149.1(1) of the *ITA*, "'charitable purposes' includes the disbursement of funds to a qualified donee, other than a gift the making of which is a political activity." This definition was introduced in 2012, in Bill C-38, *Jobs, Growth, and Long-term Prosperity Act*, 1st Sess, 41st Parl, 2012 (assented to 29 June 2012), RSC 2012 c 19.
136. *ITA*, *supra* note 4, at s 172(3)(a.1).
137. For a description of the motivation for s 168(4), see Terrance S Carter and Theresa LM Man, "March 2004 Federal Budget Rewrites Tax Rules for Charities" Charity Law Bulletin No. 41 at 1, online: <www.carters.ca/pub/bulletin/charity/2004/chylb41-04.pdf>.
138. *Prescient Foundation v Minister of National Revenue*, 2013 FCA 120 at para 12.

towards being a more comprehensive regime, therefore, neither provides a basis to presume that Parliament intended to preclude the application of equitable principles and doctrines in the determination of charitable registration appeals.

V. Conclusion

The equitable principles and doctrines that make up the common law charities tradition were, to a large extent, developed by the Court of Chancery in order to mitigate the rigour of the law and effectuate imperfect charitable gifts.[139] As the epicenter of Canadian charity law has shifted from the trust law to the tax law domain, the curative principles of equity have all but disappeared from view. The exclusion of these equitable principles from the registered charity jurisprudence has had a discernable impact on the development of charity law in Canada and has contributed to the dramatic record of failed appeals from the registration decisions of the Minister of National Revenue.

I have argued that the Federal Court of Appeal has the authority to exercise the curative powers of the Court of Chancery in determining charitable registration appeals under subsection 172(3) of the *ITA* and that there are compelling reasons for it to do so in an appropriate case. The most likely objections to the Federal Court of Appeal acting as a court of equity — that the appellants are corporations, that the Court's decisions have tax consequences and that the registered charity regime precludes such exercise — are far less potent than a first glance would admit. There remain reasons to be cautious about the exercise of equitable powers within the registered charity regime — especially where tax consequences are in issue, the public has an interest in the construction of purportedly charitable objects not being "strained". However, Parliament has signaled

139. Equity, it is often said, "was introduced to mitigate the rigour of the law": see *Re Vandervell's Trusts (No 2)*, [1974] Ch 269 (CA (Civ)(Eng)) at 322, cited in The Hon BM McLachlin, "The Place of Equity and Equitable Doctrines in the Contemporary Common Law World: A Canadian Perspective," in DWM Waters, ed, *Equity, Fiduciaries and Trusts* (Scarborough: Carswell, 1993) at 39.

through the *ITA* that the Federal Court of Appeal has a continuing role to play in developing and rationalizing the law of charity in Canada. If the Court refuses to allow a role for equity in mitigating the rigour of its decisions, that role may effectively disappear.

Compensatory Remedies for Breach of Trust

Paul S Davies*

This article considers compensatory remedies for breach of trust. The first part of analysis considers the important recent decision of the UK Supreme Court in AIB Group (UK) v Mark Redler & Co Solicitors, and explores its significance and potential impact. The second part examines more closely what is meant by 'equitable compensation', and what remedial rules might develop to govern 'equitable compensation' following a breach of trust.

* Associate Professor, Faculty of Law, University of Oxford; Fellow, St Catherine's College; Associate Member, Maitland Chambers. An earlier version of this paper was presented at a conference in July 2015 on "Modern Studies in the Law of Trusts and Wealth Management: Theory and Practice", which was held at the Supreme Court of Singapore and organised by Singapore Management University, the University of York, and the Singapore Academy of Law. I am grateful to the organisers – Richard Nolan, Hang Wu Tang and Kelvin Low – and to participants for their comments. I should also thank Sarah Green, Nick McBride and Peter Turner for their comments on an earlier draft. The usual disclaimers apply.

I. Introduction

In the important decision of the United Kingdom Supreme Court in *AIB Group (UK) v Mark Redler & Co Solicitors*[1] ("*AIB*"), Lord Toulson observed that "[t]he debate [that] has followed *Target Holdings Ltd v Redferns*[2] ("*Target*") is part of a wider debate, or series of debates, about equitable doctrines and remedies and their inter-relationship with common law principles and remedies, particularly in a commercial context".[3] As regards compensatory remedies for breach of trust, the scope of the debate has effectively been narrowed — at least in England and Wales — by the decision in *AIB*: nothing is to be gained by falsifying and surcharging the account, which is the traditional approach in equity, and the court can simply award compensation for loss caused by the trustee's

1. [2014] UKSC 58 [*AIB*].
2. [1996] AC 421 (Eng) [*Target*].
3. *AIB, supra* note 1 at para 47.

breach of duty.[4] Lord Toulson thought that "in circumstances such as those in *Target,* the extent of equitable compensation should be the same as if damages for breach of contract were sought at common law".[5] In a similar vein, in *Thanakharn Kasikhorn Thai Chamkat (Mahachon) v Akai Holdings Ltd (in liquidation)*[6] ("*Akai*"), Lord Neuberger, sitting as a non-permanent judge of the Hong Kong Court of Final Appeal, commented that:

> even if the principles … had suggested that Akai was entitled to equitable compensation in an amount greater than it should recover by way of common law damages, I would have been very sympathetic to the notion that the equitable remedy would have to be refashioned so as to equate the amount of such compensation with the common law damages.[7]

This suggests that the common law rules may have some impact upon the principles of equitable compensation.

The decision in *AIB* accelerates the need to establish clear principles that can be employed when awarding equitable compensation for breach of trust. It is suggested that there are good reasons for equity to adopt its own particular approach, but the comments of Lord Toulson and Lord Neuberger are likely to prove to be influential. It is therefore worth considering whether the rules of equitable compensation should mirror those available for breach of contract, or for tort (although if the latter it would need to be determined whether an analogy should be made with negligence, deceit, or some other tortious wrong). However, before examining the principles of equitable compensation in greater depth, the

---

4. John Heydon, Mark Leeming & Peter Turner, *Meagher Gummow and Lehane's Equity: Doctrines and Remedies*, 5d (London: LexisNexis Butterworths, 2015), as the editors of Meagher, Gummow and Lehane have recently observed "[t]he advent of the term "equitable compensation" in the last two to three decades supplied a name to a form of relief which derived from the principles of account, but was awarded without the accounting procedures" at 23-30.
5. *AIB*, *supra* note 1 at para 71.
6. [2010] HKCFA 64 [*Akai*].
7. *Ibid* at para 155. See also *Canson Enterprises Ltd v Boughton & Co*, [1991] 3 SCR 534 at 585-87 [*Canson*].

decision in *AIB* should first be explained.[8]

II. The Demise of Falsification?

"The basic right of a beneficiary is to have the trust duly administered in accordance with the provisions of the trust instrument, if any, and the general law".[9] The traditional means used to ensure the proper administration of the trust was for the beneficiary to take an account. Upon discovering a breach of trust, a beneficiary might falsify or surcharge the account. If the trustee had failed to exercise due care and skill leading to the fund not being worth as much as it would have been if managed by a reasonably prudent trustee, the account could be surcharged.[10] If the trustee misapplied trust monies in breach of trust, then the beneficiary was entitled to falsify that disbursement.[11] As Lord Sumption explained in *Williams v Central Bank of Nigeria,*[12] "[i]f the trustee misapplied the assets, equity would ignore the misapplication and simply hold him to account for the assets as if he had acted in accordance with his trust".[13]

A good example was provided by the case of *Re Dawson.*[14] In 1939, a trustee paid away NZ£4,700 in breach of trust. At the time of the improper disbursement, there was parity between the New Zealand

8. See Lusina Ho, "Equitable Compensation on the Road to Damascus?" (2015) 131 Law Quarterly Review 213; Peter Turner, "The New Fundamental Norm of Recovery for Losses to Express Trusts" (2015) 74:2 Cambridge Law Journal 188; Paul Davies, "Remedies for Breach of Trust" (2015) 78:4 Modern Law Review 681; Peter Watts, "Agents' Disbursal of Funds in Breach of Instructions" (2016) 1 Lloyd's Maritime and Commercial Law Quarterly 118.
9. *Target*, *supra* note 2 at 434; see also *AIB*, *supra* note 1 at para 64.
10. *Bristol & West Building Society v Mothew*, [1998] Ch 1 (CA (Civ)(Eng)) at 17 [*Mothew*]; *Fry v Fry* (1859), 27 Beav 144 (Ch (Eng)); *Nestle v National Westminster Bank plc*, [1993] 1 WLR 1260 (CA (Civ)(Eng)) [*Nestle*]; *Re Mulligan*, [1998] 1 NZLR 481 (HC) [*Mulligan*].
11. *Knott v Cottee* (1852), 16 Beav 77 (Ch (Eng)) [*Knott*]; *Re Massingberd's Settlement* (1890), 63 LT 296 (CA (Eng)); *Re Dawson (dec'd)*, [1966] NSWR 211 (SC (Austl)) [*Re Dawson*].
12. [2014] UKSC 10.
13. *Ibid* at para 13.
14. *Re Dawson*, *supra* note 11.

pound and Australian pound. However, by the time the claim was made against the trustee to restore the trust estate, NZ£4,700 was worth nearly A£6,000. In a judgment which has been influential throughout the Commonwealth, Justice Street held that the defaulting trustee was under a strict liability to make good the trust fund. In order to restore the trust fund to the position it would have been in had the money not wrongly been paid away, the trustee was required to pay A£6,000. Street J reviewed the authorities[15] and concluded that:

> [t]he cases to which I have referred demonstrate that the obligation to make restitution, which courts of equity have from very early times imposed on defaulting trustees and other fiduciaries is of a more absolute nature than the common law obligation to pay damages for tort or breach of contract.[16]

Indeed, the judge held that "[c]onsiderations of causation, foreseeability and remoteness do not readily enter the matter".[17]

In effect, the wrongdoing trustee in *Re Dawson* was taken to have paid away his own monies rather than the monies of the trust fund, since equity would not countenance the possibility that the trustee acted badly when it could insist that the trustee acted properly. The trustee was held up to his primary obligation to act as a reasonable and honest trustee. As a trustee was still under a primary obligation to account for NZ£4,700 to the trust fund, and this the trustee could still do. Upon the beneficiary falsifying the account to delete the unauthorised disbursement, the trustee had the option of restoring the relevant trust property *in specie* — in other words the particular monies paid away — or the money

15. Including *Caffrey v Darby* (1801), 6 Ves Jr 488 (Ch (Eng)); *Clough v Bond* (1838), 3 My & C 490 (Ch (Eng)).
16. *Re Dawson*, *supra* note 11 at 216.
17. *Ibid* at 215.

substitute,[18] out of the trustee's own funds.[19] This is a different approach from that taken by the common law. As Lord Millett has explained, extra-judicially:

> Lord Diplock has said that a contracting party is under a primary obligation to perform his contract and a secondary obligation to pay damages if he does not. It is tempting, but wrong, to assume that a trustee is likewise under a primary obligation to perform the trust and a secondary obligation to pay equitable compensation if he does not. The primary obligation of a trustee is to account for his stewardship. The primary remedy of the beneficiary – any beneficiary no matter how limited his interest – is to have the account taken, to surcharge and falsify the account, and to require the trustee to restore to the trust estate any deficiency which may appear when the account is taken. The liability is strict. The account must be taken down to the date on which it is rendered. That is why there is no question of "stopping the clock".[20]

However, the House of Lords shifted away from this approach, at least in the context of "commercial" trusts, in *Target.* Redferns was a firm of solicitors acting for both the borrowers and the lender, Target Holdings. Redferns held the mortgage advance of £1.5 million on a bare trust for Target Holdings, with authority to release the money to the borrowers only upon receipt of the executed conveyance and mortgage of the property. In breach of trust, Redferns released the money before the documents were executed. The property was in due course found to be worth only £500,000. Target Holdings argued that Redferns should reconstitute the trust fund by paying the difference between the value of the property and the money advanced to the borrowers. This was accepted by the

---

18. It is for this reason that the language of "substitutive compensation" is sometimes used: see Steven Elliott, *Compensation Claims Against Trustees* (PhD Dissertation, Oxford University Faculty of Law Library, 2002 [unpublished]; Steven Elliott & James Edelman, "Money Remedies Against Trustees" (2004) 18:3 Tolley's Trust Law International 116; *Agricultural Land Management Ltd v Jackson [No 2]*, [2014] WASC 102 (Austl) at paras 334-49 [*Jackson (No 2)*].
19. *Knott*, *supra* note 11 at paras 79, 80; *Re Bennison* (1889), 60 LT 859 (Ch (Eng); *Re Salmon* (1889), 42 Ch D 351 (CA (Eng)) at 357 [*Salmon*].
20. Peter Millett, "Equity's Place in the Law of Commerce" (1998) 114 Law Quarterly Review 214 at 255.

Court of Appeal,[21] but rejected by Lord Browne-Wilkinson, who gave the only reasoned speech in the House of Lords. Lord Browne-Wilkinson effectively reasoned in terms of the trustee being under a "secondary obligation" to compensate the beneficiary for losses caused by the breach of trust — precisely the approach criticised by Lord Millett in the passage quoted above.

Yet even the decision in *Target* did not necessarily mean that the "traditional" approach to falsification of the account was moribund, as Lord Browne-Wilkinson also said:

> I have no doubt that, until the underlying commercial transaction has been completed, the solicitor can be required to restore to client account moneys wrongly paid away. But to import into such a trust an obligation to restore the trust fund once the transaction has been completed would be entirely artificial.[22]

Significantly, in *Target* the relevant mortgage documents were subsequently executed and received by Redferns. Since Redferns remained trustees even after their breach of trust, Redferns still had the authority to receive those mortgage documents as trustees on behalf of the trust.[23] The trust fund had therefore been reconstituted and there was nothing wrong with the state of the trust fund at the time the account was taken: Target Holdings was entitled to see *either* £1.5 million *or* the mortgage documents in the trust fund. The latter was present, so there was no defect in the fund. Remedies in equity are assessed at the date of judgment, not the date of breach.[24] Using the language of Lord Browne-Wilkinson, it might be said that the transaction had been "completed" upon receipt of the mortgage documents.

Such reasoning meant that, even after *Target,* some courts continued to employ the strict approach of falsifying the account, awarding substantial relief even where the trustee's breach of duty did not cause any

---

21. This claim succeeded in the Court of Appeal: *Target Holdings v Redferns*, [1994] 1 WLR 1089 (CA (Civ)(Eng)).
22. *Target*, *supra* note 2 at 436.
23. Millett, *supra* note 20; Matthew Conaglen, "Explaining *Target Holdings* v *Redferns*" (2010) 4:3 Journal of Equity 288.
24. *Target*, *supra* note 2 at 437; see too *AIB*, *supra* note 1 at para 140.

loss. For example, in *Knight v Haynes Duffell Kentish & Co*[25] ("*Knight*"), the claimants advanced monies to its solicitors, Linnels. Linnels was to pay the monies to a company on completion of a transaction. Upon completion, a trade name was to be assigned to the claimants. In breach of trust, Linnels paid away the monies without ensuring the assignment of the trade name. The claimants therefore instructed the defendant firm of solicitors (HDK) to recover its losses from Linnels, but HDK delayed the litigation to such an extent that it was ultimately struck out for want of prosecution. The claimants therefore sued HDK for the lost opportunity to sue Linnels in breach of trust. Even though the trademark was actually worthless, the Court of Appeal held that such facts did not affect the remedy to be awarded. The Court was not concerned with compensating losses caused by Linnels' breach of trust, but rather with falsifying the disbursement made in breach of trust. As Lord Justice Aldous said:

> First, in the present case the breach was the release of the money. The trust required the money to be held against provision of both the shares and the assignment. As there had been no assignment, the money should not have been paid out. Second, the principle in *Target* only applies where the underlying transaction covered by the trust had been completed.[26]

In *Knight,* the transaction had not been completed. The Court of Appeal therefore decided that the facts fell outside the scope of Lord Browne-Wilkinson's causal analysis in *Target.*

The decision in *Knight* seems to be consistent with the traditional approach to the accounting process, but has perhaps now been undermined by the decision in *AIB.* Mark Redler & Co is a firm of solicitors which was retained to act for the Sondhi family and AIB, a bank, on the re-mortgage of the Sondhis' family home. AIB advanced £3.3 million to Redler for this purpose. The letter of instruction incorporated the Council of Mortgage Lenders' Handbook for England and Wales,[27] by virtue of which the mortgage lender required a fully enforceable first charge over the property and that all existing charges be redeemed on or

25. [2003] EWCA Civ 223.
26. *Ibid* at para 38.
27. Council of Mortgage Lenders, "CML lenders' handbook for conveyances" (2015), online: <www.cml.org.uk/cml/handbook>.

before completion. The handbook also stated: "You [Redler] must hold the loan on trust for us [AIB] until completion. If completion is delayed, you must return it to us when and how we tell you".

The Sondhis' property was already subject to a charge in favour of Barclays Bank Plc. The Barclays charge secured borrowings of about £1.5 million on two accounts. Unfortunately, Redler only paid to Barclays enough money to pay off one of the two accounts (about £1.2 million), which was insufficient to redeem the Barclays charge. £309,000 remained outstanding. Barclays refused to release its charge unless the debt was paid in full. The borrowers, who had received the balance of the £3.3 million, initially promised to do so, but never did. Redler tried to resolve its error without involving AIB but eventually told the bank of the breach of duty; AIB then negotiated directly with Barclays, and AIB's charge was registered as a second charge. The Sondhis subsequently defaulted on the loan and declared bankruptcy. The property was sold by Barclays for £1.2 million. AIB as second chargee received £867, 697.

By paying away the mortgage monies without obtaining a first legal charge over the property, Redler acted in breach of trust.[28] AIB argued that completion had not yet occurred, so Redler remained under a duty to hold the mortgage advance on trust for AIB. AIB therefore sought £3.3 million in order to reconstitute the trust fund.[29] Redler, on the other hand, argued that its liability should be limited to the difference in value of the bank's security caused by Redler's failure to pay off the entirety of the Barclays charge; this was only around £300,000 (the sum received by Barclays as first chargee).

Redler's argument succeeded at every level. A unanimous Supreme

28. *AIB*, *supra* note 1 at para 140 Lord Reed was attracted by the idea that the breach of trust only involved the misapplication of the £309,000 paid to the Sondhi's rather than Barclays, but this had been rejected by the Court of Appeal and was not challenged in the Supreme Court; the breach of trust was paying away the entire £3.3 million.
29. Strictly the claim was for £3.3 million minus the £867,697 actually received from the sale of the property.

Court[30] insisted that a causal link between Redler's breach of duty and AIB's loss needed to be established, regardless of whether the claim was brought at common law or in equity.

There was one potentially important difference between *Target* and *AIB*. In *Target,* the relevant mortgage documents were subsequently executed and received by Redferns. In *AIB,* Redler *never* obtained a first legal charge over the property in favour of AIB. AIB thought that since completion required there to be a first legal charge in its favour, and this never eventuated, completion had not occurred, and that as a result Redler was required to restore the monies wrongly paid away. Yet Lord Toulson side-stepped such arguments because:

> as a commercial matter the transaction was executed or 'completed' when the loan monies were released to the borrowers. At that moment the relationship between the borrowers and the bank became one of contractual borrower and lender.[31]

This pragmatic approach is perhaps understandable given the context of the dispute in question. AIB was anxious to push through the Sondhis' remortgage of the property, which was "driven by the need to facilitate business lending which the bank was very keen to make".[32] But the result very much broadens the scope of "completion". The Supreme Court was prepared to find that there was completion upon satisfaction of the commercial purpose, but this is less certain than insisting upon compliance with the terms of the solicitor's instructions. Redler's breach of trust meant that there was no completion in accordance with the requirements of Redler's instructions. The decision in *AIB* makes it more difficult for a lender — or any settlor — to set the terms of completion. A court might find there to be completion even if the beneficiary would not agree. Indeed, AIB did not simply seek the relationship of lender-borrower; AIB wanted to be a secured lender with priority over other chargees. It might be thought that this commercial purpose was not fulfilled.

---

30. Lord Toulson and Lord Reed gave reasoned speeches; Lord Neuberger, Lady Hale, and Lord Wilson agreed with both speeches.
31. *AIB*, *supra* note 1 at para 74.
32. See *e.g. AIB*, *supra* note 1 at para 14.

It is suggested that the approach taken in *AIB* may be unfortunate. It restricts the ability of settlors to define completion according to their wishes. For instance, imagine that, in breach of trust, a trustee purchased a second-hand car rather than a new car. Has the transaction been completed? On the approach of the Supreme Court in *AIB,* it might be tempting to conclude that, since the trustee has purchased a car, the beneficiary should simply sue for the difference in value between the second-hand car he now has rather than the new car he was entitled to under the terms of the trust. Yet it seems unsatisfactory for the wrong type of car to be the beneficiary's problem, rather than the trustee's problem.

There was a clear logic behind equity's traditional recognition of the beneficiary's ability to choose to falsify the disbursement made and treat the car as having been purchased with the trustee's own money; on that approach, the unwanted second-hand car was the trustee's problem to deal with. The hassle of selling it to realise its value, for example, lay with the wrongdoing trustee rather than the innocent beneficiary. This is admittedly different from the approach at common law, but, as Lord Millett explained,[33] this might be justified by the higher standards demanded by equity. A trustee holds particular power over a beneficiary and a beneficiary's assets, and should be held to a higher standard. It is important that a trustee complies with the terms of the trust instrument.

A second-hand car can never become a new car, whereas a second charge could become a first charge in circumstances akin to *AIB*. The two cases might therefore be distinguished. At first instance in *AIB*, His Honour Judge David Cooke observed:

> [t]here is no parallel between a charge which is, at the moment of creation, a second ranking security but can be (and is intended and required to be) promoted into a first ranking security by redeeming a prior charge, and a second hand car which can never be transmuted into a new one. The former is what the solicitors were authorised and instructed to obtain in this case and the latter is, on [counsel's] hypothesis, an unauthorised purchase.[34]

This analogy was not pursued on appeal, but the more expansive approach

33. Millett, *supra* note 20.

34. *AIB Group (UK) v Mark Redler & Co*, [2012] EWHC 35 (Ch) at para 30.

of Lord Toulson in particular suggests that the focus should be upon the loss caused by the breach of duty in both instances.

Lord Toulson clearly favoured a shift away from what might have been considered to be earlier orthodoxy. His Lordship thought that treating the trustee solicitors as having paid away their own monies in *AIB* was simply "fairy tales".[35] It might nonetheless be thought that since the solicitors had no authority to pay the money away without obtaining a first legal charge in return, any "fiction" that might be introduced could be tolerated. But in any event, even if the monies paid away are considered to be trust monies that does not inevitably lead to the conclusion that the court should only award compensation for loss. As Lord Justice Longmore recently observed: "[e]quity's response to the breach of this trust is not to give redress for the breach in the form of equitable compensation but to enforce the duty".[36] The means by which the trustee's duty is enforced is through an action in debt.

The action for the agreed sum, or debt claim, is obviously not peculiar to equity. It is not a claim for loss, or concerned with any secondary obligations that might arise after a breach of duty. Rather, the action for the agreed sum seeks to enforce the primary obligations voluntarily assumed by the defendant in the action. The action for an "equitable debt" was apparently accepted by earlier cases. For example, in *ex parte Adamson, In re Collie*,[37] Lord Justice of Appeal James and Lord Justice of Appeal Baggallay noted that relief in such cases was by way of "a suit … for equitable debt or liability in the nature of a debt. It was a suit for the restitution of the actual money or thing, or value of the thing, of which the cheated party had been cheated".[38] This also explains the outcome of cases such as *Re Dawson* and *Knight.*

In *AIB* Lord Toulson dismissed the analysis based upon debt. He recognised that the authorities do refer to "an equitable debt, or liability

35. *AIB*, *supra* note 1 at para 69.
36. *Novoship (UK) Ltd v Nikitin*, [2014] EWCA Civ 908 at para 104.
37. (1878), 8 Ch D 807 (CA (Eng)) [*Collie*].
38. *Ibid* at 819. See also *In re Smith, Fleming & Co* (1879), 11 Ch D 306 (CA (Eng)) at 311, per James LJ; *Webb v Stenton* (1883), 11 QBD 518 (CA (Eng)) at 530, per Fry LJ.

in the nature of a debt",[39] but thought that the language of debt was only used because it was necessary "long before the expression 'equitable compensation' entered the vocabulary".[40] Lord Toulson concluded that equity only ever really awarded compensation which was "clothed by the court in the literary costume of equitable debt, the debt being for the amount of the loss caused by the fraud".[41] The evocative rhetoric is not, however, entirely convincing. It is not apparent from such earlier cases that the language of debt was simply a "literary costume". Indeed, an action for the agreed sum may appear particularly appropriate in a situation where the trustee has undertaken to return trust monies to the beneficiary if completion according to the instructions does not occur.[42] This was precisely the case in *AIB.* After all, the bank had stipulated that "You [Redler] must hold the loan on trust for us [AIB] until completion. If completion is delayed, you must return it to us when and how we tell you".[43] It is possible to view this as having created an obligation for Redler to pay AIB £3.3 million. This debt could be discharged either by Redler's paying over £3.3 million to AIB, or by ensuring completion of the transaction. The completion of the transaction would have discharged the obligation to pay the debt. This analysis was rejected by Lord Toulson, who said that "a monetary award which reflected neither loss caused nor profit gained by the wrongdoer would be penal".[44] This would mean that all debt claims, even those brought at common law,[45] should be considered penal in nature. That would be a startling conclusion. It is not penal to hold a party — particularly a trustee — up to the duties to which he or she voluntarily assented.[46] This cannot be the view which

39. *AIB*, *supra* note 1 at para 61, citing *Collie*, *supra* note 37.
40. *AIB*, *supra* note 1 at para 61.
41. *Ibid*. See also Ho, *supra* note 8 at 215-16.
42. *AIB, supra* note 1 at para 4.
43. See text following note 27.
44. *AIB*, *supra* note 1 at para 64.
45. Common law claims are not at all uncommon: see, notoriously, *White & Carter (Councils) Ltd v McGregor*, [1962] AC 413 (HL).
46. Thus the remedies of specific performance and injunction – which similarly enforce the primary obligations owed – should also not be considered to be penal.

Lord Toulson truly intended to express.

Nevertheless, there are clearly significant differences between the old orthodoxy and the new approach. For instance, in *Hall v Libertarian Investments Ltd*[47] ("*Libertarian*"), Lord Millett, sitting as a non-permanent judge of the Hong Kong Court of Final Appeal, emphasised that after an unauthorised disbursement has been falsified, the trustee must make good the deficit in the trust fund either *in specie* or in money, but that this is "not compensation for loss".[48] Yet in *AIB,* Lord Reed thought that "[i]f the property cannot be restored *in specie*, the trustee must restore the trust fund to the position it would have been in but for the breach, by paying into the fund sufficient pecuniary compensation to meet that objective".[49] Shifting away from an action for the agreed sum towards compensation allows scope for arguments surrounding consequential loss which are simply irrelevant in a debt claim. After all, if a builder completes his or her work for a client and then sues the client for the agreed sum, it does not matter that the builder did not really suffer any loss at all (because if he had not done the work he would have suffered even greater losses) or that the builder suffered much more extensive losses (because the work was much more expensive than envisaged). The client would simply have to pay the builder the agreed sum.

However, according to the decision in *AIB,* this analysis based upon the liquidated sum should no longer be employed when trustees pay away money in breach of trust. The focus should instead be upon awarding compensation for loss. Before analysing how equitable compensation should be understood in greater detail, it is important to highlight some further limitations and difficulties with the decision in *AIB* which might restrict its impact elsewhere in the common law world.

47. [2013] HKCFA 93 [*Libertarian*].
48. *Ibid* at 168.
49. *AIB*, *supra* note 1 at para 90.

III. The Scope of *AIB*

A. Geographical Scope

In *FHR European Ventures LLP v Cedar Capital Partners LLC*[50] ("*FHR*"), Lord Neuberger, speaking for the Supreme Court, stated: "it seems to us highly desirable for all those jurisdictions to learn from each other, and at least to lean in favour of harmonising the development of the common law round the world".[51] This was picked up by Lord Reed in *AIB*,[52] and Lord Toulson also relied upon decisions elsewhere in the common law world in favouring a compensatory approach. However, it is uncertain whether the reasoning in *AIB* will convince the highest courts in other jurisdictions.[53] Lord Reed thought that there exists:

> a broad measure of consensus across a number of common law jurisdictions that the correct general approach to the assessment of equitable compensation for breach of trust is that described by Justice McLachlin (as she then was) in *Canson*[54] and endorsed by Lord Browne-Wilkinson in *Target.*[55]

However, it is unclear how broad any measure of consensus really is. *Canson* is a much-cited decision, but concerned a lawyer's conflict of duty rather than a breach of trust.[56] The lawyer did not hold any property on trust, so there was no breach of any custodial duty. Indeed, the court in *Canson* explicitly recognised that breach of trust cases required reconstitution of trust funds and that equitable compensation was only appropriate where reconstitution was not possible. McLachlin J even said that "compensation is an equitable monetary remedy which

---

50. [2014] UKSC 45 [*FHR*].
51. *Ibid* at para 45.
52. *AIB*, *supra* note 1 at para 121.
53. The question seems open in Singapore: see *e.g. Maryani Sadeli v Arjun Permanand Samtani*, [2014] SGCA 55; *Then Khek Koon v Arjun Permanand Samtani*, [2014] 1 SLR 245 (Sing (HC)).
54. *Canson, supra* note 7.
55. *AIB*, *supra* note 1 at para 133.
56. For thorough discussion, see Lionel Smith, "The Measurement of Compensation Claims Against Trustees and Fiduciaries" in Elise Bant & M Harding, eds, *Exploring Private Law* (Cambridge: Cambridge University Press, 2010) at 363.

is available when the equitable remedies of restitution and account are not appropriate".[57] Lord Reed cited subsequent decisions of the Supreme Court of Canada in support of McLachlin J's dissenting speech, but none concerned a misapplication of trust property.[58] Similarly, a trust was not at issue in the decision of the Supreme Court of New Zealand in *Premium Real Estate Ltd v Stevens*.[59]

There is, in principle, a distinction between breaches of the fiduciary duty of loyalty and breaches of the custodial duties of a trustee.[60] Yet by relying upon cases properly concerning breach of fiduciary duty as relevant to cases concerning breach of trust, the decisions in *Target* and *AIB* have blurred the boundaries. This might suggest a shift in approach, such that the principles relevant to claims for breach of fiduciary duty are also relevant to claims for breach of trust. That will be considered more fully below. But it is a little odd that there is no recognition in either *Target* or *AIB* that this is a potentially controversial step, and the lack of transparency over the moves made by the Supreme Court undermines the persuasiveness of its actions. Indeed, in *Target* — the origin of this shift in approach — Lord Browne-Wilkinson did not even mention the word "falsify", and instead relied upon cases concerned with surcharge[61] or breach of fiduciary duty[62] in support of his conclusion that

57. *Canson*, *supra* note 7 at 556.

58. *AIB*, *supra* note 1 at para 122, Lord Reed cited *M(K) v M(H)*, [1992] 3 SCR 6 (which concerned incest); *Cadbury Schweppes v FBI Foods*, [1999] 1 SCR 142 (which concerned breach of confidence); and *Hodgkinson v Simms*, [1994] 3 SCR 377 (which involved breach of fiduciary duty due to conflict of interest).

59. [2009] NZSC 15 cited in *AIB*, *supra* note 1 at para 126 Lord Reed explicitly recognised that the other New Zealand case upon which he relied, although a trust case, was essentially an example of surcharge rather than falsification: *Bank of New Zealand v New Zealand Guardian Trust Co Ltd*, [1999] 1 NZLR 664 (CA) [*Bank of New Zealand*], cited at *AIB*, *supra* note 1 at para 127.

60. James Penner, "Distinguishing Fiduciary, Trust and Accounting Relationships" (2014) 8 Journal of Equity 203.

61. See *e.g. Nestle*, *supra* note 10; *Bartlett v Barclays Bank Trust Co Ltd (Nos 1 and 2)*, [1980] Ch 515 (Eng) [*Bartlett Nos 1 and 2*].

62. See notably *Canson*, *supra* note 7.

the only remedy open to Target Holdings was equitable compensation.[63] Nevertheless, this inclination to blur conceptual boundaries may be gathering sufficient steam that it is too late to reverse the trend. For instance, Lord Toulson in *AIB*,[64] and Permanent Judge Ribeiro in *Libertarian*[65] cited with approval the following schema of Justice Tipping in *BNZ v NZ Guardian Trust Co Ltd*:[66]

> [b]reaches of duty by trustees and other fiduciaries may broadly be of three different kinds. First, there are breaches leading directly to damage to or loss of the trust property; second, there are breaches involving an element of infidelity or disloyalty which engage the conscience of the fiduciary; third, there are breaches involving a lack of appropriate skill or care.[67]

Ribeiro PJ thought that the principles for quantifying loss were the same in the first two categories of case. This might provide some support for the approach taken in *AIB*.

However, it is important to note that the decision of the Hong Kong Court of Final Appeal in *Libertarian* essentially concerned surcharge rather than falsification. In *Libertarian,* the trustee held monies for the purpose of acquiring shares in a company on behalf of the beneficiary. In breach of trust, the trustee misappropriated the trust monies and falsely told the beneficiary that the relevant shares had been acquired. The value of the shares rose considerably. Faced with such facts, the beneficiary could have falsified the disbursement made and sought reconstitution of the trust fund through the trustee's paying back the money taken, but this was much less valuable than a reparative claim for the losses suffered through not acquiring the shares as instructed. The beneficiary therefore sued the trustee for the lost profits that would have been made if the trustee had acted in accordance with his duties. This was therefore a case of loss: the beneficiary sought the difference between the value of the trust fund at the date of judgment and the value that

---

63. See *e.g.* Charles Mitchell "Equitable Compensation for Breach of Fiduciary Duty" (2013) 66:1 Current Legal Problems 307 at 323-27.
64. *AIB*, *supra* note 1 at para 60.
65. *Libertarian*, *supra* note 47 at para 75.
66. *Bank of New Zealand, supra* note 59.
67. *Ibid* at 678.

the trust fund would have had had the trustee acted properly. However, Lord Millett also considered the basis of falsification and insisted that it was inappropriate to consider compensation for loss in instances of restoration following a misapplication of trust assets.[68] The two sides of the account — falsification and surcharge — rest on different bases.

Lord Millett's comments in *Libertarian* do not sit comfortably with the thrust of the reasoning in *AIB.* Nor does the judgment of the High Court of Australia in *Youyang Pty Ltd v Minter Ellison Morris Fletcher*[69] ("*Youyang*") seem entirely consistent with that in *AIB.* In *Youyang,* a company paid over monies to solicitors for the purposes of an investment. Part of those monies were to be paid away by the solicitors in return for a bearer deposit certificate which would provide security for the investment. Upon receipt of that certificate, the solicitors would then pay the rest of the money to an investment company for investment in speculative market activities. In breach of trust, the solicitors paid the money away without receiving the bearer certificate. Lord Reed observed that *Youyang* was based on "broadly analogous facts" to *Target,* "with the important distinction that the security — which would have been good — was never provided".[70] This clearly enabled *Target* to be distinguished; the High Court in *Youyang* ordered the solicitors to repay the monies wrongly paid away in breach of trust. Lord Reed thought that this meant *Youyang* was consistent with *Target.* That must be right on the question of whether or not the fund was reconstituted. But it is less clear whether the reasoning in *Youyang* is truly consistent with *AIB.* The first legal charge over the Sondhis' property was not provided in *AIB,* just as the security in *Youyang* was never provided, yet Redler was not ordered to repay the monies advanced in breach of the trust. The High Court in *Youyang* was unconcerned with loss, and uninterested in the fact that the conduct of third parties meant that the loss suffered by the claimants would have happened anyway. As the High Court held that Youyang was not provided at any stage with the security for which it had bargained. It

68. *Libertarian*, *supra* note 47 at para 168.
69. [2003] HCA 15 [*Youyang*].
70. *AIB*, *supra* note 1 at para 124, per Lord Reed.

is not to the point that, in addition to the breaches of trust by Minters, there may also have been dishonest and discreditable subsequent acts by third parties which led to the loss of the funds.[71] To present the case by fixing upon those subsequent acts, to adopt the remarks of Bowen LJ in *Magnus*,[72] would be "an ocular illusion", because the loss of the trust funds occurred as soon as the trustee wrongly disbursed them, at the completion on 24 September 1993.[73]

The High Court insisted that trustees need to be "kept up to their duty".[74] In *AIB,* by contrast, the fact that the loss would have happened anyway *did* affect the compensatory remedy awarded. Perhaps the cases are distinguishable since the reason why the losses would have happened anyway in *Youyang* was the conduct of a third party, whereas in *AIB* the reason was a fall in the property market. The reasoning of *Youyang* is not obviously the same as that employed in *AIB.* The High Court of Australia did not focus upon loss caused by a breach of duty, whereas this was central to the decision in *AIB.*

B. Commercial Trusts

After *Target,* the scope of the compensatory approach was a little unclear given Lord Browne-Wilkinson's novel distinction between traditional and commercial trusts. His Lordship said:

> [t]he obligation to reconstitute the trust fund applicable in the case of traditional trusts reflects the fact that no one beneficiary is entitled to the trust property and the need to compensate all beneficiaries for the breach. That rationale has no application to a case such as the present. To impose such an obligation in order to enable the beneficiary solely entitled (*i.e.* the client) to recover from the solicitor more than the client has in fact lost flies in the face of common sense and is in direct conflict with the basic principles of equitable compensation. In my judgment, once a conveyancing transaction has been completed the client has no right to have the solicitor's client account

71. *McCann v Switzerland Insurance* [2000], 203 CLR 579 (HCA) at 18, 135; *Magnus v Queensland National Bank* (1888), 37 Ch D 466 (CA (Eng)) at 471-72, 477, 479-80 [*Magnus*] cited in *Youyang*, *supra* note 69 at para 63.
72. *Magnus*, *supra* note 71 at 480.
73. *Youyang*, *supra* note 69.
74. *Ibid.*

> reconstituted as a 'trust fund'.[75]

This distinction between "traditional" and "commercial" trusts has been questioned.[76] For example, in *Bairstow v Queens Moat Houses Plc,*[77] Lord Justice Robert Walker said:

> [i]t may be that a more satisfactory dividing line is not that between the traditional trust and the commercial trust, but between a breach of fiduciary duty in the wrongful disbursement of funds of which the fiduciary has this sort of trustee-like stewardship and a breach of fiduciary duty of a different character (for instance a solicitor's failure to disclose a conflict of interest as in Canson).[78]

In *AIB,* Lord Reed noted this controversy, and said "[t]hat it is not to say that there is a categorical distinction between trusts in commercial and non-commercial relationships".[79] The duties and liabilities of trustees depend upon the terms of the trust and relationship between the parties. Lord Toulson held that "it is a fact that a commercial trust differs from a typical traditional trust in that it arises out of a contract rather than the transfer of property by way of gift. The contract defines the parameters of the trust".[80] In such circumstances, the duties of the trustee are "likely to be closely defined and may be of limited duration".[81]

Lord Toulson cited with approval an article by Professor Hayton.[82]

---

75. *Target*, *supra* note 2 at 436.
76. See also *Youyang*, *supra* note 69 where the High Court of Australia considered it preferable to focus on "the scope and purpose of the trust" at para 49. It may be that an account should only be available where there is continuous and custodial trusteeship: Joshua Getzler, "Equitable Compensation and the Regulation of Fiduciary Relationships" in Peter Birks & Francis Rose, eds, *Restitution and Equity Volume 1: Resulting Trusts and Equitable Compensation* (London: Mansfield Press, 2000) 249 at 249-50.
77. [2001] EWCA Civ 712.
78. *Ibid* at para 53.
79. *AIB*, *supra* note 1 at para 102.
80. *Ibid* at para 70.
81. *Ibid*. See also *AIB*, *supra* note 1 at paras 33, 67.
82. *Ibid* at para 71, citing David Hayton "Unique Rules for the Unique Institution, the Trust" in Simone Degeling & James Edelman, eds, *Equity in Commercial Law* (Sydney: Lawbook Co, 2005).

Hayton argued that "where a bare trust is mere incidental machinery in the furtherance of a contractual agreement it seems that there are sufficient policy reasons to oust traditional trust law principles as to consequential losses".[83] It might therefore be arguable that the Supreme Court in *AIB* was essentially focused upon the remedies that flow from a breach of a bare, commercial trust. If so, the traditional approach towards falsification may be maintained in the context of traditional trusts. After all, the departure from the general, traditional approach in one particular area does not necessarily undermine the more general rule. Yet the tenor of the judgment in *AIB* suggests that the same rules should apply regardless of whether a "commercial" or "traditional" trust is at issue.[84] Indeed, the contrary would be somewhat strange: a lay trustee acting gratuitously for a traditional trust might be subject to more stringent liability than a professional who is paid for his services. This may be specially odd since professional trustees are more likely to enjoy the benefit of an exemption clause.[85] It is suggested that *AIB* is likely to lead to a focus upon loss caused by a breach of trust, regardless of the type of trust at issue. There is perhaps a certain irony in this conclusion. Some commentators forcefully argued that no distinction should be drawn between commercial trusts and traditional trusts in order to maintain the traditional approach of falsification in all contexts.[86] Those arguments against a commercial-traditional trust divide are persuasive, but are now likely to be used to focus attention upon equitable compensation for loss caused in respect of breaches of all types of trust.

IV. Principles of Equitable Compensation

The scope and meaning of "equitable compensation" remains unclear. It was used in *FHR* to cover a personal claim for an account of profits

83. Hayton, *supra* note 82 at 305; see also *Akai*, *supra* note 6.
84. Peter Turner, "The New Fundamental Norm of Recovery for Losses to Express Trusts" (2015) 74:2 Cambridge Law Journal 188.
85. See *e.g. Armitage v Nurse*, [1998] Ch 241 (CA (Civ)(Eng)); *Walker v Stones*, [2001] QBD 902 (CA (Civ)(Eng)).
86. See *e.g.* Millett, *supra* note 20 at 224-25.

made in breach of fiduciary duty.[87] That remedy is probably best seen as restitutionary or gain-based rather than compensatory, since the claimant need not suffer any loss.[88] Equitable compensation should exclusively denote loss-based claims. It is important to consider the principles that inform an award of equitable compensation.

It appears that similar principles might now apply to both "falsification" and "surcharge" cases. In *Target,* Lord Browne-Wilkinson relied upon cases of surcharge in the context of a claim which was really based upon falsification,[89] and the elision between the two appears to now be complete since Lord Toulson concluded that "in a practical sense both are reparative compensation".[90] It will therefore be important to consider the principles underpinning "reparative compensation" or "surcharge" since they appear to underpin the award of equitable compensation for breach of trust more generally.

Beneficiaries have often sought to surcharge the trust fund where the trustee has failed to comply with his duties of care, meaning that the fund is not worth as much as it would have been had the trustee not breached his duties. At times, the principles of surcharge have developed by reference to the principles of falsification.[91] For instance, in *Re Mulligan* Justice Panckhurst cited cases such as *Re Dawson* before saying:

> I accept that the obligation to make restitution imposed on defaulting trustees and fiduciaries is more absolute than the common law obligation to pay damages for tort or breach of contract, and the considerations of causation, foreseeability or remoteness are not of great moment […] [92]

However, the language of restitution or restoration seems inapt in the context of reparative compensation. As the editors of Meagher, Gummow and Lehane have pointed out: "it stretches belief to speak of restoration — suggesting restoration to a *prior* position — when the relief is designed to place the trust or the fiduciary's principal in the *presently*

87. See *e.g. FHR*, *supra* note 50 at para 1, per Lord Neuberger.
88. *Cf. FHR*, *supra* note 50 at para 120, per Lord Reed.
89. See *e.g. Nestle*, *supra* note 10; *Bartlett Nos 1 and 2*, *supra* note 61.
90. *AIB*, *supra* note 1 at para 54.
91. Mitchell, *supra* note 63 at 320-27.
92. *Mulligan*, *supra* note 10 at 507-509.

correct position".[93] As Lord Justice Brightman put it in *Bartlett v Barclays Bank Trust Co Ltd (Nos 1 and 2)*, "the so-called restitution which the defendant must now make … is in reality compensation for loss suffered by the plaintiffs".[94]

It therefore does not seem appropriate to look across to the falsification cases based upon an "equitable debt" when developing principles of equitable compensation in the context of surcharge. The better view is that principles of reparative compensation have traditionally been considered to be distinct from those underpinning falsification. Yet in the wake of *AIB,* it would seem that the principles underpinning reparative compensation are relevant to equitable compensation more generally, even in the context of misapplied trust property.

In *Bristol & West Building Society v Mothew*[95] ("*Mothew*") Lord Justice Millett said:

> [e]quitable compensation for breach of the duty of skill and care resembles common law damages in that it is awarded by way of compensation to the plaintiff for his loss. There is no reason in principle why the common law rules of causation, remoteness of damage and measure of damages should not be applied by analogy in such a case.[96]

This now seems to be the orthodox approach in England and Wales, even though in *AIB* Lord Reed noted that this "dictum has been questioned, or given a restrictive application, in a number of other jurisdictions".[97] However, it would seem to follow from this passage that the common law rules on damages are only applied by analogy. The equitable rules exist independently of their common law counterparts. Indeed, there is obviously no monolithic concept of the common law rules on damages.[98] The principles differ depending on the nature of the wrong involved — breach of contract, the tort of negligence, and the tort of deceit all have different rules, for example — and it is important to be clear to

93. Heydon, *supra* note 4 at 23-170.
94. *Bartlett Nos 1 and 2*, *supra* note 61 at 545.
95. *Mothew*, *supra* note 10.
96. *Ibid* at para 17.
97. *AIB*, *supra* note 1 at para 119.
98. *Ibid*.

what common law wrong an analogy should be made when elucidating equitable principles.

Where the breach of duty in question concerns the failure to exercise the necessary degree of care and diligence, the view of Millett LJ that a stricter approach of equity is not required has some attraction.[99] However, the drawing of an analogy with common law notions has not been met with universal acceptance. For instance, in *Youyang,* the High Court of Australia said:

> [h]owever, there must be a real question whether the unique foundation and goals of equity, which has the institution of the trust at its heart, warrant any assimilation even in this limited way with the measure of compensatory damages in tort and contract. It may be thought strange to decide that the precept that trustees are to be kept by courts of equity up to their duty has an application limited to the observance by trustees of some only of their duties to beneficiaries in dealing with trust funds.[100]

As McLachlin J insisted in *Canson,* the relationship at issue has "trust, not self-interest, at its core, and when the breach occurs, the balance favours the person wronged".[101] Given the control the trustee has over the beneficiary's property, leading to a sense of "vulnerability" about the beneficiary, there are strong arguments in favour of stricter rules in equity which might be employed in order to protect further the beneficiary.[102]

Upon taking an account, the focus was very clearly upon the state of the trust fund. This is why it has been said that "the relevant loss is the loss suffered by the trust estate or the trust fund, not by the beneficiaries or objects as such".[103] Yet in *Mothew*, Millett LJ stated that "[e]quitable compensation for breach of the duty of skill and care resembles common law damages in that it is awarded by way of compensation *to the plaintiff for his loss*".[104] In a similar vein, in *Target,* Lord Browne-Wilkinson appeared

---

99. *NZ Guardian Trust, Permanent Building Society (in liq) v Wheeler* (1994), 11 WAR 187 (HCA), per Ipp J.
100. *AIB*, *supra* note 1 at para 39.
101. *Canson*, *supra* note 7 at 543. See Heydon, *supra* note 4 at 23-265, 23-350.
102. See Part V, below.
103. Heydon, *supra* note 4 at 23-360. See *e.g. Re Dawson*, *supra* note 11; *Salmon*, *supra* note 19 at 371.
104. *Mothew*, *supra* note 10 at 17 [emphasis added].

to state that compensation should be paid to the beneficiary directly; this approach was endorsed in *AIB.*[105] Such comments might be limited to circumstances where there is a bare trust with a sole beneficiary, but could conceivably lead to a focus on a particular beneficiary's loss, including consequential losses.[106] This would be another significant development brought about by the *Target* line of cases. However, it is suggested that the view expressed by Lord Reed in *AIB* — that the same remedy should be available regardless of whether an account is taken or short-circuited by "equitable compensation"[107] — tends to indicate that the focus should still be on loss suffered by the trust fund, even though analogies drawn with the common law may lead to different outcomes.

In any event, it seems sensible to consider how the requirements recognised to be relevant to compensation at common law might apply in equity.[108] As Lord Steyn observed in *Smith New Court Securities Ltd v Citibank NA*[109] ("*Smith New Court*"):

> [i]t is now necessary to consider separately the three limiting principles which, even in a case of deceit, serve to keep wrongdoers' liability within practical and sensible limits. The three concepts are causation, remoteness and mitigation. In practice the inquiries under these headings overlap. But they are distinct legal concepts.[110]

In addition, considerations such as contributory negligence and the nature of recoverable loss will be considered.

A. Concurrent Liability

Most claims for breach of trust will be "stand-alone" claims in the sense that the only possible claim a beneficiary has against the trustee will be for breach of trust. But it is possible for there to be concurrent claims in contract and tort as well. Indeed, there was such concurrent liability on

105. See *e.g. AIB*, *supra* note 1 at para 91, per Lord Reed.
106. Jamie Glister, "Breach of Trust and Consequential Loss" (2014) 8:3 Journal of Equity 235.
107. See *e.g. AIB*, *supra* note 1 at para 91; *Libertarian*, *supra* note 47. *Target, supra* note 2.
108. See *e.g. AIB*, *supra* note 1 at para 81, per Lord Reed.
109. [1997] AC 254 (HL) [*Smith New Court*].
110. *Ibid* at 284.

the facts of *AIB.* Nevertheless, it is important that the equitable principles of compensation should not be distorted by considerations regarding concurrent liability, when in most instances concurrent liability simply will not arise.

On one interpretation of the judgment of Lord Neuberger in *Akai,*[111] along with that of Lord Toulson in *AIB*,[112] it may be that the equitable rules of compensation should be the same as those which apply at common law. On another view, the thrust of their Lordships' reasoning may be that, in the commercial context at least, the claimant should not recover more through an equitable claim than would be available at common law. In some respects, this is unsurprising. Even in the context of concurrent claims in contract and tort, there are strong calls for the contractual rules to trump the tortious rules, since the parties were not strangers and had the opportunity to negotiate as is the case in any contractual relationship.[113] Following this approach, there may be a hierarchy within the law of obligations which could determine which set of rules should apply to a claim for compensation on any given set of facts. At the pinnacle would be contractual rules, and these should govern the particular dispute, regardless of whether the same facts are then framed to ground a claim in tort or equity.[114]

However, this type of approach does not seem likely to prevail. Orthodoxy currently insists that where the claimant has a choice about whether to sue in contract, tort or equity, he can exercise that choice freely, taking into account which claim is likely to benefit him

111. *Akai, supra* note 7.
112. *Canson, supra* note 6.
113. See *e.g.* Andrew Burrows, *Remedies for Torts and Breach of Contract*, 3d (Oxford: Oxford University Press, 2004) at 88-94; Andrew Burrows, "Comparing Compensatory Damages in Contract and Tort: Some Problematic Issues" in Simone Degeling, James Edelman & James Goudkamp, eds, *Torts in Commercial Law* (Sydney: Thomson Reuters, 2011) 3 at 3-7. And see also the recent decision of the Court of Appeal in *Wellesley Partners LLP v Withers LLP,* [2015] EWCA Civ 1146.
114. *Cf.* Hayton, *supra* note 82.

the most.[115] This was recognised by Lord Reed in *AIB*,[116] and seems well-established.[117] Indeed, it was also accepted by Lord Neuberger in *Akai*.[118] It is suggested that the real issue is the extent to which common law notions should influence equitable reasoning. This is how best to understand the following remarks of Justice La Forest in *Canson*:

> I have no doubt that policies underlying concepts like remoteness and mitigation might have developed from an equitable perspective. However, given the paucity of authority in the field, it is scarcely surprising that courts will deal with a case falling properly within the ambit of equity as if it were a common law matter or as justifying the use of its mode of analysis.[119]

As has already been noted, *Canson* was a case concerning breach of fiduciary duty not breach of trust. But given the use of *Canson* and other cases which did not concern trusts in both *Target* and *AIB,* some of the comments made regarding breach of fiduciary duty may be exploited when determining the principles of compensation that apply in the trust context.[120] Indeed, in *Swindle v Harrison*[121] ("*Swindle*"), Lord Justice Mummery thought that "fiduciary duties are equitable extensions of trustee duties" and similar principles might apply.[122] If so, this is something of a departure from what was said by La Forest J in *Canson*:

> [w]e have been given no case where the principles applicable to trusts have been applied to a breach of a fiduciary duty of the type in question here, and for reasons already given, I see no reason why they should be transposed here. The harshness of the result is reason alone, but apart from this, I do not think that the claim for the harm resulting from the actions of third parties can fairly be looked upon as falling within what is encompassed in restoration for the harm

---

115. *Henderson v Merrett Syndicates Ltd*, [1995] 2 AC 145 (HL).
116. See *e.g. AIB*, *supra* note 1 at para 136.
117. Canson, *supra* note 7, per La Forest J ("[w]here concurrent liability lies in tort and contract and in equity, the appellants may sue in whatever manner they find most advantageous" at 565); see *Central Trust Co v Rafuse*, [1986] 2 SCR 147 at 206; *Bartlett Nos 1* and *2*, *supra* note 61 at 95-96.
118. *Akai*, *supra* note 6 at paras 130, 131.
119. *Canson*, *supra* note 7 at 580.
120. *Cf.* Mitchell, *supra* note 63; Penner, *supra* note 60.
121. [1997] 4 All ER 705 (CA) [*Swindle*]
122. *Ibid* at 723.

suffered from the breach.[123]

The departure may be justified by no longer seeking to "restore" the trust fund through an action in equitable debt, but rather awarding compensation for all breaches of trust. Concurrent liability at common law and equity has often arisen in the context of fiduciaries. In the leading case of *Nocton v Lord Ashburton*[124] ("*Nocton*"), Lord Chancellor Viscount Haldane examined the historical development of an action brought by a client against his solicitor for negligence in breach of a duty to be skilful and careful. After the development of the action of assumpsit, the Lord Chancellor thought "it probable that a demurrer for want of equity would always have lain to a bill which did no more than seek to enforce a claim for damages for negligence against a solicitor".[125] However, it is important to appreciate the limits of such comments. They only concern breach of a duty of skill and care. They do not extend to breaches of loyalty which are peculiarly fiduciary.[126]

Shortly before his judgment in *Target,* Lord Browne-Wilkinson had already expressed the view that where the claimant sought a remedy for professional negligence, the law ought to arrive at the same conclusion regardless of whether the claim was brought for breach of contract, tort, or an equitable duty of care.[127] The historical roots of the duties of care apparently made little difference to the compensatory remedies available.

However, some care should be exercised before wholeheartedly adopting such an approach to concurrent claims. As Lord Justice Evans pointed out in *Swindle,* claims in equity might be seen to differ from those in tort[128] because the aim in equity — traditionally, at least — is not to put the parties into the position they would have been in had no wrong occurred. Rather, in equity, the concern is to restore the claimant to the position he was in before the defendant committed the

123. *Canson*, *supra* note 7 at 580.
124. [1914] AC 932 (HL) [*Nocton*].
125. *Ibid* at 956.
126. *Mothew*, *supra* note 10.
127. See *e.g. Target*, *supra* note 2 at 205. See also Somer J in *Day v Mead*, [1987] 2 NZLR 443 (CA) at 458.
128. See *e.g. Livingstone v Rawyards Coal Co* (1880), 5 App Cas 25 (HL).

wrong. The different approach might be justified by the need to offer stronger protection to a beneficiary who was vulnerable under a trust and not inclined towards self-seeking behaviour.[129] As McLachlin J put it, "[i]n short, equity is concerned, not only to compensate the plaintiff, but to enforce the trust which is at its heart".[130]

It is suggested that there are good reasons why equitable rules might differ from those at common law.[131] In *Canson,* McLachlin J went on to say:

> Cooter and Freedman[132] go on to point out that because the fiduciary has superior information concerning his or her acts, it will be difficult to detect and prove breach of these wide obligations; and because the fiduciary has control based on the notion of implicit trust, there is a substantial potential for gain through such wrongdoing. This may justify more stringent remedies than for negligence or breach of contract. As Lord Dunedin put it in *Nocton* … at p. 963: "there was a jurisdiction in equity to keep persons in a fiduciary capacity up to their duty.[133]

The stricter approach in equity might extend not only to stripping wrongdoing trustees of their gains, but also to a stricter approach to compensation for loss. The key question to consider is the nature of any particular obligation owed.[134] As Lord Reed said in *AIB,* "[t]o the extent that the same underlying principles apply, the rules should be consistent".[135] This echoes the comments of McLachlin J in *Canson*: "[i]n so far as the same goals are shared by tort and breach of fiduciary duty, remedies may coincide. But they may also differ".[136] After all, equitable remedies are qualified in character in ways which are not paralleled in the common law. For instance, equitable remedies are always "discretionary"

---

129. *Canson*, *supra* note 7 at 543, per McLachlin J; *Vercoe v Rutland Fund Management Ltd,* [2010] EWHC 424 (Ch) at para 343, per Sales J.
130. *Canson*, *supra* note 7 at 543.
131. *AIB*, *supra* note 1 at para 137, per Lord Reed.
132. Robert Cooter & Bradley Freedman, "The Fiduciary Relationship: Its Economic Character and Legal Consequences" (1991) 66 New York University Law Review 1045.
133. *Canson*, *supra* note 7 at 543.
134. *AIB*, *supra* note 1 at paras 92-93, per Lord Reed.
135. *Ibid* at para 138.
136. *Canson*, *supra* note 7 at 545.

and may be withheld on the basis of delay or laches, so there exists a system of checks and balances within the equitable system which is absent at common law. As Gummow has pointed out:

> [a]ny effort to import common law concepts must be considered against this background and with an awareness that the common law has developed with an appreciation that in a court of law we cannot impose terms on the party suing; if he be entitled to a verdict, the law must take its course.[137]

Indeed, in *AIB,* Lord Reed observed that "the liability of a trustee for breach of trust, even where the trust arises in the context of a commercial transaction which is otherwise regulated by contract, is not generally the same as a liability in damages for tort or breach of contract".[138] The rules of equitable compensation deserve distinct consideration.

V. Quantifying the Loss

A. Scope of Duty

The result in *AIB* might be explained on a "scope of duty" basis: the fall in the property market and probable over-valuation of the property were outside the scope of the solicitors' duty to the bank. This parallels the development of the "scope of duty" requirement in tort.[139] Indeed, in *Nationwide Building Society v Balmer Radmore*[140] ("*Balmer Radmore*"), Mr Justice Blackburne referred to a need to have regard to the "scope of the duty which was broken" when considering equitable compensation for breach of fiduciary duty.[141]

Use of a "scope of duty" analysis strengthens the link between equitable compensation and compensation at common law. However, the concept of "scope of duty" has proven to be very difficult at common

---

137. William Gummow, "Compensation for Breach of Fiduciary Duty" in Timothy Youdan, ed, *Equity, Fiduciaries and Trusts* (Toronto: Carswell, 1989) 75, citing *Deeks v Strutt* (1794), 5 Term Rep 690 (KB (Eng)) at 693.
138. *AIB*, *supra* note 1 at para 136. See also paras 84-85.
139. *South Australia Asset Management Corporation* v *York Montague Ltd*, [1997] AC 191 (HL) [*SAAMCO*].
140. [1999] PNLR 606 (Ch (Eng)) [*Balmer Radmore*].
141. *Ibid* at 671.

law.[142] It is a malleable concept, and judges are able to mould the scope of duty analysis as they see fit.[143] Although it now appears often to be considered as an element of the remoteness enquiry,[144] this is dubious: whereas the amount of damages is often capped at a foreseeable level by the rules of remoteness, the scope of duty analysis is only able to include or exclude types of loss in a binary fashion.[145] This seems unnecessarily rigid in the context of breaches of trust and fiduciary duty. Thus in *Caffrey v Darby*[146] the Master of the Rolls thought that any other approach:

> would be an encouragement to bad motives; and it may be impossible to detect undue motives. If we get the length of neglect in not recovering this money by taking possession of the property, will they be relieved from that by the circumstance, that the loss has ultimately happened by something, that is not a direct and immediate consequence of their negligence: viz. the decision of a doubtful question of law? Even supposing, they are right in saying, this was a very doubtful question, and they could not look to the possibility of its being so decided, yet, if they have been already guilty of negligence, they must be responsible for any loss in any way to that property: for whatever may be the immediate cause, the property would not have been in a situation to sustain that loss, if it had not been for their negligence. If they had taken possession of the property, it would not have been in his possession. If the loss had happened by fire, lightning, or any other accident, that would not be an excuse for them, if guilty of previous negligence. That was their fault.[147]

This tough approach might conceivably still be defended in the equitable sphere given the flexibility of the court to excuse the trustee from personal

---

142. Edwin Peel, "SAAMCO Revisited" in Andrew Burrows & Edwin Peel, eds, *Commercial Remedies: Current Issues and Problems* (Oxford: Oxford University Press, 2003) 55; Jane Stapleton, "Negligent Valuers and Falls in the Property Market" (1997) 113 Law Quarterly Review 1.
143. See *e.g. Aneco Reinsurance Underwriting Limited v Johnson & Higgins Limited*, [2001] UKHL 51.
144. *Cf. Transfield Shipping Inc v Mercator Shipping Inc* [*The Achilleas*], [2008] UKHL 48 [*The Achilleas*]; see Andrew Burrows, "Lord Hoffmann and Remoteness in Contract" in Paul Davies & Justine Pila, eds, *The Jurisprudence of Lord Hoffmann* (Oxford: Hart Publishing, 2015).
145. Mark Stiggelbout, "Contractual Remoteness, 'Scope of Duty' and Intention" (2012) Lloyd's Maritime and Commercial Law Quarterly 97.
146. (1801), 6 Ves Jr 488 (Ch (Eng)).
147. Ibid at 495-96.

liability under section 61 of the *Trustee Act* 1925[148] ("*Trustee Act*"). But it seems unlikely that such a strict approach is consistent with the tenor of the decision of the Supreme Court in *AIB.* However, this does not mean that all the problems with scope of duty should be assimilated into equitable compensation: it is an unnecessary complication. Appropriate outcomes could be reached by instead relying on principles of causation and remoteness.[149] Where the breach of duty is deliberate there is less reason to restrict the scope of duty at common law,[150] and the higher standards generally demanded of trustees might suggest that this should be mirrored in equity.

B. Cost of Cure

An interesting question arises about whether the principles of equitable compensation require the trustee to compensate the trust fund (or possibly beneficiary) for the diminution in value suffered as a result of the breach of duty or for the cost of cure in repairing the breach of duty. In *Brudenell-Bruce v Moore & Cotton*[151] ("*Brudnell-Bruce*"), Justice Newey held that the answer to this question should reflect that given at common law. *Brudenell-Bruce* concerned the estate of the Earl of Cardigan. Lord Cardigan is the beneficiary of a bare trust administered by professional trustees. Lord Cardigan claimed that the trustees failed to maintain the Stable Block of Tottenham House, the seat of the Cardigan family. On the facts, the judge rejected the claim that the trustees had acted in breach of trust, but nevertheless went on to consider what the appropriate remedy would have been had there been a breach of trust.

The beneficiary argued that the full cost of repair should be awarded, contending that even if "it is going to cost £5 million to restore the Stable Block but fully restored it is only going to be worth £4 million, that is just the price that the trustees pay for allowing this collapse to have occurred in the first place".[152] Newey J rejected that argument. The judge

148. *Trustee Act*, 1925 (UK), 15 & 16 Geo V c 19, s 61.
149. *Cf.* Stapleton, *supra* note 142; Burrows, *supra* note 144.
150. *SAAMCO*, *supra* note 139 at 214.
151. [2014] EWHC 3679 (Ch) [*Brudenell-Bruce*].
152. *Ibid* at para 147.

could "see no reason why the Courts should be more willing to award compensation based on cost of reinstatement in circumstances such as those in the present case than they would be to measure damages in that way for breach of contract or a tort".[153] Newey J relied upon common law decisions such as *Ruxley Electronics and Construction Ltd v Forsyth*[154] and *In Southampton Container Terminals Ltd v Schiffahrtsgesellschaft "Hansa Australia" GmbH*[155] (*The "Maersk Colombo"*) to conclude that the full cost of cure measure should not be awarded where that would be unreasonable.

Equitable awards should not be unreasonable. But it is not obvious that this is an area where equity should "follow the law". The restrictions on the full cost of cure remedy in English law are controversial. For example, they have not been followed in Australia.[156] If what the claimant really wants is performance of the bargain or transaction, why should the courts not protect that performance interest?[157] This question might be thought to be particularly difficult to answer in the context of breach of trust. After all, it is important that equity hold trustees up to their primary obligations to perform the trust properly, which would suggest that the cost of reinstatement be the *prima facie* remedy available. Yet Newey J only thought that this "may be the case where equitable compensation is awarded as a substitute for performance of a trustee's obligation to deliver up trust assets in specie".[158]

Newey J relied upon *AIB* to justify his approach:

> [i]n AIB, Lord Reed explained that equitable compensation for breach of trust "aims to provide the pecuniary equivalent of performance of the trust" (paragraph 93) and that the measure of compensation for a breach of trust "will generally be based upon the diminution in the value of the fund caused by the trustee's default" (paragraph 94). The present case is, in my view, plainly one where, had a relevant breach of trust been established, it would have been

153. *Ibid* at para 152.
154. [1996] AC 344 (HL).
155. [2001] EWCA Civ 717.
156. *Tabcorp Holdings Ltd v Bowen Investments Pty Ltd*, [2009] HCA 8.
157. See generally James Edelman, "Money Awards for the Cost of Performance" (2010) 4 Journal of Equity 122.
158. *Brudenell-Bruce*, *supra* note 151 at para 151.

> appropriate to measure compensation by the resulting 'diminution in the value of the fund,' not by the cost of reinstating the Stable Block.[159]

It is not clear whether *AIB* fully supports such restrictions on the cost of cure remedy. Equitable compensation will generally be the diminution in value, because this will generally also equate the cost of cure remedy. But where the two measures differ — as in, for example, *Ruxley* and *Brudenell-Bruce* — then a choice has to be made, and Lord Reed's insistence on "the pecuniary equivalent of performance" might favour a cost of cure award. It is suggested that the cost of cure remedy has been unduly restricted at common law, and that equity should not be bound to follow suit.[160]

C. Presumption of Cheapest Means of Performance

At common law, it seems that when assessing the value of chances — and indeed of compensation generally — it should be borne in mind that the court should assess damages on the basis that the contract-breaker performed in a manner most advantageous to himself.[161] This is a sensible rule. It accords with the need for the defendant to protect its own position. Nevertheless, in *Durham Tees Valley Airport v BMI Baby*,[162] the Court of Appeal held that, where the defendant has a wide discretion about how to perform its obligations, the court should ascertain how the defendant would, in fact, have performed. That may have been a pragmatic decision on the facts of the case, given the difficulties involved in determining the minimum performance required (about operating flights from an airport), but it is suggested that the decision should be treated with some caution; the established rule that the defendant perform in the manner most advantageous to himself should be maintained at common law.

At common law, the defendant owes no duty to the claimant to look after the latter's interests. The same is not true in equity. There is therefore no room for a similar presumption to apply in equity. Indeed, the contrary presumption seems more appropriate; there should be a "presumption

159. *Ibid* at 155.
160. See *Elder's Trustee and Executor Co v Higgins* (1963), 113 CLR 426 (HCA) at 473 [*Elder's*].
161. *Lavarack v Woods of Colchester Ltd*, [1967] 1 QB 278 (CA (Eng)).
162. [2010] EWCA Civ 485.

that trust funds will be put to the most profitable use",[163] and the burden should shift to the trustee to prove that that is not the case.[164] At the very least, the court should have regard to what a reasonable, prudent trustee would have managed to obtain for the trust fund, rather than the very minimum which would have been achieved without there being a breach.[165]

D. Date of Assessment

The date of assessment of loss at common law is a matter of some controversy. Orthodoxy suggests that damages are assessed at the date of breach. This has the advantage of providing commercial certainty and allows the victim of the wrong to assess its losses immediately upon breach in order to determine what steps it should take to mitigate its loss and whether or not to settle and compromise its claim against the defendant.[166] This has recently been challenged,[167] but, in any event, the position in equity is clear and different from the traditional understanding of the common law. In equity, compensation is assessed at the date of judgment, not breach.[168] The breach of duty does not "stop the clock";[169] since the trustee must continue to act as a good trustee unless and until he is removed from office. The trustee is unable simply to breach his trust obligations and then walk away from his duties upon compensating the trust fund or beneficiary. It may be that at the date of breach no losses occurred, yet substantial losses have arisen by the date of judgment.[170] A "breach-date rule" would be inapt in such circumstances.

---

163. Canson, *supra* note 7 at 545, per McLachlin J; see also Heydon, *supra* note 4 at 23-260, 23-330.
164. *Mulligan*, *supra* note 10 at 508.
165. *Nestle*, *supra* note 10 at 1280, per Staughton LJ.
166. See *e.g. Golden Strait Corp v Nippon Yusen Kubishika Kaisha*, [2007] UKHL 12 at paras 10, 11, per Lord Bingham.
167. Andrew Dyson & Adam Kramer, "There is No Breach Date Rule: Mitigation, Difference in Value and Date of Assessment" (2014) 130 Law Quarterly Review 259.
168. *Target*, *supra* note 2 at 437; *AIB*, *supra* note 1 at para 140, per Lord Reed.
169. *Target*, *supra* note 2 at 437.
170. *Cf. Youyang*, *supra* note 69.

However, the general rule that losses are assessed at the date of judgment need not be rigidly applied. In some situations it may be appropriate to demand that the trustee pay compensation assessed on the basis of the highest intermediate value of the property which was improperly sold, for example.[171] This is justified because of the continuing duties owed by trustees, and flexible approach that equity should be prepared to adopt.

E. Causation

On a traditional approach, as has been seen above, causation was irrelevant to claims of falsification. Thus Justice Edelman has said that "when payment was sought following an account in common form there was a direct analogy with an order for specific performance or payment of a liquidated debt which was due. In each case it is no answer for the defendant to allege that the plaintiff had suffered no loss".[172] This helps to explain why the appeal was allowed in *Youyang*. The Court of Appeal had held that the beneficiary's claim should fail since "the acceptance of the defective deposit certificate was a breach of trust which nevertheless did not cause any loss of Youyang's funds".[173] The High Court of Australia, on the other hand, insisted that it was "not to the point"[174] that the conduct of third parties would have caused the loss anyway "because the loss of the trust funds occurred as soon as the trustee wrongly disbursed them".[175]

Although it has also been suggested that a strict approach to causation should be adopted in the context of equitable compensation for breach of

171. *Libertarian*, *supra* note 47 at para 171, per Lord Millett; see generally Glister, *supra* note 106 at 529-34. See also *Elder's*, *supra* note 160 at 473.
172. *Jackson (No 2)*, *supra* note 18 at para 337.
173. *Youyang*, *supra* note 69 at para 29.
174. *Ibid* at para 63.
175. *Ibid*.

fiduciary duty,[176] it has since become clear[177] that there is no "equitable by-pass" of the need to prove a causal link.[178] The same now appears to be true in the context of all claims for breach of trust. As Lord Reed pointed out in *AIB*:

> since the concept of loss necessarily involves the concept of causation, and that concept in turn inevitably involves a consideration of the necessary connection between the breach of duty and a postulated consequence (and therefore of such questions as whether a consequence flows 'directly' from the breach of duty, and whether loss should be attributed to the conduct of third parties, or to the conduct of the person to whom the duty was owed), there are some structural similarities between the assessment of equitable compensation and the assessment of common law damages.[179]

Indeed, Lord Toulson thought that a remedy that did not reflect the loss caused (or gain caused) by a breach of duty would be penal. It is therefore necessary to be clear about what causal link is required. In both *Canson*[180] and *Target*,[181] reference was made to some sort of "common sense" test of causation; yet, as Lord Reed commented in *AIB*, "[d]ifficult questions of causation do not however always have an intuitively obvious answer".[182] In some areas, equity has favoured rather claimant-friendly approaches. For instance, in the context of account of profits for breach of fiduciary duty, only "some reasonable connection" between the gain and the breach is required;[183] and a misrepresentation only needs to be one reason, but not necessarily a "but-for" reason, for the claimant's entering into the

176. *Brickenden* v *London Loan & Savings Co*, [1934] 3 DLR 465 (PC (Canada)).
177. At least in England and Wales: for comparative discussion see Jamie Glister, "Equitable Compensation" in Jamie Glister & Pauline Ridge, eds, *Fault Lines in Equity* (Oxford: Hart Publishing, 2012).
178. *Swindle*, *supra* note 121; see also *Gwembe Valley Development Co Ltd v Koshy (No 3)*, [2003] EWCA Civ 1048 at para 147.
179. *AIB*, *supra* note 1 at para 136.
180. *Canson*, *supra* note 7 at 163, per McLachlin J.
181. *Target*, *supra* note 2 at 439, per Lord Browne-Wilkinson.
182. *AIB*, *supra* note 1 at para 95.
183. *CMS Dolphin Ltd v Simonet*, [2001] 2 BCLC 704 (Ch (Eng)) at para 97, per Lawrence Collins J.

contract to ground a claim for rescission.[184] Nevertheless, *AIB* appears to demand the familiar "but-for" test,[185] and this seems sensible when focusing on a trustee's responsibility for loss.[186] It is clearly insufficient that the wrongdoing trustee simply provided an opportunity for the loss to occur; the trustee must cause the loss.[187] However, as in *AIB*, this may lead to different outcomes than the traditional approach. Ho has given the example of a trustee who wrongfully disposes of a seaside bungalow shortly before a tsunami would have destroyed it in any event.[188] Traditionally, a beneficiary would still be able to falsify the wrongful misapplication of trust property. Yet it is difficult to see how the wrongful act of the trustee causes the beneficiary's loss, when that loss would have been suffered in any event.[189]

In *Libertarian*, Ribeiro PJ held that:

> [w]here the plaintiff provides evidence of loss flowing from the relevant breach of duty, the onus lies on a defaulting fiduciary to disprove the apparent causal connection between the breach of duty and the loss (or particular aspects of the loss) apparently flowing therefrom.[190]

---

184. See *e.g. Attwood v Small*, [1838] 6 Cl & F 232 (HL); *Reynell v Spyre* (1852), 1 De G M & G 660 (Ch (Eng)) at 708, per Cranworth LJ.
185. See *e.g. AIB*, *supra* note 1 at paras 73, 132, per Lord Toulson & Lord Reed. See also *Target*, *supra* note 2 at 431 (this is perhaps what Lord Justice Patten meant when he spoke of "a proper causal connection between the breach and eventual loss" in the Court of Appeal in *AIB Group (UK) v Mark Redler & Co Solicitors*, [2013] EWCA Civ 45 at para 47).
186. *Cf.* John Cartwright, *Misrepresentation, Mistake and Non-Disclosure,* 3d (London: Sweet & Maxwell, 2012) at 5-38; Ken Handley, "Causation in Misrepresentation" (2015) 131 Law Quarterly Review 275.
187. *Swindle*, *supra* 121 at 727, per Hobhouse LJ.
188. Ho, *supra* note 8 at 217.
189. *Ibid,* Ho recognises that this result might not be desirable, and suggests that "the court will need to adjust the causal test to deal with multiple sufficient causes such as these". It is not clear how this should be done, and given the complexities of causation at common law it is unlikely that causation in equity will prove to be simple.
190. *Libertanan*, *supra* note 47 at para 93. See too *Stevens v Premium Real Estate Ltd*, [2009] NZSC 15 at para 85 [*Stevens*].

Given the control over the trust property that the trustee enjoys, and the consequent difficulties that a beneficiary faces when seeking to establish and prove a breach of duty, it seems entirely appropriate to put the onus on the trustee to disprove an apparent causal connection.[191] However, it will not be sufficient for a trustee to show that if he had not committed a breach of trust the same loss would have been caused by some other third party's dishonest conduct.[192] Moreover, it is difficult to see much scope for the principle of *novus actus interveniens* in the context of equitable compensation, since rarely will anything happen to trust property which is truly independent of a breach of the duty to safeguard it. For instance, imagine that one trustee carelessly allows trust property to come exclusively under the control of another trustee. [193] The latter then misappropriates the trust assets. The first trustee, who only breached a duty of care, is nonetheless liable for all losses suffered, even though the immediate cause of the loss is the latter trustee's misappropriation of the trust assets.

F. Remoteness

Compensation requires some rules of remoteness. Whilst issues of remoteness are irrelevant to actions for an agreed upon sum,[194] equitable compensation must establish principles of remoteness. If the "the relentless contractualisation of trust law"[195] continues apace, and the principles of equitable compensation mirror the contractual principles,[196] then cases such as *Hadley v Baxendale*[197] and *C.f. Transfield Shipping Inc v Mercator Shipping Inc*[198] might be thought to be relevant in equity. But that is surely misguided. The point of the contractual rules is that

191. See also *Re Brogden* (1888), 38 Ch D 546 (CA)(Eng)) at 567-68, 572-73.
192. *AIB*, *supra* note 1 at para 58, per Lord Toulson. *Cf. Youyang*, *supra* note 69 at 23-170.
193. Heydon, *supra* note 4 at [23-340].
194. *Re Dawson*, *supra* note 11 at 215, per Street J.
195. Getzler, *supra* note 76 at 257.
196. *Ibid*.
197. (1854), 9 Exch 341 (Eng).
198. *The Achilleas*, *supra* note 144.

there exists an agreement between the parties and when making that agreement each party could bring the risk of certain losses to the other party's attention. Yet in the trust context, the beneficiary may not always have a contract with the trustee. Additionally, the trustee at the time of the dispute may not be the same person that originally agreed to the terms of the trust instrument. In most commercial trusts, admittedly, there will be a concurrent contractual claim, but that contractual claim should stand apart. It may be that the contractual claim should trump the equitable claim, but the equitable claim — and certainly any free-standing equitable claim — should not adopt the contractual principles of remoteness. Under the contractual approach, foreseeability is assessed at the date of entering into the contract.[199] But given the higher standard expected of trustees, and the different situations that can evolve over the course of a trust relationship, it is surely more appropriate for any foreseeability requirement to be assessed at the date of breach.

If an analogy is to be drawn to the common law, it would be more sensible to look across to tort law. There is a split between the "reasonable foreseeability" approach of *Overseas Tankship (U.K.) Ltd v Morts Dock and Engineering Co. Ltd*[200] (*The Wagon Mound*) in the tort of negligence and the "direct" approach[201] adopted in the context of the intentional torts, such as deceit.[202] It would be possible for equity similarly to adopt different approaches depending on whether or not the breach of duty was deliberate, and this differentiation may be evolving in the context of breach of fiduciary duty.[203] Given the higher standards demanded in equity, there is a strong argument for a stricter approach to be taken for all breaches of equitable duty.[204] As McLachlin J commented in *Canson*:

199. See *e.g. Jackson v Royal Bank of Scotland*, [2005] UKHL 3.
200. [1961] AC 388 (PC (Austl)).
201. *Re Polemis & Furness, Withy & Co Ltd*, [1921] 3 KB 560 (CA (Eng)).
202. *Smith New Court*, *supra* note 109.
203. See the differing judgments in *Swindle*, *supra* note 121. For a common law analogy in the tort of conversion, see *Kuwait Airways Corp v Iraqi Airways Co (Nos 4 and 5)*, [2002] UKHL 19 at paras 100-104, per Lord Nicholls.
204. *Cf.* Steven Elliott, "Remoteness Criteria in Equity" (2002) 65:4 Modern Law Review 588.

> [i]n negligence, we wish to protect reasonable freedom of action of the defendant, and the reasonableness of his or her action may be judged by what consequences can be foreseen. In the case of a breach of fiduciary duty, as in deceit, we do not have to look to the consequences to judge the reasonableness of the actions. A breach of fiduciary duty is a wrong in itself, regardless of whether a loss can be foreseen. Moreover the high duty assumed and the difficulty of detecting such breaches makes it fair and practical to adopt a measure of compensation calculated to ensure that fiduciaries are kept 'up to their duty'.[205]

This passage was cited with approval in *Libertarian*,[206] and surely applies equally to breach of trust. Indeed, *Canson* was a case where the claim failed because the losses suffered were too remote from the breach of fiduciary duty on any test.[207]

In *AIB,* Lord Reed said that "the foreseeability of loss is generally irrelevant, but the loss must be caused by the breach of trust, in the sense that it must flow directly from it".[208] It is suggested that foreseeability of loss should be irrelevant in the context of the misapplication of trust property.[209] The risk of unforeseeable consequential loss should be visited upon the wrongdoing fiduciary rather than the vulnerable beneficiary.[210] Perhaps foreseeability may be relevant where a duty to take reasonable care has been breached such that — in traditional language — the

---

205. *Canson*, *supra* note 7 at 553. See also *Stevens*, *supra* note 191 paras 24, 34, per Elias CJ.
206. *Libertarian*, *supra* note 47 at para 80.
207. *Canson*, *supra* note 7 at 590, per Stevenson J.
208. *AIB*, *supra* note 1 at para 135.
209. *Clough v Bond*, (1838), 3 My & C 490 (Ch (Eng)) at 496, per Cottenham LC; *Re Dawson*, *supra* note 11 at 215, per Street J; *Canson*, *supra* note 7 at 555-56, per McLachlin J; *Target*, *supra* note 2 at 438-39, per Lord Browne-Wilkinson; *Bank of New Zealand*, *supra* note 59 at 687, per Tipping J.
210. *Cf. Smith New Court*, *supra* note 109.

account could be surcharged,[211] and this may explain the qualification of "generally" in Lord Reed's statement. This depends upon how the nature of the trustee's obligation is explained. If it is akin to a duty to take care in tort, then foreseeability should be relevant. If a higher standard is demanded of the trustee as a fiduciary, then foreseeability should not limit the recoverable losses.

G. Mitigation

Mitigation is a general principle that might reduce the amount of compensation a claimant can recover. It even applies in the context of deceit. As Lord Steyn commented in *Smith New Court*: "[t]he third limiting principle is the duty to mitigate. The plaintiff is not entitled to damages in respect of loss which he could reasonably have avoided. This limiting principle has no special features in the context of deceit".[212] The victim of a wrong must not act in an unreasonable manner which would exacerbate his losses.

A common justification for mitigation is that it helps to avoid waste and promote efficient outcomes. Concerns of efficiency may be more important in the context of contracts than trusts. This might explain why McLachlin J said in *Canson* that "[t]he plaintiff will not be required to mitigate, as the term is used in law, but losses resulting from clearly unreasonable behaviour on the part of the plaintiff will be adjudged to flow from that behaviour, and not from the breach".[213] However, the essence of this difficult passage is essentially to introduce a requirement

211. See *e.g. Mothew*, *supra* note 10 at para 17, per Millett LJ; *Libertarian*, *supra* note 47; *cf. Youyang*, *supra* note 69 at para 39. See further Darryn Jensen, "Compensation for Breach of Trust — the Remoteness Impasse" in Charles Rickett, ed, *Justifying Private Law Remedies* (Oxford: Hart Publishing, 2008); Joshua Getzler, "Am I My Beneficiary's Keeper? Fusion and Loss-Based Fiduciary Remedies" in Simone Degeling & James Edelman, eds, *Equity in Commercial Law* (Sydney: Lawbook Co, 2005); John Heydon, "Are the Duties of Company Directors to Exercise Care and Skill Fiduciary?" in Simone Degeling & James Edelman, eds, *Equity in Commercial Law* (Sydney: Lawbook Co, 2005).
212. *Smith New Court*, *supra* note 109 at 285.
213. *Canson*, *supra* note 7 at 554.

that the beneficiary acted reasonably after learning of the breach of trust, which seems tantamount to mitigation.[214] This view gains some support from Lord Reed in *AIB*,[215] and is consistent with the majority view in *Canson*.[216] Once the beneficiary knows of the breach of trust, he should take reasonable steps to minimise his loss.

H. Contributory Negligence

The role of contributory negligence in equity is unclear.[217] This issue has mainly been discussed in the context of breach of fiduciary duty. Different jurisdictions have adopted different approaches. In *Day v Mead*,[218] the New Zealand Court of Appeal reduced the amount a principal could recover because he was partly the author of his own loss. Sir Robin Cooke thought this was the "obviously just course, especially now that law and equity have mingled or are interacting".[219] This approach to "fusion" is controversial. Other courts have not been so willing to recognise the influence of the common law in this area. For example, in *Pilmer v Duke Group Ltd*,[220] the High Court of Australia cited with approval the comments of McLachlin J in *Norberg v Wynrib*[221] which recognised the "conceptual and functional uniqueness" of fiduciary obligations, particularly since "one party exercises power on behalf of another and pledges himself or herself to act in the best interests of the other".[222] The High Court further noted, "the severe conceptual difficulties in the path of acceptance of notions of contributory negligence as applicable to diminish awards of equitable compensation for breach of

214. Smith, *supra* note 56 at 368.
215. *AIB*, *supra* note 1 at para 135.
216. *Canson*, *supra* note 7 at 581, LaForest J; Derk Davies, "Equitable Compensation: Causation, Foreseeability and Remoteness" in Donovan Waters, ed, *Equity, Fiduciaries and Trusts* (Toronto: Carswell, 1993) 317.
217. See generally *Balmer Radmore*, *supra* note 140 at 672-77, per Blackburne J and the academic commentary cited at 676.
218. [1987] 2 NZLR 443 (CA) [*Day*].
219. *Ibid* at 451.
220. [2001] HCA 31 [*Pilmer*].
221. [1992] 2 SCR 226 at 272.
222. *Pilmer*, *supra* note 220 at para 71.

fiduciary duty"[223] and that "the attempt to push common law notions of contributory negligence, as now modified by statute, into equitable remedies collapses in the face of insurmountable obstacles".[224]

In England and Wales, one obstacle is the *Law Reform (Contributory Negligence) Act 1945*[225] ("*1945 Act*"). That makes no provision for the operation of a defence of contributory negligence in the context of equitable claims. If the statute is to operate at all in this area, it could only be on the same basis as *Forsikringsaktieselskapet Vesta v Butcher*[226] ("*Vesta*") applies to claims in breach of contract, such that where an equitable duty of care overlaps with a tortious duty of care, then in those circumstances only might contributory negligence be a defence. This would be a very limited category of case and may mean that a trustee would be in a better position if able to establish a breach of a tortious duty of care as well a (potentially strict) duty imposed by the trust.[227]

It is suggested that the *1945 Act* does not provide any basis for a defence of contributory negligence in equity. Nor does section 61 of the *Trustee Act*, concerning the power of the court to relieve trustees from personal liability for breach of trust,[228] encompass contributory negligence. At first instance in *Markandan & Uddin,*[229] Roger Wyand QC, sitting as a Deputy High Court Judge, noted that there was nothing in *Vesta* which indicated that contributory negligence should be extended to instances of breach of trust.[230] The judge then went on to state that section 61 of the *Trustee Act* "could have provided for the conduct of

---

223. *Ibid* at para 86, per McHugh, Gummow, Hayne & Callinan JJ.
224. *Ibid* at para 174, per Kirby J.
225. (UK), 8 & 9 Geo VI, c 28.
226. [1989] AC 852 (HL) [*Vesta*].
227. A similar oddity arises in the contractual context, since a defendant who breaches a strict contractual duty is actually better off (since the defence of contributory negligence applies) if he can establish that he also breached a tortious duty of care. Yet the innocent party will seek to avoid showing that the defendant acted negligently, in order to escape the ambit of contributory negligence.
228. See Part X below.
229. [2010] EWHC 2517 (Ch) [*Markandan*].
230. *Vesta*, *supra* note 226.

the beneficiary to be taken into account as the Defendant here wishes. It did not and it is not for the Court to extend the law in a way that was not done by the legislature".[231] The judge's conclusions on contributory negligence were not pursued on appeal,[232] and should be supported: the focus of section 61 is firmly placed on the trustee who has committed the breach of trust, not the beneficiary. Similarly, section 62 of the *Trustee Act*, concerning the power of the court to make the beneficiary indemnify the trustee for breach of trust, only applies if the beneficiary instigated or consented to the breach of trust, which does not cover the same ground as contributory negligence.

If contributory negligence is truly to be accepted in equity, then this is perhaps best explained on the basis that equity seeks a just result, and "he who seeks equity should do equity". It may be that insisting that a beneficiary is entitled to trust his trustee absolutely, without accepting any responsibility to safeguard his own position at all, is too harsh an approach in some instances (particularly in the commercial sphere). This may explain Lord Reed's comment that "losses resulting from unreasonable behaviour on the part of the claimant will be adjudged to flow from that behaviour and not from the breach".[233] This appears to present trustees with an opportunity to raise arguments on contributory negligence. Where a breach of trust is concerned, it could perhaps be expected that such arguments will find favour. After all, "the line between failure to take due care and mitigation could in some instances become a fine one and mitigation is certainly a defence in fiduciary law".[234]

However, this approach may not be universally popular. Indeed, Fleming has described contributory negligence as "an adventure of the common law which represents one of its outstanding failures"[235] and the unsatisfactory nature of the defence led Gummow to argue that it should not be recognised in equity: "[a]ll this suggests the unwisdom [of entangling the already complex law as to fiduciary duties with notions

231. *Markandan*, *supra* note 229 at para 42.
232. *Markandan & Uddin*, [2012] EWCA Civ 65 at para 7.
233. *AIB*, *supra* note 1 at para 135.
234. Davies, *supra* note 216 at 317.
235. *Law of Torts* 2d (Sydney: Law Book, 1961) 214.

of contributory negligence".[236] In England and Wales, it appears that contributory negligence is not a defence to claims for breach of fiduciary duty. The headnote to *Balmer Radmore* states that "no such reduction could be made where a breach of fiduciary duty was shown. Breach of fiduciary duty was not covered by the provisions of the 1945 Act".[237] However, this is perhaps slightly misleading. The *ratio* of the decision of Blackburne J is restricted to the rejection of contributory negligence in instances of "conscious disloyalty".[238] Blackburne J reasoned that even at common law, contributory negligence is no defence to an intentional tort, and in equity, it was equally important "to keep persons in a fiduciary capacity up to their duty".[239] But it is not yet entirely clear whether the same approach would be taken where the fiduciary committed a breach of duty negligently rather than deliberately. The need to hold such parties up to a higher standard might suggest that no defence of contributory negligence should be available.

In any event, even if contributory negligence were to be recognised in the equitable sphere, it will be very difficult to satisfy the court that the defence has been made out. As Justice Somers commented in *Day v Mead*:

> [o]f course, before reducing an award on the ground that the claimant has been partly the author of his or her own loss, the Court will have to give much weight to the well-established principle that, largely for exemplary purposes, high standards are expected of fiduciaries. A strong case is needed to relieve the fiduciary of complete responsibility.[240]

VI. Section 61 of *The Trustee Act* 1925

The fact that liability for breach of trust is strict makes it likely that claims for breach of trust will continue to prove attractive to claimants. Unlike common law claims for breach of a contractual or tortious duty of care, the claimant beneficiary does not need to prove that the trustee acted

236. Gummow, *supra* note 137 at 86.
237. *Balmer Radmore*, *supra* note 140 at 610.
238. *Ibid* at 676-77.
239. *Ibid* at 677, citing *Nocton*, *supra* note 124 at 963, per Lord Dunedin.
240. *Day*, *supra* note 218 at 452.

negligently in failing to comply with the terms of the trust instrument.[241] Instead, the onus is clearly placed on the trustee to prove that he or she acted reasonably. This is obviously advantageous to the beneficiary, and it seems appropriate. After all, the trustee undertakes to act in the best interests of the beneficiary, and "[t]he beneficiary may not be in a position to know all that occurred in the chain of events leading to the breach of trust".[242] In some instances, holding the trustee liable when he or she acted wholly reasonably may be too harsh; this may explain the existence of section 61 of the *Trustee Act*. There is no counterpart of section 61 at common law.

Section 61 provides the courts with a statutory discretion to relieve trustees from personal liability for breach of trust where the trustee "has acted honestly and reasonably, and ought fairly to be excused for the breach of trust and for omitting to obtain the directions of the court in the matter in which he committed such breach". Although, in *AIB,* Lord Toulson thought that section 61 might be used as a "*deus ex machina*",[243] it is suggested that relief under section 61 is guided by recognised principles.[244] In any event, the burden is clearly placed on the trustee to prove that he or she acted honestly and reasonably and ought fairly to be excused for the breach of trust.[245] This equitable approach has further advantages over common law claims if the beneficiary wishes to sue third parties to the breach of trust. This is because section 61 is personal to the trustee: it excuses the liability of the trustee, but does not mean that there

241. *Santander UK v RA Legal Solicitors*, [2014] EWCA Civ 183, per Briggs LJ ("[i]t is precisely because of the strictness of the solicitor's trust liability that lenders which are the victims of such fraud prefer to base their claims for recovery upon breach of trust rather than breach of contract or negligence, because both of those causes of action generally require the lender to prove that the solicitor has been guilty of a breach of a duty of care" at para 19) [*Santander*].
242. *Ibid* at para 111, per Count Etherton.
243. *AIB*, *supra* note 1 at para 69, per Lord Toulson.
244. See Paul Davies, "Section 61 of the Trustee Act 1925: *deus ex machina*?" (2015) 6:5 The Conveyancer and Property Lawyer 379-94.
245. *Santander*, *supra* note 241 at para 55, per Briggs LJ.

has been no breach of trust.[246] Thus a third party might still be liable for dishonestly assisting that breach of trust, or for knowingly receiving property in breach of trust. For example, in *Re Smith*,[247] the trustee was a widow who lived in the country and employed a firm of solicitors to act as her agents. The solicitor's clerk fraudulently obtained the trustee's signature on certain cheques and induced her to initial alterations to the cheques. He then absconded with the money. The trustee was held to have committed a breach of trust, but her liability was excused. If the clerk had been sued as an accessory to the breach of trust, the morally innocent quality of the trustee's breach of trust should clearly not afford him a defence.

VII. Conclusion

At the very start of his judgment in *AIB,* Lord Toulson said that "140 years after the *Judicature Act* 1873, the stitching together of equity and the common law continues to cause problems at the seams".[248] Even though the *Judicature Act 1873* was only concerned with the fusion of the administration of common law and equity, rather than the fusion of the substantive rules of each jurisdiction, there is now a need to analyse further the principles underpinning equitable compensation. As Lord Reed observed:

> a trust imposes different obligations from a contractual or tortious relationship"[249] but "[t]o the extent that the same underlying principles apply, the rules [of compensation] should be consistent. To the extent that the underlying principles are different, the rules should be understandably different.[250]

Clarity surrounding equitable compensation is required, both where equitable claims exist alongside common law claims and where the only

246. As section 61 itself says, "the court may relieve [the trustee] either wholly or partly from personal liability". *Cf. Perrins v Bellamy*, [1899] 1 Ch 797 (CA (Eng)) at 80, per Lindley MR.
247. (1902), 86 LT 401 (Ch (Eng)). *Cf. Re Stuart*, [1897] 2 Ch 583 (Eng) at 590.
248. *AIB*, *supra* note 1 at para 1.
249. *Ibid* at para 137.
250. *Ibid* at para 138.

claim available is for breach of trust. The need to hold trustees up to a high standard suggests that a strict approach to causation and remoteness, for example, will generally be appropriate, especially given the difficulties beneficiaries face in discovering breaches of trust.[251]

251. See *e.g. Santander*, *supra* note 241 at para 112, per Count Etherton.

A Proposal for Flexibility in Private and Public Express Trust Enforcement

Mark Gillen*

The Uniform Trustee Act provides for the enforcement of certain non-charitable purpose trusts. This paper draws on that concept but extends it by recommending an expansion of the range of potential enforcement methods for all types of express trusts whether they are for non-charitable purposes, charitable purposes or for persons. In particular, it recommends that, for all types of express trusts, settlors be allowed to indicate possible enforcers. It also recommends that courts be allowed, unless the settlor indicates otherwise, to grant standing to persons to enforce express trusts whether for persons, non-charitable purposes or charitable purposes. It further recommends allowing for Crown enforcement of all types of express trusts (not just charitable purpose trusts) in statutorily specified situations and it recommends extending the availability of cy-près and administrative scheme orders to statutorily-identified non-charitable purpose trusts. The paper also argues that with Crown enforcement of all types of express trusts in statutorily specified situations, it is no longer necessary to retain several of the distinct features of charitable purpose trusts, such as exclusivity, public benefit and the invalidity of political purpose trusts.

---

* Mark Gillen (B. Com, MBA, LL.B., LL.M.) is a professor of law at the Faculty of Law, University of Victoria, Canada. He is co-author, with Donovan WM Waters and Lionel D Smith, of *Waters' Law of Trusts in Canada* (Toronto: Carswell, 2012); co-editor (with Faye Woodman) and co-author of *The Law of Trusts: A Contextual Approach*, 3rd ed. (Toronto: Emond Montgomery, 2015); author of *Securities Regulation in Canada*, 3rd ed. (Carswell, 2007); co-author of *Corporations and Partnerships in Canada* (Wolters Kluwer Law & Business, 2014); and co-author of *Business Organizations: Principles, Policies and Practice* (Toronto: Emond Montgomery, 2007).

I. Introduction

Equity has long held, subject to limited exceptions, that non-charitable purpose trusts are not legally valid trusts. While a provision in Canadian wait-and-see perpetuities legislation[1] provides that trusts for specific non-charitable purposes are valid trusts, it provides for their enforcement only as a power (*i.e.* the trustee may carry out the terms of the non-charitable purpose trust but is not legally obliged to do so). The *Uniform Trustee Act (2012)*[2] ("*Uniform Trustee Act*") proposed by the Uniform Law Conference of Canada in 2012 contains a provision

1. Wait-and-see perpetuity legislation modifies the common law rules against perpetuities. Many common law jurisdictions have enacted such legislation including the provinces of Alberta (*Perpetuities Act*, RSA 2000, c P-5 [*Alberta Perpetuities Act*]); British Columbia (*Perpetuity Act*, RSBC 1996, c 358 [*BC Perpetuity Act*]); Ontario (*Perpetuities Act*, RSO 1990, c P.9 [*Ontario Perpetuities Act*]); Yukon Territories (*Perpetuities Act*, RSY 2002, c 168 [*Yukon Perpetuities Act*]); Northwest Territories (*Perpetuities Act*, RSNWT 1988, c P-3 [*NWT Perpetuities Act*]); and Nunavut (*Perpetuities Act*, RSNWT 1988, c P-3 (as duplicated for Nunavut by s 29 of the *Nunavut Act*, SC 1993, c 28) as amended by the *Miscellaneous Statutes Amendment Act*, S Nu 2010, c 4, s 25 [*Nunavut Perpetuities Act*]). See *infra* notes 34-36 and the accompanying text.
2. *Uniform Trustee Act (2012)* adopted by the Uniform Law Conference of Canada, online: Uniform Law Conference of Canada <www.ulcc.ca> [*Uniform Trustee Act*].

for the enforcement of certain non-charitable purpose trusts as trusts (*i.e.* as legally enforceable obligations). This paper draws on this *Uniform Trustee Act* concept, but recommends expanding the range of potential enforcement for all types of express trusts whether they are for non-charitable purposes, charitable purposes or for persons. In particular, it recommends that, for all types of express trusts, settlors be allowed to indicate possible enforcers and courts be allowed to grant standing to persons to enforce, subject to the settlor indicating otherwise and subject to controls against vexatious actions or actions that would not be in the best interests of the trust.

Following a suggestion in the *Uniform Trustee Act* which allows for Crown enforcement of certain non-charitable purpose trusts, this paper recommends allowing for Crown enforcement of all types of express trusts (not just charitable purpose trusts) in statutorily specified situations. It is also argued that with Crown enforcement of all types of express trusts in statutorily specified situations, it is no longer necessary to retain several of the distinct features of charitable purpose trusts, such as exclusivity, public benefit and the invalidity of political purpose trusts. If the rule against perpetuities is abrogated, as it now has been in three provinces, there would also be no need for the perpetuity exception for charitable purpose trusts. If, however, the rule against perpetuities is retained, it is recommended that the perpetuity exception for charitable purpose trusts be extended to statutorily-identified non-charitable purpose trusts. In either case, the availability of *cy-près* orders should be extended to statutorily-identified non-charitable purpose trusts. In addition, it is suggested that the availability of court ordered administrative schemes be extended to statutorily-identified non-charitable purpose trusts.

In short, this paper recommends making quite drastic changes to trust law developments dating back hundreds of years. The path to the arguments for such significant changes begins, in Part II, with some background on the distinction between express trusts for persons, non-charitable purpose trusts and charitable purpose trusts. Part III indicates how express trusts for persons, sometimes referred to as "private trusts", can provide benefits to a community broader than persons with a close connection to the settlor and may do so in a way that might be considered

socially beneficial or fall into traditional legal concepts of charitable purposes. It also indicates that non-charitable purpose trusts may have purposes that, while not fitting the legal definition of charitable purpose, may nonetheless provide societal benefits. Part IV focuses on enforcement issues, noting that problems of weak enforcement, while perhaps not as severe as for non-charitable purpose trusts, can nonetheless also occur in express trusts for persons and charitable purpose trusts. Part V notes the *Uniform Trustee Act* provision for enforcement of non-charitable purpose trusts and recommends extending the concept of settlor designated enforcers to express trusts for persons and charitable purpose trusts. Part VI recommends allowing the court to grant standing to persons to enforce all types of express trusts subject to the settlor indicating otherwise and subject to controls against vexatious actions or actions that would not be in the best interests of the trust. Part VII then suggests providing for Crown enforcement of all types of express trusts that have aspects that justify expenditure of public funds on enforcement determined according to statutorily provided circumstances for Crown enforcement. With the recommended common approach to the enforcement of express trusts, Part VIII goes on to suggest the elimination of other distinguishing features of charitable purpose trusts such as exclusivity, public benefit, the political purposes doctrine and perpetuities.

II. Relevant Trust Law Background

Some aspects of trust law that provide helpful background to the discussion that follows are briefly reviewed here. Subpart A, below, sets out a typology of trusts to highlight the focus of the paper on express trusts. It also identifies a common division of express trusts into trusts for persons and trusts for purposes and briefly notes the distinction between charitable and non-charitable purposes. Subpart B discusses the general invalidity of non-charitable purpose trusts with emphasis on the enforcement concern asserted to be the main reason for general invalidity. It then briefly notes judicial and statutory exceptions that have been made to general invalidity. Subpart C notes the validity of charitable purpose trusts and how the purpose trust enforcement concern is addressed for charitable purpose trusts. It also notes the requirement that charitable

purpose trusts be for exclusively charitable purposes and the enforcement concern that has been given as the reason for that requirement.

A. A Typology of Trusts: Express Trusts, Trusts by Operation of Law, and Statutory Trusts

A common typology of trusts distinguishes between express trusts and trusts by operation of law.[3] Express trusts are trusts that a person (or persons) intends to create.[4] The intention may have been clearly expressed in words the person used or, where the words are less clear, may be implied from words used, the conduct of the person or the particular circumstances in which the person used the words or engaged in the conduct.[5] Trusts by operation of law include constructive trusts and, in

3. See *e.g.* John Mowbray et al, *Lewin on Trusts*, 18d (London: Sweet & Maxwell, 2008) at 23; David Hayton & Paul Matthews, *Underhill and Hayton Law Relating to Trusts and Trustees*, 18d (London: LexisNexis, 2010) at 79-86; Geraint Thomas & Alastair Hudson, *The Law of Trusts*, 2d (Oxford: Oxford University Press, 2010) at 19-21; Jill E Martin & Harold G Hanbury, *Modern Equity*, 19d (London: Thomson, Sweet & Maxwell, 2012) at 71-74; John McGhee, ed, *Snell's Equity*, 32d (London: Thomson, 2010) at 630; Eileen E Gillese, *The Law of Trusts*, 3d (Toronto: Irwin Law, 2014) at 39, 105; and Albert H Oosterhoff, Robert Chambers & Mitchell McInnes, *Oosterhoff on Trusts: Text, Commentary and Materials*, 8d (Toronto: Carswell, 2014) at 24-28.
4. See *e.g.* Mowbray, *supra* note 3 at 23; Hayton & Matthews, *supra* note 3 at 79-80; Thomas & Hudson, *supra* note 3 at 19; Martin & Hanbury, *supra* note 3 at 71; McGhee, *supra* note 3 at 630; Gillese, *supra* note 3 at 39; Oosterhoff, Chambers & McInnes, *supra* note 3 at 24; and Philip H Pettit, *Equity and the Law of Trusts*, 11d (Oxford: Oxford University Press, 2009) at 67.
5. See *e.g.* Mowbray, *supra* note 3 at 23; Hayton & Matthews, *supra* note 3 at 80; Pettit, *supra* note 4 at 67; Thomas & Hudson, *supra* note 3 at 19-20; and McGhee, *supra* note 3 at 630.

the typology of some, may include resulting trusts.[6] Statutes often create trusts that can have unique aspects deriving explicitly or implicitly from the particular statute.[7] The focus in this paper is on express trusts.

Express trusts have been divided into trusts for persons and trusts for purposes. Trusts for persons provide benefits for persons who are called "beneficiaries". For example, a parent with two minor children, Nancy and Dave, and not expecting to have any more children, might provide in a will that the residue of her or his estate be held in trust and invested, with the investment income to be used to provide for Nancy and Dave

6. Donovan WM Waters, Mark R Gillen & Lionel D Smith, *Waters' Law of Trusts in Canada*, 4d (Toronto: Carswell, 2012) notes that "[a]n express trust arises out of the intention of the settlor; a constructive trust comes into existence, regardless of any party's intent, when the law imposes upon a party an obligation to specific property for the benefit of another" at 478. It also notes at 394, citing *Pecore v Pecore*, 2007 SCC 17, that "[b]roadly speaking, a resulting trust arises whenever legal or equitable title to property is in one party's name, but that party is under an obligation to return it to the original title owner, or to the person who paid the purchase money for it" at para 20. Oosterhoff, Chambers & McInnes, *supra* note 3, says that "[a] constructive trust … defies simple description. Whereas an express trust is defined by its origins in the settlor's intention and a resulting trust is defined by its operation in reversing a transfer, a constructive trust is simply a trust that equity has 'constructed' for some good reason" at 709. In Waters, Gillen & Smith at 394-97, 478 different uses of the term "implied trust," "resulting trust" and "constructive trust" are discussed. While some treat the resulting trust as arising by operation of law, others treat it as an express trust arising out of implied intention (see *e.g.* Peter D Maddaugh & John McCamus, *The Law of Restitution* (Toronto: Canada Law Book, 2014) (loose-leaf 5:200), at 5-6, nn 22a, 23, and the accompanying text; and A J Oakley, *Constructive Trusts*, 2d (London: Sweet & Maxwell, 1987) at 14-18).

7. There are many statutory trusts. Examples include trusts of the property of bankrupts under the *Bankruptcy and Insolvency Act*, RSC 1985, c B-3; builders' lien trusts such as under the *Builders Lien Act*, SBC 1997, c 45, s 10; or trusts for purposes such as habitat conservation an example of which is the Habitat Conservation Trust Fund under the *Wildlife Act*, RSBC 1996, c 488, as amended by SBC 2007, c 24, ss 51-52.

during their minority.[8] Nancy and Dave would both be beneficiaries specifically named in the trust. A trust for persons can also describe a class of persons. A grandparent might, for example, during her or his life, transfer money to a person to hold in trust and invest the money, accumulating the income and making distributions to provide for the post-secondary education of her or his grandchildren. The described class of persons would be the "grandchildren" of the grandparent and could include as yet unborn grandchildren.[9] The object of a trust for purposes differs in that it is not directed to specific persons or to a defined class of persons, but to the pursuit of one or more specific purposes. A person might, for example, provide in her or his will that some specified amount of money out of the person's estate be set aside for the erection of a monument at the person's grave site.

Trusts for purposes are typically divided into two types: charitable purpose trusts and non-charitable purpose trusts. A legally valid charitable purpose trust requires that the purpose(s) of the trust: (i) be charitable purposes; (ii) be exclusively charitable; and (iii) provide a "public benefit".[10] The meaning of "charitable purpose" is based on the

8. The trust could go on to provide that when Nancy and Dave reach the age of majority any remaining funds be, among numerous other possibilities, distributed equally between Nancy and Dave, distributed to some named person other than Nancy or Dave or returned to the settlor. If nothing is said about the distribution of any remaining funds, they would go to the settlor (or the settlor's estate) on a resulting trust.

9. If the applicable jurisdiction of the trust continued to have the common law rule against perpetuities (or a modified version thereof), care would need to be taken to ensure that the potential gift to unborn grandchildren does not vest outside the perpetuity period. The description of a class of beneficiaries needs to meet the test of certainty of beneficiaries. In *McPhail v Doulton* (1970), [1971] AC 424 (HL), it was held that the test of certainty of beneficiaries for a discretionary trust is "that the trust is valid if it can be said with certainty that any given individual is or is not a member of the class" at 16 [*McPhail*].

10. The "charitable" requirement is discussed here in this Part II.A. The exclusivity requirement is discussed further in Part II.C.2 below and the public benefit requirement is discussed further in Part III.C below.

preamble to the *Statute of Charitable Uses of 1601*.[11] The preamble listed examples of charitable purposes to which funds had been provided at that time. In *Morice v The Bishop of Durham*[12] it was held that charitable purpose had a restricted meaning. For it to be a charitable purpose it had to be mentioned in the preamble to the *Statute of Charitable Uses* or

11. 43 Eliz I, c 4. That was a statute that provided a mechanism for the enforcement of purpose trusts through publicly appointed "commissioners". The preamble to the statute said:
"WHEREAS Landes Tenements Rentes Annuities Profittes Hereditamentes, Goodes Chattels Money and Stockes of Money, have bene heretofore given limitted appointed and assigned, as well by the Queenes moste excellent Majestie and her moste noble Progenitors, as by sondrie other well disposed persons, some for Reliefe of aged impotent and poore people, some for Maintenance of sicke and maymed Souldiers and Marriners, Schooles of Learninge, Free Schooles and Schollers in Universities, some for Repaire of Bridges Portes Havens Causwaies Churches Seabankes and Highewaies, some for Educacion and prefermente of Orphans, some for or towardes Reliefe Stocke or Maintenance for Howses for Correccion, some for Mariages of poore Maides, some for Supportacion Ayde and Helpe of younge Tradesmen, Handiecraftesmen and persons decayed, and others for reliefe or redemption of Prisoners or Captives, and for aide or ease of any poore Inhabitants concerninge paymente of Fifteenes, settinge out of Souldiers and other Taxes, Whiche Landes Tenements Rents Annuities Profitts Hereditaments Goodes Chattells Money and Stockes of Money nevertheless have not byn imployed accordinge to the charitable intente of the givers and founders thereof, by reason of Fraudes breaches of Truste and Negligence in those that shoulde pay delyver and imploy the same … ".
12. (1804), 32 ER 656 (Ch) [*Morice* Ch]; (1805), 32 ER 947 (Ch) [*Morice* Ch D].

be analogous to a purpose mentioned therein.[13] In the 1891 House of Lords decision in *Commissioners for Special Purposes of the Income Tax v Pemsel*[14] ("*Pemsel*") Lord McNaughten said the charitable purposes in the preamble fell into four categories:[15]

i. the relief of poverty;
ii. the advancement of education;
iii. the advancement of religion; and
iv. other purposes beneficial to the community.

The categorization in *Pemsel* has been accepted by the Supreme Court of

13. In the words of Master of the Rolls Grant in *Morice* Ch, *supra* note 12 (what is a charitable purpose, "is derived chiefly from the Statute of Elizabeth (stat. 43 Eliz. c. 4). Those purposes are considered charitable, which that Statute enumerates, or which by analogies are deemed within its spirit and intendment" at 659). In upholding the decree of Grant MR, Lord Eldon in *Morice* Ch D, *supra* note 12 noted that the duty of trustees of a disposition to charity is "to apply the money to charity, in the sense, which the determinations have affixed to that word in this Court: viz. either such charitable purposes as are expressed in the Statute (stat. 43 *Eliz*. c. 4), or to purposes having analogy to those. I believe the expression 'charitable purposes,' as used in this Court, has been applied to many Acts described in that Statute and analogous to those, not because they can with propriety be called charitable, but as that denomination is by the Statute given to all the purposes described" at 954.
14. [1891] AC 531 (HL) [*Pemsel*].
15. at 538. This categorization, or something close to it, appears to have been contemplated much earlier given the following comment by Mr. Romilly for the petitioners in *Morice* Ch D, *supra* note 12 ("There are four objects, within one of which all charity, to be administered in the Court, must fall: 1st, relief of the indigent; in various ways: money: provisions: education: medical assistance; &c.: 2dly, the advancement of learning: 3dly, the advancement of religion: 4thly, which is the most difficult, the advancement of objects of general public utility" at 951).

Canada.[16] The fourth category appears on its face to say that all one needs to show to establish a charitable purpose is that the purpose is beneficial to the community. This, however, is not the approach courts, including courts in Canada, have taken. The majority decision of the Supreme Court of Canada in *Vancouver Society of Immigrant and Visible Minority Women v MNR*[17] noted that it is not enough to find that a purpose is beneficial to the community to make it charitable. The benefit must also be one that the law regards as charitable.[18] The majority said the way to determine whether a purpose is charitable is to look to the preamble to the *Statute of Charitable Uses*, to analogies to the preamble and then to analogies upon analogies found in previous cases.[19]

B. Invalidity of Non-Charitable Purpose Trusts

Trusts for non-charitable purposes have traditionally been held not to be legally valid trusts.[20] The idea behind a legally valid trust, whether a trust for persons or a trust for purposes, is that the court will enforce the obligation undertaken by the trustee. The main reason given for the non-validity of non-charitable purpose trusts is that, without named beneficiaries or a defined class of beneficiaries, there is no one to enforce

16. See *e.g. AYSA Amateur Youth Soccer Assn v Canada (Revenue Agency)*, 2007 SCC 42 at 233 [*AYSA*]; *Vancouver Society of Immigrant and Visible Minority Women v MNR*, [1999] 1 SCR 10 at paras 40-41, 102-103 [*Vancouver Society*]; *Guaranty Trust Company of Canada v Minister of National Revenue*, [1967] SCR 133 at 141 [*Guaranty Trust*]; and *The King v Assessors of Sunny Brae (Town)*, [1952] 2 SCR 76.
17. See *Vancouver Society*, *supra* note 16.
18. See the majority decision in *Vancouver Society*, *supra* note 16 at paras 148, 176. The same point is made in the minority judgment at paras 43-49.
19. See the majority decision of *Vancouver Society*, *supra* note 16 at paras 148, 176-79, 200-203. The minority, after discussing several cases and the approach to the fourth head of charitable purposes, also concluded that ("courts should consider whether the purpose under consideration is analogous to one of the purposes enumerated in the preamble of the *Statute of Elizabeth*, or build analogy upon analogy" at 49-51). This approach to charitable purposes under the fourth head was confirmed more recently in *AYSA*, *supra* note 16 at paras 27-28, 31.
20. See the cases cited *infra* note 31.

the trust by seeking a court order to hold the trustee to obligations he or she has undertaken.[21] The phrase often noted in the context of the reason for non-enforcement of purpose trusts is that of Master of the Rolls Grant who, in the 1804 decision of *Morice v The Bishop of Durham*,[22] said that there needs to be someone "in whose favour the court can decree performance".[23]

Over the course of the 19th century some limited exceptions were made to the general rule of non-validity of non-charitable purpose trusts.[24] These were for:

i. the maintenance of a gravesite;[25]
ii. the erection of a monument at a gravesite;[26] and
iii. the provision of food and shelter for specified animals.[27]

Another possible exception was where there was a gift over to some

---

21. See *e.g.* James R Phillips, "Purpose Trusts" in Mark R Gillen & Faye Woodman, eds, *The Law of Trusts: A Contextual Approach* 3d (Toronto: Emond Montgomery, 2015) 165 at 166. See also Donovan WM Waters, "Non-Charitable Purpose Trusts in Common Law Canada" (2008) 28 Estates Trusts & Pensions Journal 16 at 18-19.
22. *Morice* Ch, *supra* note 12 (The purported trust was for "such objects of benevolence and liberality as the trustee [the Bishop of Durham] in his own discretion shall most approve" at 656). Master of the Rolls Grant's decree that the trust was invalid was upheld on appeal by Lord Eldon (see *Morice* Ch D, *supra* note 12).
23. *Morice* Ch, *supra* note 12 at 658. Master of the Rolls Grant's decision was upheld by Lord Eldon on appeal – see *Morice* Ch D, *supra* note 12.
24. While these exceptions allow the trust to be valid so the funds directed to such purposes do not revert to the estate of the deceased, the trust is, practically speaking, only enforced as a power since the fundamental problem remains that there is no one to enforce the trust.
25. See *Trimmer v Danby* (1856), 25 LJ Ch 424 (Eng); *Pirbright v Salwey*, [1896] WN 86 (Ch (Eng)); *Re Hooper*, [1932] 1 Ch 38 (Eng). This was accepted in Canada in *Crocker v Senior* (1971), 2 Nfld & PEIR 179 (SCTD).
26. *Ibid*. A different position was, however, taken in the Canadian case of *Re Jefferson Estate*, [1929] 3 WWR 690 (Man KB).
27. *Pettingall v Pettingall* (1842), 11 LJ Ch 176 (Eng); *Mitford v Reynolds* (1848), 60 ER 812 (Ch).

person after a period, complying with the rule against perpetuities, during which the trust for the purpose could operate.[28] The exceptions led to the argument in *Re Astor's Settlement Trusts*[29] that these exceptions meant there was no general rule against the non-validity of non-charitable purpose trusts.[30] It was, however, held that there was a general rule of non-validity and that the list of exceptions should be considered closed.[31]

The possible exception where there was a gift over to some person after a period for the operation of the trust provided a means of enforcement of the purpose trust. The person to whom there was a gift over at the end of the period would have an incentive to monitor the trustee and, if warranted, pursue a breach of trust action if the trustee expended the

---

28. See *Re Thompson*, [1934] Ch 342 (Eng) [*Re Thompson*] commented on in *Re Astor's Settlement Trusts*, [1952] Ch 534 (Eng) [*Re Astor's*]. *Re Thompson* involved a trust for the purpose of furthering fox-hunting leaving the residue after a specified period to Trinity College of Cambridge University. The purpose trust was held to be valid since the purpose was found to be reasonably certain and since Trinity College of Cambridge University could enforce the trust if the funds were misapplied. Pettit, *supra* note 4 at 60-61, suggests that there may also be an exception for trusts for the saying of masses if they are not found to be for charitable purposes (saying of masses in public have been found to be for charitable purposes) and are restricted to the perpetuity period citing *Bourne v Keane*, [1919] AC 815 (HL).
29. *Re Astor's*, *ibid*.
30. See the arguments of Gray QC and Wilfred M Hunt in *Re Astor's*, *ibid* at 539 and Justice Roxburgh's summary of their argument at 541.
31. *Re Astor's, ibid* at 547; *Leahy v Attorney General of New South Wales*, [1959] 2 All ER 300 (PC (Austl)); and *Re Endacott*, [1960] Ch 232 (CA) (LJ Harmon at 250-51, and Lord Evershed, at 247) [*Re Endacott*]. In Canada see *e.g. Wood v R* (1977), 1 ETR 285 (Alta SC (TD)) at paras 18-19 [*Re Russell*]; *Rowland v Vancouver College Ltd*, (2000), 78 BCLR (3d) 87 (SC) at paras 68-73; (aff'd on appeal (2001), 94 BCLR (3d) 249 (CA) but the non-charitable purpose trust question was not discussed); *Dionisio v Mancinelli* (2004), 12 ETR (3d) 296 (Ont Sup Ct J); and *Keewatin Tribal Council Inc v City of Thompson* (1989), 61 Man R (2d) 241 (QB) at para 69 [*Keewatin*]. *Re Astor's*, *supra* note 28, and *Re Endacott* were considered in Canada in *Re Lerner and Society for Crippled Children & Adults of Manitoba*, [1979] 3 ACWS 442 (Man QB).

trust funds in a way that was not consistent with the purpose of the trust. This, it has been pointed out, provides only "negative enforcement" since the person to whom there was a gift over at the end of the period would not have an incentive to ensure the trustee actually used trust funds to pursue the intended purpose.[32] Indeed, the person to whom there was a gift over at the end of the period would be better off if the trustee never expended any of the trust funds on the intended purpose. This would effectively make the trust operate as a power since it would be unlikely that any action would be taken against the trustee for not expending funds for the purpose, although the trustee could expend the funds for the purpose if the trustee so chose.[33]

Wait-and-see perpetuities legislation often contains a provision for the validity of non-charitable purpose trusts along similar lines to the possible exception of a purpose trust for a period with a gift over at the end of the period.[34] The typical provision provides that where there is a trust for a specific non-charitable purpose that creates no enforceable equitable interest in a specific person, the "trust" is valid but is to be construed as a power and can operate as such for a maximum period

32. See *e.g.* Oosterhoff, Chambers & McInnes, *supra* note 3 at 545.
33. On the effect of enforcing a purpose trust where there is a gift over of the residue as being similar to a power see *e.g.* James E Penner, *The Law of Trusts*, 7d (Oxford: Oxford University Press, 2010) at 248.
34. See *e.g. Alberta Perpetuities Act*, *supra* note 1, s 20; *BC Perpetuity Act*, *supra* note 1, s 24; *NWT Perpetuities Act*, *supra* note 1, s 17; *Nunavut Perpetuities Act*, *supra* note 1, s 25; *Ontario Perpetuities Act*, *supra* note 1, s 16; *Yukon Perpetuities Act*, *supra* note 1, s 20. See also American Law Institute, *Restatement of the Law of Trusts*, 3d (St. Paul: American Law Institute Publishers, 1992) at 47; and prior to that American Law Institute, *Restatement of the Law of Trusts*, 2d (Washington, DC: American Law Institute Publishers, 1957) at 124. See also the United States of America's *Uniform Trust Code* § 409 (2010); and see *e.g. California Probate Code* § 15211.

of up to twenty-one years.[35] At the end of the period during which the purpose is to operate, subject to a maximum of twenty-one years, the capital and unexpended income is to go to the person who would have been entitled to the property at the end of twenty-one years.[36]

The general rule of non-validity of non-charitable purpose trusts does not apply if the trust, although expressed as being for a purpose, can be construed as being a trust for persons, since there are then persons in whose favour the court can decree performance — *i.e.* persons who could enforce the trust. A common example is a trust "for the education of my children" that, while expressed to be for the purpose of educating

35. See *e.g. Alberta Perpetuities Act*, *supra* note 1, s 20(1); *BC Perpetuity Act*, *supra* note 1, ss 24(1)-(3); *Ontario Perpetuities Act*, *supra* note 1, s 16(1). Two of the questions that arise in the interpretation of this legislation are: (i) when does the provision apply; and (ii) what is the meaning of the phrase "a specific non-charitable purpose trust". As to the first question, one might argue that since this purpose trust provision appeared in legislation modifying the common law rule against perpetuities, it would only apply if the particular trust violated the common law rule (*i.e.* if the purpose trust was for a perpetual duration). In *Re Russell*, *supra* note 31, the court held a non-charitable purpose trust to be valid applying the provision in the Alberta wait-and-see perpetuities legislation (*The Perpetuities Act*, SA 1972, c 121, s 20, now *Perpetuities Act*, RSA 2000, c P-5, s 20) even though the purported trust did not violate the common law perpetuity rule. As to the second question, the court in *Re Russell*, *supra* note 31, applied the test in *McPhail, supra* note 9 at 28 (modifying it to purpose trusts so that there is "a specific non-charitable purpose" if one can say whether any given use of the trust funds would qualify as a proper use).

36. See *e.g. Alberta Perpetuities Act*, *supra* note 1, s 20(2); *BC Perpetuity Act*, *supra* note 1, s 24(4); *Ontario Perpetuities Act*, *supra* note 1, s 16(2). If, for example, there was a gift in a will to a person who was to use the property for a specific non-charitable purpose, that gift could take effect. It would not revert to the estate on the basis that the purported trust was an invalid non-charitable purpose trust.

the children, can be read as "to my children for their education".[37] In this example, whichever of the two wordings are used, the beneficiaries are the children and the problem of there being no one "in whose favour the court can decree performance" does not arise.[38] *Re Denley's Trusts*[39] ("*Re Denley's*") took this approach to finding a trust expressed ostensibly as for a purpose to be a valid trust. Land was given to trustees to create a recreational or sports ground for the benefit of the employees of a particular company.[40] While the trust was expressed as being for the purpose of creating a recreation or sports ground, it was clear that the beneficiaries were the employees. The court could decree performance in favour of the employees as beneficiaries — *i.e.* there were beneficiaries who could enforce the trust. Justice Goff held the trust to be a valid trust saying, "[w]here, then, the trust, though expressed as a purpose,

37. See *e.g.* Phillips, *supra* note 21 at 171; and Gillese, *supra* note 3 at 59. A case involving a trust remarkably similar to this example is *Sacks v Gridiger* (1990), 22 NSWLR 502 (SC (Austl)) where the settlor created a trust "to pay the school tuition fees for the children of Dr Marcus L Sachs … while both or either of them remain at school".

38. If necessary, enforcement on behalf of the children can be through a Public Guardian and Trustee, Children's Lawyer or similar official. See *e.g. The Public Trustee Act*, SA 2004, c P-44.1, s 5(c) [*Alberta Public Trustee Act*]; *The Public Guardian and Trustee Act*, RSBC 1996, c 383, s 7 [*BC Public Guardian and Trustee Act*]; *The Public Guardian and Trustee Act*, CCSM, c P205, s 5(2)(b) [*Manitoba Public Guardian and Trustee Act*]; *The Children's Law Reform Act*, RSO 1990, c C 12, s 47; and *The Public Trustee Act*, RSNS 1989, c 379, s 4(2) [*Nova Scotia Public Trustee Act*].

39. [1968] 3 All ER 65 (Ch) [*Re Denley's*].

40. *Ibid* at 67-69. The facts were a bit more complicated. The trustees wanted to sell part of the lands to raise funds to improve the recreational facility and asked the court to declare that they had a power to sell part of the land. The deed, however, said that if less than 75% of the employees were subscribing for the use of the facilities then the property was to go to the General Hospital Cheltenham. The General Hospital Cheltenham opposed the application since less than 75% of the employees were subscribing to use the facility and sought an order that the land be conveyed to it. The trustees responded by saying the trust was a purpose trust and was therefore not a valid trust so the land should be held by them on a resulting trust for the company.

is directly or indirectly for the benefit of an individual or individuals, it seems to me that it is in general outside the mischief of the beneficiary principle".[41]

Re Denley's might be interpreted to simply mean that a trust expressed as being for a purpose is valid as long as it can be interpreted as really being a trust for persons. This is the interpretation it has been given in England and Australia.[42] There is, however, Canadian authority relying on *Re Denley's* for the broader proposition that a purpose trust is valid as long as there is some person who can be given standing to enforce the trust. In 1989 in *Keewatin Tribal Council Inc v City of Thompson*,[43] Justice Jewers made reference to *Re Denley's* commenting, "[t]he real question is one of enforceability and nothing else … there should be no problem with a non-charitable purpose trust where there are any number of persons with standing to enforce it". [44]

More recently in *Peace Hills Trust Co v Canada Deposit Insurance Corp*,[45] Justice Thomas said, "having particular regard to the approach taken in *Keewatin* … a non-charitable purpose trust may be created in

41. *Ibid* at 69.
42. See *e.g.* in England *Re Grant's Will Trusts*, [1980] 1 WLR 360 (Ch) at 370-71; in Australia *Tidex v Trustees Executors and Agency Co Ltd*, [1971] 2 NSWLR 453 (WASC (Austl)) at 465-66; and *Strathalbyn Show Jumping Club Inc v Mayes* (2001), 79 SASR 54 (SASC (Austl)) at 64-66.
43. *Keewatin*, *supra* note 31.
44. *Ibid* at para 72. No one was actually granted standing to sue as a beneficiary in the case. The issue in the case was whether an exemption from municipal taxation could be claimed under the particular municipal taxing statute that provided an exemption for "lands held in trust for any tribe or body of Indians" at para 33. The applicant was Keewatin Tribal Council Inc which held title to the lands but purported to hold the lands in trust for a "tribe or body of Indians". It was appealing from a decision by the "Board of Revision" which had upheld the imposition of the municipal tax on Keewatin Tribal Council Inc. The respondent, the City of Thompson, was the taxing municipality.
45. 2007 ABQB 364 [*Peace Hills*].

Canada and would be recognized by the courts of this country".[46]

C. Validity and Enforcement of Charitable Purpose Trusts

1. Crown Enforcement of Charitable Purpose Trusts

An exception to the invalidity of purpose trusts has long been made for charitable purpose trusts.[47] The Crown, as *parens patriae*, enforces charitable purpose trusts under its prerogative power. Responsibility for the enforcement of charitable purpose trusts in Canada is normally exercised by the relevant provincial Attorney General.[48] The Crown can take legal action in response to fraudulent or negligent acts of trustees of charitable purpose trusts or the Crown may be called upon in the context

46. *Ibid* at para 29. *Keewatin*, *supra* note 31, and *Peace Hills*, *ibid*, raise issues about when the wait-and-see perpetuities legislation provisions on purpose trusts apply. Do these provisions only now apply in Canada where *Keewatin* and *Peace Hills* have no application? Does that leave much room for the operation of the wait-and-see legislation provisions? See the discussion in Waters, *supra* note 21.
47. This goes back at least as far as *Morice* Ch, *supra* note 12. See per Grant MR in *Morice* Ch, *supra* note 12 at 658 and per Lord Eldon on appeal at *Morice* Ch D, *supra* note 12 at 954.
48. This role of the Attorney General or Public Trustee has been recognized and discussed in several Canadian cases. See *e.g. L'Évêque Catholique Roman de Bathurst v New Brunswick (Attorney General)* 2010 NBBR 372 (QB); *Pathak v Hindu Sabha* (2004), 8 ETR (3d) 151 (Ont Sup Ct J); *Re Baker* (1984), 47 OR (2d) 415 (H Ct J) at 420; *Re Stillman Estate* (2003), 5 ETR (3d) 260 (Ont Sup Ct J). See the discussion in Waters, Gillen & Smith, *supra* note 6 at 123-25. On the historical development of the *parens patriae* role of the Attorney General in England and Wales and in Canada see Kathryn Chan, "The Role of the Attorney General in Charity Proceedings in Canada and in England and Wales" (2010) 89:2 Canadian Bar Review 373.

of litigation involving the trust.[49]

Some supervision of charitable purpose trusts in Canada is done by the Charities Directorate of the Canada Revenue Agency (the "CRA") in the context of the federal *Income Tax Act*[50] exemption of a "registered

49. See Waters, Gillen & Smith, *supra* note 6 at 835. The Crown might, for example, be called upon in the context of litigation by residuary legatees of the estate of a settlor arguing that the trust is not valid and that the funds allocated to the trust remain in, or result to, the estate of the settlor. The Crown might also be called upon in the context of an administrative scheme or *cy-près* application.

50. RSC 1985, c 1 (5th Supp) [*Income Tax Act*].

charity"[51] from income tax[52] and the tax credit for donations to registered

51. *Income Tax Act*, *supra* note 50, a "registered charity" is defined in the Act to include a "charitable organization, private foundation or public foundation" with the terms "charitable organization", "private foundation" and "public foundation" defined in ss 149.1(1)) and 248(1). Section 149.1(1) defines a "private foundation" as a "charitable foundation that is not a public foundation" and defines a "public foundation" as a "charitable foundation" that meets certain conditions set out in the definition. A "charitable foundation" is defined in s 149.1(1) as "a corporation or trust that is constituted and operated exclusively for charitable purposes, … , and that is not a charitable organization". A "charitable organization" is defined in s 149.1(1) as "an organization, whether or not incorporated, (a) all the resources of which are devoted to charitable activities carried on by the organization itself". The *Income Tax Act* does not define the word "charitable" for the "charitable purposes" or "charitable activities" parts of the definitions of "charitable foundation" or "charitable organization". In interpreting the meaning of "charitable" for these purposes courts have relied on the common law meaning given to "charitable purposes" in the context of trust law. See *Vancouver Society*, *supra* note 16 at para 143.

52. *Income Tax Act*, *supra* note 50, s 149(1)(f) which says that "no tax is payable under this Part on the taxable income of a person for a period when that person was … (f) a registered charity". Trusts are not recognized as separate legal persons but the *Income Tax Act* treats them as taxpayers (see s 104(2)). As taxpayers, trusts are required to pay tax on income earned by the trust (*e.g.* rent on leasehold property, dividends on shares or interest on debentures). As with other taxpayers, the trust can, in calculating trust income, deduct expenses reasonably incurred in earning the income. A particular feature unique to trust taxation under the *Income Tax Act* is that the trust may deduct amounts that became payable to beneficiaries during the year (s 104(6)(b)) and amounts so deducted with respect to any given beneficiary are to be included in that beneficiary's income for the year (s 104(13)). The trust, therefore, operates as a conduit through which income earned on trust investments can flow through to the beneficiaries to be taxed in the hands of the beneficiaries. The trust, therefore, is only taxed on the income that is retained (*i.e.* accumulated) in the trust). See *e.g.* Faye Woodman, "Introduction to the Taxation of Trusts and Beneficiaries" in Mark R Gillen & Faye Woodman, eds, *The Law of Trusts: A Contextual Approach*, 3d (Toronto: Emond Montgomery, 2015) at 25-28; and Peter W Hogg, Joanne E Magee & Jinyan Li, *Principles of Canadian Income Tax Law*, 8d (Toronto: Carswell, 2013) at 537.

charities.[53] Registration as a "registered charity" is required to take advantage of the exemption from income tax,[54] or to provide receipts required for the tax credit.[55] The Charities Directorate of the CRA assesses applications for registered charity status[56] and can remove that status where the registered charity is not devoting its funds to charitable purposes or engaging in charitable activities.[57] Registered charities must file reports to facilitate the assessment of whether they are complying with

53. See *Income Tax Act*, *supra* note 50 (s 118.1(3) for the tax credit for donations by individuals that provides for a deduction from tax payable for "total gifts" for the year, and see s 118.1(1) for the definition of "total gifts" that refers to "total charitable gifts" which is also defined in s 118.1(1). The definition of "total charitable gifts" refers to gifts to a "qualified donee" which is defined in s 149.1(1) to include a "registered charity". See *supra* note 51 for the definition of "registered charity". See also *Income Tax Act*, *supra* note 50, s 110.1(1) for a deduction for charitable donations by corporations).
54. See *supra* note 52.
55. See *supra* note 53. See also *Income Tax Act*, *supra* note 50 that says "a gift shall not be included in the total charitable gifts … of an individual unless the making of the gift is proven by filing with the Minister (a) a receipt for the gift that contains prescribed information" at s 118.1(2). The prescribed information is set out in *Income Tax Regulation,* CRC, c 945, s 3501(1) that refers to "every official receipt issued by a registered organization", and s 3500 defines a "registered organization" to include "a registered charity".
56. To obtain registered charity status as a charitable foundation under the *Income Tax Act*, *supra* note 50 the foundation must show it will be providing its funds to organizations operated for "charitable purposes". See the definition of "registered charity" in *supra* note 51. To obtain registered charity status as a charitable organization under the *Income Tax Act*, the organization must show it will devote its resources exclusively to "charitable activities". See the definition of "charitable organization" in *Income Tax Act*, *supra* note 50, s 149.1(1). The *Income Tax Act* does not define "charitable purposes" or "charitable activities". In interpreting these terms courts have relied on the common law meaning of "charitable purpose" developed in the trust law context (see *Vancouver Society*, *supra* note 16).
57. See *Income Tax Act*, *supra* note 50, ss 149(2)-(4), 168(1).

the requirements for income tax exempt status.[58] The CRA, however, as discussed further below,[59] likely does not have the incentive or mandate to ensure charitable purpose trust funds are being used for the settlor's intended charitable purposes as long as charitable purpose trust funds are being used for charitable purposes. It also likely does not have the incentive or mandate to ensure the charitable purpose trust funds are being used efficiently for the settlor's intended charitable purposes.

Ontario is unique among the Canadian common law jurisdictions in having legislation relating to the enforcement of charitable purpose trusts.[60] In particular, the Ontario *Charities Accounting Act* requires a donee of property to be held for charitable purposes to give notice of the trust to the Public Guardian and Trustee.[61] The Public Guardian and Trustee may request information of the trustee concerning the trust including the condition, disposition or other such particulars as requested of the property held in trust or as to any other matter relating to the administration or management of the trust.[62] The Public Guardian and Trustee may also request of the trustee information concerning dealings with the property "to be passed and examined and audited by a judge of the Superior Court of Justice".[63] The Public Guardian and Trustee may apply to court for an order where, among other matters, the trustee is found to have misapplied or misappropriated trust property, made an improper or unauthorized investment or is not applying the trust property in the manner directed by the trust instrument.[64] The Act also allows any two or more persons who allege a breach of a charitable purpose trust or who seek the direction of the court for the administration

58. *Ibid*, s 149(14) together with Forms T3010, T1235, TF725 and publication T4033 "Completing the Registered Charity Information Return".
59. See Part IV.B, below.
60. *Charities Accounting Act*, RSO 1990, c C 10.
61. *Ibid*, s 1 which deals with an *inter vivos* instrument – presumably the Public Trustee and Guardian will have notice of a will that has received probate.
62. *Ibid*, s 2. This applies to both testamentary and *inter vivos* trusts.
63. *Ibid*, s 3.
64. *Ibid*, s 4.

of a charitable purpose trust to apply to court, with notice to the Public Guardian and Trustee, to hear the application and make such order as the court considers just for the carrying out of the trust.[65]

In England and Wales there is a Charity Commission that performs its functions on behalf of the Crown.[66] Its functions include: (i) determining whether institutions are charities or not; (ii) identifying and investigating apparent misconduct or mismanagement in the administration of charities and taking remedial or protective action in that connection;[67] (iii) giving such advice or guidance it considers appropriate with respect to the administration of charities;[68] and (iv) the maintenance of an accurate and up-to-date register of charities.[69] Subject to certain exemptions,[70] every charity is required to be registered with the registry, including the name of the charity and other information as the Charity Commission thinks fit.[71] The Charity Commission has the power to make inquiries concerning a charity[72] and has the discretion to publish the results of the

65. *Ibid*, s 10. Section 10 also allows the court to make an order directing the Public Guardian and Trustee to conduct an investigation of a charitable purpose trust and make a report to the Attorney General.
66. The Charity Commission is continued under the *Charities Act 2011* (UK), c 25, s 13. The Charity Commission had its origins in mid-19th century legislation (*Charitable Trusts Act, 1860* (UK), 23 & 24 Vict, c 136). The powers of the Charity Commission were increased in the *Charities Act 1960*, 8 & 9 Eliz II, c 58 (replaced by the *Charities Act 1993* (UK), c 10 then by the *Charities Act 2006* (UK), c 50 and subsequently by the *Charities Act 2011* (UK), c 25 [*Charities Act 2011*]). The Charity Commission is not subject to the direction or control of any Minister of the Crown or of another government department (see *Charities Act 2011*, s 13(4)) although it is subject to administrative controls by the Treasury over its expenditures (see *Charities Act 2011*, s 13(5)).
67. *Charities Act 2011*, *supra* note 66, s 15(1).
68. *Ibid*, s 15(2).
69. *Ibid*, s 15(4).
70. There is, for example, an exemption for charities with less £5,000 in gross income. See *ibid*, s 30(2)(d).
71. *Ibid*, ss 29(2), 30.
72. *Ibid*, s 46 (including broad powers in connection with inquiries – see also ss 47-53).

inquiry.[73] Proceedings concerning a charity can be brought by any of the trustees of the charity, "any person interested in the charity" or by any two or more inhabitants of the area of a charity where it is a local charity[74] and where the Charity Commission authorises such proceedings.[75]

2. Exclusivity and the Enforcement of Charitable Purpose Trusts

A charitable purpose trust principle that is related to the question of enforcement of charitable purpose trusts is the requirement that such a trust be exclusively for charitable purposes.[76] A purported trust will not be a valid charitable purpose trust if it is partly for a charitable purpose and partly for some non-charitable purpose or partly for persons. The reason given for this principle is that if the trust had a charitable purpose combined with a non-charitable purpose or was, in part, for a defined class of persons, the trustee would have the discretion to use the funds for the charitable or non-charitable purposes or for persons. That would make it hard to enforce use of the trust property for the charitable purposes since the trustee might simply answer that it was within his discretion to use

73. s 50. Australia also has a statutory scheme for charities regulation and enforcement including charitable trusts. See *e.g.* the *Australian Charities and Not-for-profits Commission Act 2012* (Cth); and see "Australian Charities and Not-for-profits Commission Statement: Regulatory Approach", online: Australian Charities and Not-for-profits Commission <www.acnc.gov.au>.

74. *Charities Act 2011*, *supra* note 66, s 115(1).

75. *Ibid*, s 115(2). Where the Charities Commission refuses to grant authority an application can be made to a judge of the High Court attached to the Chancery Division for leave to take proceedings.

76. On this exclusivity requirement see *e.g. Morice* Ch, *supra* note 12; *Vezey v Jamson* (1822), 57 ER 27 (Ch); *Hunter v Attorney-General*, [1899] AC 309 (HL) at 323-24; *Re MacDuff*, [1896] 2 Ch 451 (CA (Eng)) at 464-66 [*Re MacDuff*]; *Re Eades*, [1920] 2 Ch 353 (Eng); and *Chichester Diocesan Fund & Board of Finance Inc v Simpson*, [1944] 2 All ER 60 (HL). In Canada see *e.g. Brewer v McCauley*, [1954] SCR 645 [*Brewer*]; and *Jones v Executive Officers of the T Eaton Company Limited*, [1973] SCR 635 [*Eaton*].

the funds for the non-charitable purposes or for persons.[77]

This exclusivity principle has the unfortunate capacity to invalidate many charitable purposes that have been inadvertently mixed with non-charitable purposes or with trusts for persons.[78] Courts have used several techniques to avoid such unfortunate outcomes.[79] Several common law

77. *Morice* Ch D, *supra* note 12. This reason was given by Lord Eldon where he says that "the true question is, whether, if upon the one hand he might have devoted the whole to purposes, in this sense charitable, he might not equally according to the intention have devoted the whole to purposes benevolent and liberal, and yet not within the meaning of charitable purposes, as this Court construes those words; and, if according to the intention it was competent to him to do so, I do not apprehend, that under any authority upon such words the Court could have charged him with mal-administration, if he had applied the whole to purposes, which according to the meaning of the testator are benevolent and liberal" at 955. Similarly, Grant MR's earlier decision in *Morice* Ch, *supra* note 12 said, "[t]he question is, not, whether he may not apply it upon purposes strictly charitable, but whether he is bound so to apply it?" at 659. This reasoning was accepted in *Re MacDuff*, *supra* note 76 at 465.

78. The settlor may, for instance, have directed funds to be used for "philanthropic" purposes (see *e.g. Brewer*, *supra* note 76), "benevolent" purposes (see *e.g. Lawrence v Lawrence* (1913), 42 NBR 260 (SC), "worthy" purposes (see *e.g. Planta v Greenshields*, [1931] 1 WWR 401 (BCCA)) or "public" purposes (see *e.g. Cox v Hogan* (1924), 35 BCR 286 (SC) although the majority held the gift to be charitable on the basis of added words that limited the public purposes to specified purposes that were charitable). However commendable such purposes may be, they do not fall within the legal meaning of "charitable purposes".

79. Courts have, for instance, interpreted the intention apparently expressed as an intention to devote the trust property to both charitable and non-charitable purposes as really being an intention to have all the property devoted to charitable purposes (see *e.g. Eaton*, *supra* note 76). They have also read the non-charitable part as charitable based on who the donee is (see *e.g. Blais v Touchet*, [1963] SCR 358 where the donee was a bishop and the donor was a priest). Courts have also sometimes severed the non-charitable portion. See *e.g. Re Coxen*, [1948] 2 All ER 492 (Ch); *Re Porter*, [1925] Ch 746 (Eng); *Re Clarke*, [1923] 2 Ch 407 (Eng); and *Salusbury v Denton* (1857), 69 ER 1219 (Ch). They have also sometimes found the non-charitable purpose to be merely ancillary to the charitable purposes (*Guaranty Trust*, *supra* note 16).

jurisdictions have enacted a provision to address this concern,[80] although the provisions typically contain what might be said to be flaws limiting their application.[81] The *Uniform Trustee Act* addresses this concern by providing, in section 75(1), that, "a trust is not void by reason only that the objects of the trust consist of a charitable purpose and a non-charitable purpose". Section 75(3) gives the court powers to make orders addressing issues of combined valid charitable and non-charitable trusts, separating them where possible and otherwise dealing with situations

80. Early examples include, for example, in Australia *Property Law Act 1958* (Vic), s 131; *Conveyancing Act 1919* (NSW), s 37D; *Trustees Act 1962* (WA), s 102; in New Zealand, *Charitable Trusts Act 1957* as amended by *An Act to Amend Charitable Trusts 1957* (NZ), 1957/18 s 4(1) adding s 61B to the *Charitable Trusts Act 1957*; in Ireland, *Charities Act, 1961*, Act No 17 of 1961, s 49; and in Northern Ireland, *Charities Act* (Northern Ireland), 1964, c 33, s 24. For British Columbia see the *Law and Equity Act*, RSBC 1996, c 253, s 47 [*Law and Equity Act*]. For Alberta see the *Wills Act*, RSA 2000, c W-12, s 32; for New Brunswick see *Wills Act*, RSNB 1973, c W-9, s 30; and for Manitoba see the *Trustee Act*, CCSM, c Tl60, s 91.

81. British Columbia *Law and Equity Act*, *ibid*, says, for example, that "[i]f a person gives, devises or bequeaths property in trust for a charitable purpose that is linked conjunctively or disjunctively in the instrument by which the trust is created with a non-charitable purpose, and the gift, devise or bequest would be void for uncertainty or remoteness, the gift, devise or bequest is not invalid as a result but operates solely for the benefit of the charitable purpose" at s 47. One of the problems with this provision is that the charitable purpose must be linked "linked conjunctively or disjunctively … a non-charitable purpose" and, therefore, does not deal with expressions such as "such benevolent purposes as my trustees choose", "such worthy objects as my trustees shall select" or "the general benefit of children in the Sick Children's Hospital". Another problem is that the provision only applies if the "gift, devise or bequest would be void for uncertainty or remoteness". If the gift, devise or bequest is not uncertain and not void for remoteness, the provision would not apply. It is not clear how the provision would fit with the wait-and-see perpetuity legislation provisions that treat a trust for a "specific non-charitable purpose that creates no enforceable equitable interest in a specific person" as a valid trust but to be construed "as a power to appoint the income or capital".

where they are not separable.[82]

3. Political Purposes and Charitable Purpose Trusts

It is also important to note for the discussion in Part VIII, below, that even if a trust is for an exclusively charitable purpose that provides a public benefit, it will not be a valid charitable purpose trust if it is for "political purposes".[83] "Political purposes" include not just the promotion of a particular candidate in an election, a political party or political ideas, but also any attempt to influence the legislative or executive process, to influence domestic, or foreign, laws or government policy or to improve international relations.[84] The reason given for the rule is that the court has no means of judging whether a proposed change in the law will or

82. *Uniform Trustee Act*, *supra* note 2, s 75(3). Paragraphs (a), (b), (c) provide for the separation of mixed charitable and non-charitable purposes where it is "practicable" to separate them. Paragraph (d) through (g) provide for various situations where the court finds it is not practicable to separate mixed charitable and non-charitable purposes.

83. See *Bowman v Secular Society*, [1917] AC 406 (HL) [*Bowman*]. In Canada see *Re Loney Estate* (1953), 9 WWR (2d) 366 (Man QB); and *Ontario (Public Trustee) v Toronto Humane Society* (1987), 60 OR (2d) 236 (HCt J).

84. *McGovern v Attorney-General*, [1982] Ch 321 (Eng) at 340 followed in Canada in *Human Life International in Canada Inc v Minister of National Revenue*, [1998] 2 FC 202 (CA).

will not be for the public benefit.[85] Many have been critical of this rule.[86]

III. The Potential Public-Private Trust Overlap

This part begins by noting the public-private distinction that has sometimes been used to distinguish charitable purpose trusts from express trusts for persons. Subpart B notes the public benefit requirement for the validity of charitable purpose trusts and discusses how courts have defined "public" in assessing that requirement. Subpart C argues that express trusts for persons and non-charitable purpose trusts can potentially have "public" qualities.

A. Public Trusts and Private Trusts

A public-private distinction is sometimes made between charitable purpose trusts and express trusts for persons. Charitable purpose trusts, are sometimes referred to as "public trusts". Express trusts for persons and non-charitable purpose trusts are sometimes referred to as "private trusts".[87] Charitable purpose trusts are, in part, "public" in the sense that the purposes are normally directed to a broader community and not just

85. See *Bowman*, *supra* note 83 at 417. In other words, in assessing public benefit the court would be put in the position of saying that there would be a public benefit to changing the law (*i.e.* implicitly saying the law was a bad law) when it might subsequently be bound to enforce that law.

86. See *e.g.* Paul Michell, "The Political Purposes Doctrine in Canadian Charities Law" (1995) 12:4 The Philanthropist 3 at 4-6; Henry G Intven, "Political Activity and Charitable Organisations" (1983) 3:3 The Philanthropist 35; LA Sheridan, "Charity Versus Politics" (1973) 2 Anglo-American Law Review 47; LA Sheridan, "The Charpol Family Quiz: A Game of Skill and Luck Played on the Boundaries of Charity and Politics" (1977) 2:1 The Philanthropist 14; and Abraham Drassinower, "The Doctrine of Political Purposes in the Law of Charities: A Conceptual Analysis" in Jim Phillips, Bruce Chapman & David Stevens, eds, *Between State and Market: Essays on Canadian Charities Law and Policy* (Montreal: McGill-Queen's University Press, 2001) 288.

87. Mowbray, *supra* note 3 at 22; Pettit, *supra* note 4 at 74; Thomas & Hudson, *supra* note 3 at 24; McGhee, *supra* note 3 at 683-84; Oosterhoff, Chambers & McInnes, *supra* note 3 at 24; Waters, Gillen & Smith, *supra* note 6 at 29.

the benefit of particular private interests.[88] Charitable purpose trusts are also "public" in the sense that the Crown (*i.e.* "the State") is given standing to enforce such trusts. Enforcement by the Crown involves expenditure, although perhaps modest, of government funds provided at taxpayer expense. Charitable purpose trusts have another "public" quality in that they can benefit, as described above, from tax subsidies under the federal *Income Tax Act*.[89]

B. The "Public" and "Private" Aspects of Charitable Purpose Trusts

It has been noted that a charitable purpose trust is a hybrid of private and public relationships.[90] This might be articulated in terms of a distinction between public law and private law that is said to have its roots in the jurisprudential model established by the Romans in which law is

88. See the reference to the public benefit requirement, *supra* note 10 and the accompanying text, and see *infra* Part III.C on the public aspects of charitable purpose trusts. Charitable purpose trusts are "normally," not always, directed to a broader community since there are recognized exceptions to the public benefit requirement with respect to "poor relations" and "poor employees" (see Part III.C, below).

89. See *supra* notes 50-58 and the accompanying text.

90. See Mayo Moran, "Rethinking Public Benefit: The Definition of Charity in the Era of the Charter" in Jim Phillips, Bruce Chapman & David Stevens, eds, *Between State and Market: Essays on Charities Law and Policy in Canada* (McGill-Queen's University Press 2001) 251 "the charitable trust is neither a purely private nor a purely public institution; instead it is a hybrid of public and private relationships and must be assessed as such" at 257. On the public and private overlap of charitable purpose trusts see also *e.g.* Mark Freedland, "Charity Law and the Public/Private Distinction" in Charles Mitchell & Susan R Moody, eds, *Foundations of Charity* (Oxford: Hart Publishing, 2000); Adam Parachin, "Regulating Philanthropy: Is Charity Public or Private" (CBA National Charity Law Symposium, Toronto, 4 May 2012) 111; and Evelyn Brody & John Tyler, "Respecting Foundation and Charity Autonomy: How Public is Private Philanthropy?" (2010) 85 Chicago-Kent Law Review 571 which focuses largely on charities organized as not-for-profit corporations, but much of what it says about the public and private aspects of charities could be extended to charities organized as trusts (*i.e.* charitable purpose trusts).

conceived of as having relationships between person and person, person and thing and person and the state.[91] Relationships between person and person gave rise to actions *in personam*.[92] Relationships between person and thing gave rise to actions *in rem*.[93] Public law would involve relationships between persons and the state and would include actions by the state against persons.[94] It is also said to include laws of general application.[95] Charitable purpose trusts have a public law quality to them in the sense that they can be enforced by the Crown and therefore involve actions by the state against persons (the trustees).

There is also a private law quality to charitable purpose trusts in that charitable purpose trusts can be created by persons exercising rights with respect to property they own. Persons creating the trust can create their own scheme for the administration of the trust under which persons assenting to act as trustees will be under a legal obligation to carry out. It is, in effect, a legally enforceable promise to the person, or persons, creating the trust made by the persons assenting to be trustees to act for the benefit of others according to the terms of the trust instrument.[96]

There is another "public" quality to charitable purpose trusts, although perhaps not "public law" in the sense described above. A

---

91. See Geoffrey Samuel, "Public and Private Law: A Private Lawyer's Response" (1983) 46 Modern Law Review 558 at 558-59.
92. *Ibid* at 559.
93. *Ibid*.
94. See *e.g.* Herbert M Kritzer, *Legal Systems of the World: A Political, Social, and Cultural Encyclopedia* (Santa Barbara: ABC-CLIO, 2002) vol III at 1339; and *Black's Law Dictionary*, 10d, *sub verbo* "public law" [*Black's Law Dictionary*].
95. Kritzer, *supra* note 94 at 1340; and *Black's Law Dictionary*, *supra* note 94, *sub verbo* "public law".
96. See John H Langbein, "The Contractarian Basis of the Law of Trusts" (1995) 105:3 Yale Law Journal 625 at 627. This connection between trust and contract was also suggested long ago by the well-known equity scholar Maitland. According to Maitland, "the use or trust was originally regarded as an obligation, in point of fact a contract though not usually so called". See Frederic W Maitland, *Equity: A Course of Lectures*, in AH Chaytor and WJ Whittaker, eds, 1d (1909), John Brunyate, ed, rev'd 2d (1936) (Cambridge: Cambridge University Press, 1936; reprinted 1969) at 110.

charitable purpose trust is "public" in the sense of the definition of "public" as "members of the community in general or a particular section of this".[97] This is reflected in the requirement that, for a trust to be valid as a charitable purpose trust, the trust must provide a "public benefit".[98] The main, oft-referred to, case on what constitutes the "public" aspect of "public benefit" is the 1951 decision of the House of Lords in *Oppenheim v Tobacco Securities Trust Company*[99] ("*Oppenheim*"). The purported trust in issue provided for "the education of children of employees or former employees of British-American Tobacco Co Ltd … or any of its subsidiary or allied companies".[100] Since the trust was for education, it fit the charitable purpose of advancement of education. The question was whether there was a public benefit. Lord Simonds held that to constitute the "public" the beneficiaries must be a "section of the community"[101] and he went on to say that:

> these words 'section of the community' have no special sanctity, but they conveniently indicate, that the possible (I emphasize the word 'possible') beneficiaries must not be numerically negligible, and secondly, that the quality which distinguishes them from the community, so that they form by themselves a section of it, must be a quality which does not depend on their

97. This wording is from *The Oxford Paperback Dictionary*, 4d, *sub verbo* "public". The *Oxford English Dictionary*, 2d, *sub verbo* "public" has a similar definition that says "[a] particular section, group, or portion of a community or of mankind". Brody & Tyler, *supra* note 90 argue against the notion that the state is justified in subjecting charities to broad-based government control over their governance, structures, missions, effectiveness, programs and other operations on the basis that property directed to charitable purposes is "public money" (in the sense that the purposes are public purposes) and on the basis that charities are enforced by the state.
98. See *supra* Part II.A.
99. [1951] AC 297 (HL) [*Oppenheim*]. The meaning of "public benefit" is canvassed in Gerald Henry Louis Fridman, "Charities and Public Benefit" (1953) 31 Canadian Business Review 537 which reviews the *Oppenheim* case and cases leading up to the *Oppenheim* case.
100. *Oppenheim*, *supra* note 99 at 298.
101. *Ibid* at 306.

relationship to a particular individual.[102]

It was held that the purported trust did not satisfy the public benefit requirement and therefore did not qualify as a charitable purpose trust. There were over 110,000 employees of the various companies, but whether one was a beneficiary depended on whether one was a child of an employee of the companies and it, therefore, ultimately depended on the relationship to a particular person, namely, the British-American Tobacco Co Ltd as employer.[103]

A justification for the first part of this public benefit or section of the community test (the "not numerically negligible" part) might be that there must be a sufficient benefit from the trust to warrant the potential expenditure of taxpayer funds to enforce it. The test seems to suggest that, even if the benefit is small on a per person basis, the total public benefit is more likely substantial if there are many persons who might benefit. The test, however, refers to the "number of *possible* beneficiaries" and it was specifically noted in the case that the number of persons actually benefiting can be quite small as long as the *possible* number of beneficiaries is high.[104]

It is the second part of the test that attempts to distinguish between trusts created for a public purpose and those created for private reasons. This arguably makes sense in terms of the potential expenditure of taxpayer funds on enforcement. Taxpayer funds should presumably not be expended to enforce a trust created for purely private interests. This justification would, presumably, be much stronger for a charitable purpose trust that is exempt from income tax and for which a tax credit is given for donations to the trust.[105] If a person wants to create a trust that provides private benefits, the mechanism for enforcement should

102. *Ibid*, per Lord Simonds: "a group of persons may be numerous but, if the nexus between them is their personal relationship to a single propositus or to several propositi, they are neither the community nor a section of the community for charitable purposes".

103. *Ibid* at 299.

104. *Ibid* at 306.

105. On these *Income Tax Act* provisions see *supra* notes 50-58 and the accompanying text.

arguably be one that does not rely on a public enforcement mechanism or for which tax subsidies are provided.

While the second part of the test in *Oppenheim* attempts to make a distinction between trusts that provide "public" benefits and those that provide "private" benefits, there are, arguably, problems with how effectively it performs that function. In *Oppenheim* itself Lord MacDermott, dissenting, questioned whether a distinction should be drawn between "those who are employed in a particular industry before it is nationalized and those who are employed therein after that process has been completed and one employer has taken the place of many".[106] Lord MacDermott was, in other words, noting that under the second part of the test a trust for the education of children in a coal industry with many employers would be valid, but after nationalization, with just one employer, the same trust would not be valid.

Another curious qualification on the public-private distinction in the context of charitable purpose trusts is that an exception has been made for charitable purpose trusts for the relief of poverty. Courts have accepted as valid, charitable purpose trusts that are for relief of poverty of "poor relations"[107] or "poor employees".[108] While the House of Lords expressed reservations about such exceptions, these exceptions had been

106. *Oppenheim*, *supra* note 99.

107. See *e.g. Isaac v Defriez* (1754), 27 ER 387 (Ch); and *Re Scarisbrick*, [1951] Ch 622 (CA (Eng)) (a trust for "such relations" as "shall be in needy circumstances"). The exception did not apply on the facts of the *Oppenheim*, *supra* note 99, since there the purported trust was for the advancement of education not the relief of poverty. In *Re Cox* [1951] OR 205 at 224 (CA), aff'd [1953] 1 SCR 94, aff'd [1955] 2 All ER 550 (PC (Canada)) [*Re Cox*], Justice Roach of the Ontario Court of Appeal could see no reason for there being an exception for relief of poverty but not for advancement of education.

108. For England see *e.g. Dingle v Turner*, [1972] 1 All ER 878 (HL); and *Gibson v South American Store (Gath & Chaves) Ltd*, [1950] Ch 177 (CA (Eng)) at 191-97. For Canada see *Re Massey Estate*, [1959] OR 608 (HCt J); and *Eaton*, *supra* note 76.

long accepted in a series of first instance English decisions.[109]

C. Potential "Public" Aspect for Trusts for Persons and Non-Charitable Purpose Trusts

While charitable purpose trusts have been referred to as "public trusts" and express trusts for persons referred to as "private trusts," even private trusts can provide benefits to a community of persons and can, therefore, take on a "public" quality. Non-charitable purpose trusts that were not considered legally valid under the common law can also take on a "public" quality. The potential for each of these types of private trusts to take on a "public" quality is discussed below.

1. Express Trusts for Persons

As noted above, an express trust for persons can describe a class of beneficiaries. The example of a class of beneficiaries given in Part II.A, above, was the "grandchildren" of a particular person. Even if the particular person has hundreds of grandchildren, most persons, it is submitted, would be likely to consider trusts with such a class of beneficiaries, or potential beneficiaries, to be "private" in nature. The class defined in an express trust for persons can, however, be much broader. In the 1971 House of Lords decision in *McPhail v Doulton*[110] ("*McPhail*"), for example, Mr. Baden, by deed, purported to settle property on trust instructing the trustees to "apply the net income of the fund in making at their absolute discretion grants to or for the benefit of any of the officers and employees or ex-officers or ex-employees of the company [Mathew Hall & Co Ltd] or to any relatives or dependents of any such persons in such amounts and on such conditions (if any) as they think fit".[111] The purported trust was not expressed to be for the relief of poverty of these persons or to advance their education. It was, therefore, not a charitable

109. See *Oppenheim*, *supra* note 99 at 306-307. See also, the reservations of the Privy Council as to common employment trusts in *Re Cox*, *supra* note 107 at 552, affirming (1952), [1953] 1 SCR 94. For some of the first instance cases the House of Lords was referring to see *supra* notes 107-108.

110. *McPhail*, *supra* note 9.

111. *Ibid* at 437, 447.

purpose trust. The class of beneficiaries was potentially quite wide. One of the requirements for the establishment of a valid express trust for persons is that there must be certainty of beneficiaries. It was held that the requirement of certainty of beneficiaries is met if the class of beneficiaries is sufficiently clearly defined that "it can be said with certainty that any given individual is or is not a member of the class".[112] When the matter returned to be reassessed by the High Court armed with this guidance from the House of Lords it was found that Mr. Baden's purported trust met this test.[113] This was affirmed by a unanimous Court of Appeal with the words "relatives" and "dependents" being held to meet the test set out by the House of Lords.[114] The test for certainty of beneficiaries set out by the House of Lords in *McPhail* has been accepted in Canada[115] and could, presumably, allow for a very large group of persons. One qualification suggested by Lord Wilberforce in the House of Lords was that the trust objects could not be so wide as to be "administratively unworkable".[116] They could not, he said, be so hopelessly wide as not to form "anything like a class".[117]

The *McPhail* test for certainty of beneficiaries for a discretionary trust can allow for a very broad class of beneficiaries having potentially little direct connection with the settlor. The trust in *McPhail* had such a quality in that it provided that the trustees apply the net income of the fund "for the benefit of any of the officers and employees or ex-officers or ex-employees of the company or to any relatives or dependants of any such persons in such amounts at such times and on such conditions ...

112. *Ibid* at 450.
113. *Re Baden's Trust Deeds (No 2)*, [1971] 3 All ER 985 (Ch), aff'd on appeal, [1972] 2 All ER 1304 (CA).
114. *Ibid*.
115. See *e.g. Eaton*, *supra* note 76; *Dickson v Richardson* (1981), 32 OR (2d) 158 (CA); *Canada Trust Co v Ontario Human Rights Commission* (1990), 74 OR (2d) 481 (CA); and *Lewis v Union of BC Performers* (1996), 18 BCLR (3d) 382 (CA).
116. *McPhail*, *supra* note 9 at 441.
117. *Ibid*. Lord Wilberforce suggested that "[a]ll the residents of Greater London," might be an example.

as they think fit".[118] The range of potential beneficiaries in such a trust could be quite broad. The nature of the "benefit" was apparently not constrained, and therefore was not "charitable" in the legal sense[119] and might also have not been considered charitable in a popular sense, except, perhaps that it likely had a wealth redistribution quality since the settlor, Mr. Baden, was likely reasonably well off relative to the employees, their relatives and dependents.

Eaton[120] also provides an example of the potential for a discretionary express trust to have a public quality in providing for a potentially wide range of persons in a way that might be considered socially beneficial. In that case the testator provided a specified amount be paid to the executive officers of the Eaton Company to be used "for any needy or deserving Toronto members of the Eaton Quarter Century Club". The club was not a charitable organization, but a social club, the members of which were persons who had been in the service of the Eaton Company for twenty-five years or more. By the time the litigation arose the club had 7,000 members who had worked, or were still working, for the Eaton Company, or who had retired. "Needy" was accepted as a charitable purpose relating to relief of poverty. The issue from the charitable trust perspective was whether "deserving" was for a charitable purpose. The trust was upheld as a charitable purpose trust interpreting "deserving" as what the testator presumably had in mind in the 1934 Great Depression context of the trust.[121] The trust was also upheld as a trust for persons applying the *McPhail* test interpreting "Toronto members" to mean "those members who were employed by the company in Toronto at the time when they became members".[122] While the number of beneficiaries would have been less than all the 7,000 members of the club at the time of the litigation,

118. *Ibid* at 433.
119. It was not "charitable" in the legal sense since it was not limited to the traditional heads of charitable purposes and the funds could, therefore, have been used for non-charitable purposes. It would, therefore, not have been charitable under the exclusivity requirement for a valid charitable purpose trust (see Part II.C.2 above).
120. *Eaton*, *supra* note 76.
121. *Ibid* at 642-43.
122. *Ibid* at 651.

it would, given the number likely employed by the company in Toronto, presumably have included a fairly large number of persons. As *Eaton* demonstrates, an express trust for persons can have a large number of beneficiaries and can have a public quality with the beneficiaries having a very limited connection to the settlor of the trust and with the trust having a public purpose quality given that the funds were to be used for persons who were "needy or deserving".

While the trust in *Oppenheim* was found, in 1951, not to meet the requirement of "public benefit" given the way "public benefit" was defined in that case, the funds were to be used for the education of the beneficiaries which was considered "charitable" under the traditional legal definition of "charitable purposes" and otherwise might be said to provide a benefit to a potentially large number of persons. If one were to argue the *Oppenheim* case today, one might well argue that it meets the 1972 *McPhail* test for certainty of beneficiaries.[123] The trust might, therefore, be valid as an express trust for persons with the potential number of beneficiaries being very large and with a purpose that might be considered to have a public quality.

Even broader examples of the potential use of an express trust for persons with public qualities could be provided. An express trust for persons might, for instance, be used to create a form of governance for a broad group of persons such as a particular Aboriginal group, assuming such persons could be sufficiently clearly defined to meet the *McPhail* test.[124] An express trust for persons might also be used to set

123. *Oppenheim*, *supra* note 99 was decided in 1951 long before *McPhail*, *supra* note 9, which was decided in 1972. The beneficiaries in *Oppenheim* were "children of employees or former employees of British-American Tobacco Co Ltd … or any of its subsidiary or allied companies" at 298 (see *supra* note 100 and the accompanying text) and this description of a class of beneficiaries, it is submitted, would probably meet the test of certainty for a discretionary trust for persons under the *McPhail* test.

124. See *e.g.* Daryn R Leas & Victoria B Fred, "Yukon First Nation Trusts: Establishing a Legacy for Future Generations", online: Council of Yukon First Nations <www.sgsyukon.ca> (on this use of trusts for First Nations).

up a community ecosystem trust[125] or for an environmental restoration trust.[126]

2. Non-Charitable Purpose Trusts

A non-charitable purpose trust can have a public quality both in terms of the number of potential beneficiaries (who may well not be related to the settlor), and because the purpose may provide a public, or broad community, benefit. The facts in *Re Lipinski's Will Trusts*[127] provide an example of this. In that case the testator made a bequest to the Hull Judeans (Maccabi) Association "in memory of my late wife to be used solely in the work of constructing the new buildings for the Association and/or improvements to the said buildings".[128] The Hull Judeans (Maccabi) Association was an unincorporated association that had been established to provide "social, cultural and sporting activities for Jewish youth in Hull, or for such Jewish youth in Hull as become members

125. On community ecosystems trusts see Michael McGonigle, Brian Egan & Lisa Ambus, *The Community Ecosystem Trust: A New Model for Developing Sustainability* (Victoria: POLIS Project on Ecological Governance, University of Victoria, 2001), online: POLIS Project on Ecological Governance < www.polisproject.org>.

126. On the use of an express trust to set up an environmental restoration trust see Donovan WM Waters, "The Role of the Trust in Environmental Protection Law" in Donovan WM Waters, ed, *Equity, Fiduciaries and Trusts 1993* (Toronto: Carswell, 1993) 383 at 405-19.

127. [1976] Ch 235 (Eng) [*Re Lipinski's*]. The focus here is on borrowing the facts from *Re Lipinski's* case as an example of a non-charitable purpose trust that has a public benefit quality. The testator's gift was upheld in the case but while different possible bases for upholding the gift were discussed, it is not clear from the case on which basis it was upheld (or if it was upheld on each basis discussed). On the difficulty in pinning down on just what basis the trust was upheld see *e.g.* Nigel P Gravells, "Gifts to Unincorporated Associations: Where There's a Will There's a Way" (1977) 40:2 Modern Law Review 231 generally and especially where it is said that, "the reasoning in the judgment and the criteria adopted to avoid the conclusion of total invalidity of the gift may create yet more uncertainty for draftsmen in the future" at 236.

128. Gravells, *supra* note 127 at 237.

of the association".[129] Such a trust might be valid as a trust for persons (consistent with *Re Denley's*[130]). If, however, it could not be found valid as a trust for persons and was held to be a non-charitable purpose trust,[131] it would nonetheless potentially benefit a large number of persons (Jewish youth in Hull) who might be said to represent a section of the community. That section of the community may well have had a minimal connection to the settlor and the purpose of the intended trust, "social, cultural and sporting activities," might have been considered one that provided a public benefit.

Many other non-charitable purpose trusts might have a public quality in the sense of benefiting a substantial community of persons and in a way that might be considered to be for the general benefit of society. Suppose, for example, funds were raised to be held on trust for the purpose of promoting the development of technologies that would reduce global warming. Such a trust would likely not be charitable in many common law jurisdictions but it could provide what, it is submitted, would be accepted by many persons as a public benefit and that benefit could be a benefit for virtually everyone in the world.

IV. Enforcement Problems Across All Three Types of Express Trusts

This part notes the potential for weak enforcement is not limited to non-charitable purpose trusts.[132] Weak enforcement can also arise for charitable purpose trusts and even in express trusts for persons.

---

129. *Ibid* at 238.
130. See *Re Denley's*, *supra* note 39.
131. Justice Oliver held in *Re Lipinski's*, *supra* note 127 at 243, that the purposes of the unincorporated Hull Judeans (Maccabi) Association were not charitable purposes. To the extent, therefore, that the members of the association were to hold the funds on trust to build or improve buildings for the association, or otherwise for the purposes of the association, it would not be a trust for charitable purposes.
132. See *supra* notes 20-23 and the accompanying text on enforcement problems as a primary reason for the non-validity of non-charitable purpose trusts.

A. Addressing Enforcement Problems for Non-Charitable Purpose Trusts

The lack of a person in whose favour the court can decree performance has been given as the basis for the non-validity of non-charitable purpose trusts. There are no beneficiaries to enforce such trusts. Where the creator of the trust has retained the services of a knowledgeable trust draftsperson, the trust draftsperson, recognizing that the settlor's purpose is non-charitable or that there is significant doubt as to whether the purpose is charitable, has to find a way to either reconceive of the trust as a trust for persons, fit it into a recognized charitable purpose or abandon the trust form in favour of some other form such as a corporate form.[133] Such attempts to reconceive the trust are likely not to achieve what the settlor wants. They may, indeed, fall considerably short of what the settlor wants and may fall so far short that the settlor abandons the notion of devoting funds to his or her intended purpose. If the settlor's non-charitable purpose is reasonably certain and is not illegal or contrary to public policy, there seems no reason not to attempt to achieve the settlor's purpose by permitting the purpose to be enforced as a trustee obligation and not just as a trustee power as long as some means of enforcement can be provided.[134] Where the settlor's purpose happens to benefit a reasonably substantial section of the community there would seem to

133. A corporate form (likely a not-for-profit corporate form) would have members who could enforce the objects of the corporation. An unincorporated association might also be considered in which members would have a contractual basis for enforcing compliance with the objects of the unincorporated association.

134. Enforcement as a power is the approach taken by wait-and-see perpetuities legislation in Canada (see *supra* notes 34-36 and the accompanying text) and the approach effectively taken in *Re Thompson*, *supra* note 28 (see *supra* notes 32-33 and the accompanying text).

be all the more reason to facilitate its enforcement.[135] The grounds for facilitating enforcement of the settlor's purpose would, it is submitted, be that much stronger where the settlor has provided funds to be used for a non-charitable purpose to which taxpayer funds are currently directed.[136]

While other forms, such as a not-for-profit corporation or an unincorporated association might achieve the same objective, the trust may provide the advantage of flexibility. There will also be situations in which persons without legal assistance, or sometimes even with legal assistance, purport to create trusts for non-charitable purposes instead of relying on alternate forms. It is arguably unfortunate that such trusts, particularly when they are for public purposes, are invalid and either operate not at all or only as powers.

B. Enforcement Problems for Charitable Purpose Trusts

Although the Crown can enforce charitable purpose trusts, it has been suggested that:

> [i]t is probable that, without any previous complaint or critical information being brought to the Attorney General's attention, the prerogative to investigate charities, if it exists, would not be exercised. And, even when the Attorney General does receive such a complaint or information, it is entirely within the

135. In *Re Lipinski's*, *supra* note 127 discussed above, Part III.C.2, the gift in question was upheld but the basis for enforcement was not clear. James Phillips notes in Phillips, *supra* note 21 that "commentators have had difficulty knowing how to characterize *Lipinski*" at 180. It has been said that *Re Lipinski* "confirmed in this particular sphere a continuing determination on the part of the judiciary not to allow the clear intentions of testators to be defeated" see Gravells, *supra* note 127 at 231.

136. A trust that, for example, provided for assisting immigrants in finding jobs and integrating into the society of the country they had migrated to might reduce tax dollars spent on supporting such persons and might lead to more tax revenues from immigrant employment income. A trust to provide for recreational facilities might reduce the expenditure of local taxpayer dollars on recreational facilities.

> attorney's discretion whether the office should take any legal or other action. [137]

Indeed, Canadian cases dealing with charitable purpose trusts rarely, if ever, arise in the context of a breach of trust action pursued by the Crown.[138] While the CRA may also have a role with respect to the enforcement of charitable purpose trusts, its focus may not be entirely consistent with the interests of the persons who might benefit from pursuit of the purposes of a charitable purpose trust. If, for instance, charitable purpose trust funds are being used for charitable purposes but not for the particular charitable purposes the settlor clearly intended, the CRA, given that the funds are at least being used for some public benefit falling within the legal definition of charitable purpose, may not be particularly concerned, from the income tax perspective, that the funds are not being directed to the purpose the settlor intended. The CRA is unlikely to have the resources, and perhaps, from an income tax administration perspective, the incentive, to monitor charitable purpose trusts to assess whether they are devoting trust funds to their intended

---

137. Waters, Gillen & Smith, *supra* note 6 at 835. See also the discussion of the lack of Crown enforcement of charitable purpose trusts in Chan, *supra* note 48 at 389-93. It is perhaps because, practically speaking, the enforcement of charitable purpose trusts is relatively weak that charities in Canada are typically organized as corporations. With a charity organized as a not-for-profit corporation the prospect of the corporation's members enforcing the corporation's charitable objects is likely greater than the prospect of Crown enforcement of charitable purposes of a charity organized as a trust.

138. While the author has not made a proper attempt to verify this empirically, a search through numerous (not all) Canadian cases involving an allegation of breach of trust did not yield a single case in which the Crown in right of a province (either through the Attorney General or nominee such as a Public Trustee or Public Trustee and Guardian) appeared as a party in the case name. The author also does not recall a single Canadian breach of trust case in which the Crown in right of a province pursued a breach of trust claim in respect of a charitable purpose trust. The assertion that Canadian cases dealing with charitable purpose trusts rarely, if ever, arise in the context of a breach of trust action pursued by the Crown appears to the author to be a reasonable one. See also the discussion in Chan, *supra* note 48 at 389-93.

charitable purposes in an efficient manner. It is also doubtful whether the CRA has the statutory mandate to prevent a trustee from deviating from specific intended charitable purposes as long as the trustee is spending the funds for charitable purposes (or activities) or to deal with a trustee's failure to expend the trust funds in an efficient manner.[139]

Charitable purpose trusts may, therefore, also operate with very limited enforcement. A trust draftsperson aware of this might consider advising the settlor of this problem and suggest as an alternative that the trust be reconceived as a trust for persons. This would at least provide beneficiaries who could enforce the trust instead of leaving enforcement to the Crown. If, however, the class of beneficiaries was quite large and if the trustee was given discretion over to whom distributions of income and capital were to go, it is likely none of the beneficiaries would have much of an incentive to enforce the trust and one might be no further ahead in terms of increasing the likelihood that the trust would be enforced.

C. Potential Enforcement Problems in Express Trusts for Persons

Weak enforcement may also arise with an express trust for persons. While the beneficiaries of express trusts are often individuals with a close relationship to the settlor, an express trust can, as noted in Part III.C.1 above, be a discretionary express trust with a broadly-defined, but conceptually certain, class of beneficiaries. The probability of the trustees exercising their discretion in favour of any particular persons within the class of beneficiaries may be low enough that none of the beneficiaries has

139. The *Income Tax Act*, *supra* note 50, allows for a tax credit for donations to registered charities. A "registered charity" is a "charitable organization" that devotes its resources exclusively to charitable activities or a "charitable foundation" that distributes is funds to charitable organizations. See *supra* note 51. The Canada Revenue Agency could refuse registration or remove registration if these requirements were not met. While trustees of a charitable organization must devote the trust funds to charitable activities, the *Income Tax Act* does not specifically provide for action to be taken against trustees who fail to use trust funds for the specific purposes intended by the settlor or who fail to exercise a fiduciary duty of care in their use of trust funds for charitable purposes.

much of an incentive to enforce the trust. Expressing it in the way it was expressed in *Morice v The Bishop of Durham*, there would be "[persons] in whose favour the court [could] decree performance"[140] and, assuming requirements for a valid express trust were met (in addition to certainty of beneficiaries), the trust would be a valid trust, but the probability of enforcement would, nonetheless, be low. With no one likely to ever enforce the obligations the trustee has assented to, the court enforcement feature that arguably gives a trust its "legal" quality is, practically speaking, absent.

V. Expanding the Range of Persons Who Can Enforce Express Trusts

A. Proposed Enforcement of Certain Non-Charitable Purpose Trusts under Section 74 of the *Uniform Trustee Act*

The *Uniform Trustee Act* was the result of a project initiated by the Civil Section of the Uniform Law Conference of Canada in 2008.[141] It is intended to be suitable for enactment across the country.[142] Section 74 of the *Uniform Trustee Act* would allow some non-charitable purpose trusts to operate as valid trusts and, unlike the non-charitable purpose trust provisions in perpetuities acts,[143] would allow for the enforcement of trust duties instead of simply having the trust operate as a power.[144] Like the perpetuity act provisions, the trust would have to be one that "does not create an equitable interest in any person"[145] and that is

140. *Supra* note 12 at 658.
141. See Uniform Law Conference of Canada, Civil Section, *Uniform Trustee Act: Final Report to the Working Group*, (Whitehorse: 2012) at para 1.
142. *Ibid.*
143. The wait-and-see perpetuities legislation provisions are discussed in *supra* notes 34-36 and the accompanying text.
144. *Uniform Trustee Act*, *supra* note 2, s 74(10) allows a court to make an order it considers appropriate to enforce a non-charitable purpose trust. The comment on s 74 says that, "this section validates the creation of trusts for certain non-charitable public purposes".
145. *Uniform Trustee Act*, *supra* note 2, s 74(2)(a).

"sufficiently certain to allow the trust to be carried out".[146] The types of non-charitable purpose trusts that are made valid by section 74(3) are those that are "recognized by law as capable of being a valid object of a trust",[147] those with "purposes for which a society may be incorporated

146. *Ibid*, s 74(3)(a). This appears to capture the essence of the "specific non-charitable purpose trust" language in the wait-and-see perpetuity legislation as it was interpreted in *Re Russell*, *supra* note 31. *Uniform Trustee Act*, *supra* note 2, s 74(3)(c) adds that the non-charitable purpose not be contrary to public policy.

147. These would presumably include the recognized exceptions of: (i) the erection of a monument at a gravesite; (ii) the maintenance of a gravesite; and (iii) the provision of food and shelter for specified animals. See the discussion *supra* notes 24-31 and the accompanying text. This is, indeed, what the commentary to *Uniform Trustee Act*, *supra* note 2, s 74 says. Would it, however, include non-charitable purpose trusts recognized on the basis suggested in *Keewatin*, *supra* note 31, that there are persons to whom the court could give standing to enforce the trust?

under the jurisdiction's legislation respecting societies",[148] those for the performance of a function of government in Canada or those provided

148. *Uniform Trustee Act*, *supra* note 2, s 74(3) is in italics and set out in square-brackets presumably as a suggested way to identify the "public purposes" the section is intended to validate (see the commentary to s 74). For provinces or territories with Acts styled "Societies Act" or "Society Act" see *e.g.* the Nova Scotia *Societies Act*, RSNS 1989, c 435; the British Columbia *Society Act*, RSBC 1996, c 433 soon to be replaced by the *Societies Act* SBC 2015, c 18; the Alberta *Societies Act*, RSA 2000, c S-14; the Yukon *Societies Act*, RSY 2002, c 206; and the Northwest Territories *Societies Act*, RSNWT 1988, c S-11. Not all provinces or territories have Acts styled "Societies Act" or "Society Act" but have instead Acts for "not-for-profit" or "non-profit" corporations – see *e.g.* the Ontario *Not-for-Profit Corporations Act*, SO 2010, c 15; and in Saskatchewan *The Non-profit Corporations Act, 1995*, SS 1995, c N-4.2. Some provinces provide for non-profit corporations in a statute that covers both for-profit and non-profit corporations – see *e.g.* for Manitoba *The Corporations Act*, CCSM c C225 [*Manitoba Corporation Act*]; and the Newfoundland *Corporations Act*, RSNL 1990, c C-36. British Columbia *Societies Act*, provides that: "a society may be formed under this Act for one or more purposes, including, without limitation, agricultural, artistic, benevolent, charitable, educational, environmental, patriotic, philanthropic, political, professional, recreational, religious, scientific or sporting purposes" at s 2(1). The Nova Scotia *Societies Act* provides that "[a] society may be incorporated under this Act to promote any benevolent, philanthropic, patriotic, religious, charitable, artistic, literary, educational, social, professional, recreational or sporting or any other useful object, but not for the purpose of carrying on any trade, industry or business" at s 3(1). In Prince Edward Island Part II of the *Companies Act*, c C-14 allows for the incorporation by letters patent of a body corporate "without share capital, for the purpose of carrying on in Prince Edward Island, without pecuniary gain to its members, objects of a patriotic, religious, philanthropic, charitable, scientific, artistic, social, professional or sporting character, or the like" at s 89. Similarly, the *Manitoba Corporation Act* provides that a corporation without share capital must "restrict its undertaking to one that is only of a patriotic, religious, philanthropic, charitable, educational, agricultural, scientific, literary, historical, artistic, social, professional, fraternal, sporting or athletic nature or the like" at s 267(1).

for by regulation.[149]

Section 74(10) would allow a court to make an order it considers appropriate to enforce a non-charitable purpose trust or to enlarge, or otherwise vary, the powers of the trustee of a non-charitable purpose trust.[150] Section 74(11) provides that:

> [a]n application for an order under this section may be made by any of the following:
>
> (a) the Attorney General;
>
> (b) a person appointed specifically by the settlor in the trust instrument to enforce the trust;
>
> (c) the settlor;
>
> (d) the personal representative of the settlor;
>
> (e) the trustee; or
>
> (f) a person appearing to the court to have a sufficient interest in the matter.

This list of persons who can seek a court order to enforce the non-charitable purpose trusts made valid by section 74 addresses the main long-standing reason for the non-validity of non-charitable purpose trusts, namely, the lack of a beneficiary to enforce such trusts.

The provision in section 74(3)(c)(i) providing for the validity of a non-charitable purpose trust with a purpose for which a society may be incorporated is set out in italics and with square brackets perhaps because it is just a suggestion. It is a very broad suggestion that, given the breadth of societies legislation purpose provisions,[151] might well include trusts

149. *Uniform Trustee Act*, *supra* note 2, ss 74(3)(c)(iii) and (12). Each jurisdiction would provide its own regulations and could add to the list of purposes of non-charitable purpose trusts to which s 74 would apply. Section 74 would also apply to non-charitable purpose trusts created by a court making an order under s 75(3)(b), (d) or (f) which gives the court certain powers to deal with mixed charitable and non-charitable purpose trusts.

150. *Uniform Trustee Act*, *supra* note 2, ss 74(5)-(10) set out court powers to make orders. Subsections 74(5)-(9) set out a somewhat modified version of court *cy-près* powers with respect to charitable purpose trusts. These *cy-près* powers are necessary because s 74(4) allows for the limited types of non-charitable purpose trusts made valid by s 74 to exist indefinitely.

151. See *supra* note 148 on the scope of societies' legislation provisions on possible society purposes.

that have little or no public quality. In that respect it may be too broad to the extent it might allow for enforcement of effectively private trusts by the Attorney General which would arguably not be a proper expenditure of taxpayer funds.

The *Uniform Trustee Act* effectively creates a fourth type of express trust since it provides in section 74 for the enforcement of trust duties for only certain types of non-charitable purpose trusts as described in section 74(3), leaving the remaining non-charitable purpose trusts to be dealt with under section 76 in a manner similar to the non-charitable trust provision in wait-and-see perpetuities legislation. Section 76(2) provides that,

> If the terms of a disposition of property purport to create a trust that
>
> > (a) does not create an equitable interest in any person, and
> > (b) is for a specific non-charitable purpose, other than a non-charitable purpose described under s. 74(3),
>
> the terms of the disposition must be construed, subject to this section, as constituting a power to appoint the income or the capital, as the case may be, for a period not exceeding 21 years.

Under the *Uniform Trustee Act* there would, therefore, be express trusts for persons, express trusts for non-charitable purposes that can be enforced under section 74, express trusts for specific non-charitable purposes that would, under section 76, operate as powers to appoint income or capital for the specific non-charitable purposes[152] and express trusts for charitable purposes.

B. Allowing the Settlor to Add Trust Enforcers

The persons who can apply for a court order in the list in section 74 that are of particular interest are: the settlor, the personal representative of the settlor or a person appointed specifically by the settlor. None of

152. If the suggestion in *Uniform Trustee Act*, *supra* note 2, s 74(3)(c)(i) of allowing for purposes for which a society might be incorporated were followed there would seem to be relatively little room left for non-charitable purpose trusts under s 76 given the scope of purposes for which a society may be incorporated. On the scope of purposes for which a society may be incorporated see *supra* note 148.

the persons in this list can apply to court to enforce an express trust for persons or a charitable purpose trust. It seems a curious result that the set of non-charitable purpose trusts made valid by section 74 would have access to a greater range of enforcement mechanisms than express trusts for persons or charitable purpose trusts. It might be said that the reason for the difference is that express trusts for persons and charitable purpose trusts have their own enforcement mechanisms — the beneficiaries for express trusts for persons and the Crown for charitable purpose trusts. Perhaps, however, it would make sense to expand the scope of enforcement mechanisms for express trusts for persons and for charitable purpose trusts, that could give the trusts draftsperson greater scope for creating an enforcement mechanism for the trust. Perhaps also a greater scope of enforcement mechanisms could be extended to non-charitable purpose trusts more generally and not just to those that are made valid by section 74.

The expanded scope of enforcement mechanisms for express trusts for persons and non-charitable purpose trusts not made valid by section 74 should not include enforcement by the Attorney General for all such trusts since many such trusts are likely to have private objects.[153] It arguably does not make sense to have the Attorney General, at taxpayer expense, enforce trusts created for the pursuit of private objects. Crown enforcement should, however, as discussed in Part VII below, extend not to just charitable purpose trusts but to non-charitable purpose trusts and express trusts for persons that provide sufficient public benefits to justify expenditure of taxpayer funds on enforcement.

1. Express Trusts for Persons

While the rights of beneficiaries to enforce express trusts for persons should remain, legislation might allow the settlor to indicate in the trust instrument that any one or more of the persons in (b) through (d)

153. Also, as noted in notes 149-50 above and the accompanying text, a scope of non-charitable purpose trusts made valid by the *Uniform Trustee Act*, *supra* note 2, s 74 that includes any purpose for which a society may be incorporated is likely to be too broad since it may well include trusts that have little or no public benefit.

of section 74(11) be given standing to enforce the trust (*i.e.* a person appointed specifically by the settlor in the trust instrument to enforce the trust, the settlor or the personal representative of the settlor). In the context of a discretionary express trust for persons with a broadly defined class of beneficiaries who might have little incentive to enforce the trust, legislative support for such a broader range of persons able to enforce the trust would allow the trust draftsperson (in consultation with the settlor) to consider and, if thought appropriate, allow for these additional means of enforcement. It might also be helpful to allow the trust draftsperson (in consultation with the settlor) to indicate in the trust instrument that any one or more of such persons be given standing to enforce the trust where it has unborn, minor or other incapacitated beneficiaries to better ensure enforcement on behalf of such beneficiaries.[154]

One problem with naming persons to enforce a trust is that they may have no more incentive to enforce the trust than the trustee has to carry out its terms. One might, therefore, ask: who monitors the enforcer? This may also be true of the personal representative of the settlor. Perhaps even the settlor will have little incentive to enforce once she or he has disposed of her or his interest in the trust property to the trustee. Why would the settlor not simply start by naming as trustee a person whom the settlor genuinely trusts to carry out the obligation regardless of whether there will ever be, for practical reasons, any enforcement of the legal obligation? There may, however, be situations where the person chosen as trustee has particular skills that will be beneficial in managing the trust. For instance, the trustee may have particular skills in investing trust funds and skills with respect to managing other trust assets or in the distribution of trust funds. The settlor may feel more comfortable with other persons seeing

154. The person chosen by the settlor to enforce the trust might have a greater incentive to monitor the day-to-day operation of the trust than a public official who might be subject to many demands on limited resources. Even if the settlor appointed someone as enforcer, it might, nonetheless, be prudent, in the interests of protecting the minor, unborn or incapacitated beneficiary, to retain enforcement powers of a public official, such as a Public Trustee and Guardian, even where the settlor had appointed an enforcer (and even where the appointed enforcer refused to enforce the trust).

to it that the chosen trustee, through the threat of enforcement, has an incentive carry out the trust obligations and do so in good faith and with care.

2. Charitable Purpose Trusts

Consideration should also be given to allowing settlors to name enforcers for the enforcement of charitable purpose trusts. This would allow them to provide for ongoing monitoring and enforcement of a charitable purpose trust that likely would be more effective than enforcement by the Crown. It would also allow for enforcement without expenditure of taxpayer funds.

3. Non-Charitable Purpose Trusts

As noted above, section 74 of the *Uniform Trustee Act* only provided for the enforcement of trust duties for certain types of non-charitable purpose trusts, leaving other types of non-charitable purpose trusts to operate as powers pursuant to section 76. Legislative support might also be provided to allow these other non-charitable purpose trusts to be enforced not just as powers, but as trust duties by giving standing to enforce to persons the settlor appoints in the trust instrument to enforce the trust. This is an approach that was suggested by the Manitoba Law

Reform Commission in 1992.[155] The persons entitled to enforce non-charitable purpose trusts might also be extended to the settlor or the settlor's personal representative. The trustee of a non-charitable purpose should also be given standing to apply to court to vary a non-charitable

155. Manitoba Law Reform Commission, *Report on Non-Charitable Purpose Trusts*, no 77 (1992). This Manitoba Law Reform Commission report also recommended that the enforcer be someone other than the trustee and that the enforcer be subject to the supervision of the court. Further, if no enforcer had been appointed, the trustee would have a duty to appoint one. It is interesting to compare the notion in *Uniform Trustee Act*, *supra* note 2, s 74(11)(b) of giving standing to a person appointed specifically by the settlor in the trust instrument to enforce the trust to opting instead to make a gift to a non-profit corporation (or society) with purposes set out in its constitution and a board of directors appointed by members of the corporation. The purposes might well be purposes similar to those in a non-charitable purpose trust (although they would have to fit the permitted purposes for a non-profit corporation). The board of directors might be seen as functionally equivalent to trustees of a purpose trust and the members who elect, or remove, members of the board of directors might be seen as an enforcement mechanism for seeing to it that the directors carry out the purposes of the corporation and do so efficaciously. With this comparison in mind, the linking of the limited set of non-charitable purpose trusts made valid by s 74(3)(c)(i) of the *Uniform Trustee Act*, *supra* note 2, to the range of purposes set out in societies acts or their equivalents makes sense. Section 74 could, in this light, be seen as facilitating the use of an alternative organizational form, a trust, for the pursuit such purposes.

purpose trust.[156]

Allowing non-charitable purpose trusts to be enforced as trust duties, and not just as powers, gives greater scope to carrying out the intent of the settlor. It seems reasonable to carry out a settlor's purpose trust intent if the trust intent is sufficiently clear to permit enforcement of the trust and the purpose is not illegal or contrary to public policy. The interest in carrying out the settlor's intent would arguably be strengthened where the non-charitable purpose happens to be one that provides a benefit to a reasonably broad section of the community.

VI. Allowing Courts to Grant Standing to Persons to Enforce Express Trusts of Any Type

As noted in Part IV, enforcement can be weak for charitable purpose trusts and for discretionary express trusts for persons with large numbers of beneficiaries. Empowering the court to grant standing to persons to enforce not only non-charitable purpose trusts, but also charitable purpose trusts and express trusts for persons could facilitate the enforcement of

156. In 1991 the British Columbia Law Reform Commission suggested in British Columbia Law Reform Commission, *Working Paper on Non-Charitable Purpose Trusts*, no 128 (1991) that standing to enforce non-charitable purpose trusts be extended to a trustee [*BC Working Paper*]. *Uniform Trustee Act* s 74(11)(e) would also allow the trustee to enforce a non-charitable purpose trust. This would allow the trustee to apply pursuant to subsections 74(5)-(10) of the *Uniform Trustee Act*, *supra* note 2, to vary a non-charitable purpose trust where it is of a type made valid by s 74. A provision for the variation of a non-charitable purpose trust is necessary because a non-charitable purpose trust would not have identifiable beneficiaries and therefore could not be varied by invoking *Saunders v Vautier* (1841), 49 ER 282 (Ch) or under variations of trusts legislation. It would be particularly important to allow for the variation of non-charitable purpose trusts if the rule against perpetuities is abrogated as is proposed under s 88 of the *Uniform Trustee Act*, *supra* note 2, and as has occurred in in three provinces (for Manitoba see the *Perpetuities and Accumulations Act*, CCSM, c P33, ss 2-3 [*Manitoba Perpetuities and Accumulations Act*]; for Saskatchewan see *The Trustee Act*, 2009, SS 2009, c T-23.01, s 58 [*Saskatchewan Trustee Act*]; and for Nova Scotia see the *Perpetuities Act*, SNS 2011, c 42, ss 2-3 [*Nova Scotia Perpetuities Act*].

such trusts when other means of enforcement are weak.

Section 74(11) of the *Uniform Trustee Act* takes a similar approach by providing, in section 74(11)(f), that an application may be brought by "a person appearing to the court to have a sufficient interest in the matter".[157] A similar provision could be enacted for express trusts for persons or charitable purposes allowing a person appearing to the court to have a sufficient interest in the matter to make such an application.

Opening the scope for applications by anyone appearing to the court to have a sufficient interest in the matter might raise a concern for vexatious actions or actions not in the interests of the trust.[158] This might be addressed by constraints similar to those used in corporate statutory derivative action provisions. For instance, one might require that the trustees be notified prior to the application to court for standing so the trustees have an opportunity to respond to alleged breaches and avoid potentially expensive litigation.[159] One might also require that the applicant be acting in good faith and that the action appear to be in the best interests of the trust (*i.e.* in the interests of accomplishing the intent of the trust and in a cost-justified manner).[160] Other features of the corporate statutory derivative action might also be borrowed. For instance, requiring court approval of settlements might discourage attempts by the applicant to pressure trustees into giving the applicant favourable treatment and for court orders giving directions with respect to the conduct of an action.[161] Provisions relating to interim costs and

157. The British Columbia Law Reform Commission in *BC Working Paper*, *supra* note 156, would also have allowed the court to grant standing to enforce a non-charitable purpose trust to "any person appearing to the Court to have a sufficient interest in the enforcement of the trust" at 33.
158. For instance, actions that could significantly drain trust assets with little to be gained for the beneficiaries or purposes of the trust.
159. For a roughly corresponding corporate derivative action provision see *Canada Business Corporations Act*, RSC 1985, c C-44, s 239(2)(a) [*CBCA*].
160. For roughly corresponding corporate derivative action requirements see *CBCA*, *supra* note 159, ss 239(2)(b), (c).
161. For related corporate derivative action provisions see *CBCA*, *supra* note 159, ss 240(b) and 242(2).

final cost awards might also be considered.[162]

In keeping with the approach in Part V.B, above, of allowing settlors to add to the list of persons who can enforce a trust, one might allow settlors to decide whether the court should have a power to grant standing to persons to enforce the trust. This could be an opt-in approach where the statutory provision granting the court such a power would only operate if the settlor so indicated. It might instead be designed as an opt-out approach in which the settlor could indicate that the statutory provision granting the court a power to grant standing to persons (other than beneficiaries) was not to operate.

VII. Crown Enforcement of Specified Express Trusts

The Crown, as noted in Part II.C.1 above, has a prerogative power to enforce charitable purpose trusts. Trusts other than charitable purpose trusts, including both non-charitable purpose trusts and express trusts for persons, may, as noted in Part III.C above, have a public quality in some circumstances. That public quality may justify expenditure of Crown funds in the enforcement of such trusts. The *Uniform Trustee Act* appears to recognize this in the context of non-charitable purpose trusts since it proposes providing for potential Crown enforcement of certain specified non-charitable purpose trusts. Since it is possible for express trusts for persons to have a public quality, it arguably makes sense to extend Crown enforcement to express trusts for persons where they have a similarly sufficient public quality to justify the expenditure of public funds in enforcing such trusts.

If the *Uniform Trustee Act* proposal for potential Crown enforcement of certain types of non-charitable purpose trusts is followed, there would be two bases for potential Crown enforcement of trusts. One would be the traditional enforcement of charitable purpose trusts said to be

162. For corporate derivative action provisions see *CBCA*, *supra* note 159, ss 240(d), 242(3), 242(4). Section 240(d) allows the court to order the corporation to pay reasonable legal fees incurred by the complainant (and here corporation might be replaced with payments out of trust funds). Section 242(3) concerns security for costs, and s 242(4) concerns interim costs.

based on a Crown prerogative. The other would be potential Crown enforcement of non-charitable purpose trusts under a statutory authority. The justification for both forms of Crown enforcement is presumably the same — the presence of sufficient public interest in enforcement to make the potential expenditure of Crown funds on enforcement worthwhile. If the justification for the potential for Crown enforcement is the same for both charitable and non-charitable purpose trusts, then why should the basis of Crown enforcement be different? Grounding Crown enforcement on a statutory power makes Crown enforcement more directly part of the political process since elected representatives would have to turn their minds to appropriate circumstances for Crown enforcement and be responsible to the electorate for their decisions in that regard. It is, therefore, also recommended that the Crown prerogative to enforce charitable purpose trusts be replaced with a statutory authority. The statute would set out the types of trusts for which Crown enforcement would be available. The potential for Crown enforcement would then be based on a modern statute rather than the preamble to an English statute from over 400 years ago.[163] In other words, the statutorily identified types of trusts for which there would be potential Crown enforcement would replace the current common law on what constitutes a charitable purpose. Potential for Crown enforcement would extend to all express trusts that were identified as having the requisite public quality to justify

163. This approach is similar to the approach in the English *Charities Act 2011*, *supra* note 66. This act lists charitable purposes in ss 3(1)(a) to (l) that capture the main elements of the charitable purpose categories in *Pemsel*, *supra* note 14 and accompanying text, while adding several new charitable purposes. It therefore effectively adds to the list of the types of trusts that can be subject to Crown enforcement (although through the various mechanisms provided for in the Act). The English *Charities Act 2011* differs, however, from what is proposed here in that it would allow the Charity Commission or a court to add to the list of charitable purposes by including under "any other purpose … (ii) that may reasonably be regarded as analogous to, or within the spirit of, any purposes falling within any of paragraphs (a) to (l) … or (iii) that may reasonably be regarded as analogous to, or within the spirit of, any purposes which have been recognized, under the law relating to charities in England and Wales, as falling within sub-paragraph (ii) or this subparagraph" s 3(1)(m).

Crown enforcement whether they were traditional charitable purpose trusts, non-charitable purpose trusts or express trusts for persons.

One danger in trying to set out in statutory form all the types of trusts for which Crown enforcement is possible is that it might, perhaps through inadvertence, leave out many trusts that, as currently legally recognized charitable purpose trusts, have potential Crown enforcement. That problem might be addressed by grand-parenting in existing charitable purposes, but without a court power to add to these by reference to the preamble to the 1601 *Statute of Charitable Uses* or analogies thereto.[164] Instead the court would be assessing whether the Crown could enforce a particular express trust and would do so through interpretation of a statute identifying the types of express trusts that could be enforced by the Crown.

A possible advantage to the existing common law approach to charitable purpose trusts is that it allows the law to move forward by gradually adding to valid charitable purposes by way of analogy and analogy upon analogy. A statutory approach might be too slow to respond to current societal needs. The statutory approach to Crown enforcement suggested above could, however, include additions to potential Crown enforcement by regulation allowing for a relatively quick response to societal need without having to get the question on the legislative agenda. The additions would then be part of a political process and would not be subject to constraints of a common law approach wedded to the preamble of a 400 year old statute. While the statute might list various purposes that would typically tend to have a public quality justifying Crown enforcement, it would likely also have to have a more general category that would be subject to incremental modifications through

164. See *e.g.* English *Charities Act 2011*, *supra* note 66 which lists charitable purposes in s 3(1)(a) to (l) that, while capturing the main elements of the charitable purpose categories in *Pemsel*, *supra* note 14 and accompanying text, and while expanding on the scope of charitable purposes by adding several items, goes on to add, in s 3(1)(m), "any other purposes – (i) that are not within paragraphs (a) to (l) but are recognised as charitable purposes … under the old law".

court statutory interpretation.[165]

The statutory provisions identifying express trusts that could be enforced by the Crown would presumably have to address the question of public benefit since that is arguably the key to justifying Crown enforcement. While, for instance, a trust for the purpose of education normally would be inclined to have a public quality to it, it would presumably not be considered to have the requisite public quality for Crown enforcement if it was for exclusively for the education of the settlor's children. Consideration should, however, be given to providing some statutory guidance on the meaning of public benefit.

In providing statutory authority for Crown enforcement of specified purpose trusts one might also revisit the question of how to make Crown enforcement more effective. In doing so one might consider various models currently used for the enforcement of charitable purpose trusts such as the Ontario *Charities Accounting Act* or the English *Charities Act 2011*, discussed in Part II.C.1 above.[166]

VIII. Dispensing With the Distinctions Between Non-Charitable Purpose and Charitable Purpose Trusts (or Between Private Purpose Trusts and Statutorily Identified Public Purpose Trusts)

Once charitable and non-charitable purpose trusts are treated on the same basis with respect to Crown enforcement in the way described above, one might consider whether the distinction between these types of purpose trusts needs to be maintained. In addition to the issue of enforcement addressed by a Crown prerogative to enforce, charitable purpose trusts have a number of other distinct features such as the public benefit requirement, an exception with respect to the application of the rule against perpetuities, the exclusivity requirement, the political purposes doctrine and a qualification of the requirement of certainty of

165. The listed purposes would also be subject to statutory interpretation, therefore, leaving some room for court modification over time.

166. See notes 60-75 and the accompanying text.

objects. This part addresses these distinct features of charitable purpose trusts arguing that once a common approach to enforcement of express trusts is provided for and if the historical concept of charitable purpose trusts is replaced with statutorily identified public benefit express trusts that can be enforced by the Crown, several distinct features of current charitable purpose trusts with respect to exclusivity, public benefit, the invalidity of political purpose trusts and possibly also perpetuity, need no longer be maintained. This part also notes implications for the *Income Tax Act* approach to registered charities if the traditional approach to charitable purpose trusts is replaced in the way recommended herein.

A. Exclusivity

The exclusivity requirement for charitable purpose trusts should be dispensed with even if the current approach to charitable versus non-charitable purpose trusts remains. For similar reasons, the exclusivity requirement should be dispensed with if non-charitable purpose trusts are treated as valid trusts that impose trust obligations (not mere powers) on trustees. Also for similar reasons, the requirement should be dispensed with if the concept of charitable purpose trusts is replaced with broader statutorily identified public purpose trusts.

1. Dispensing with the Exclusivity Requirement

One might approach the exclusivity requirement for charitable purpose trusts from the standpoint of encouraging the creation of charitable purpose trusts on the assumption that it is desirable to encourage the creation of such trusts. In other words, one might ask what the effect would be on the creation of charitable purpose trusts if a settlor was allowed to mix charitable and non-charitable purposes and give the trustee an unfettered discretion to choose the extent of dispositions for either type of purpose. This is an empirical question that would be difficult to assess. However, there would be three possible outcomes. Dispensing with the exclusivity requirement could reduce the creation of charitable trusts, it could increase them or it could have no significant effect.

i. Reduction Unlikely

Is dispensing with the exclusivity requirement likely to reduce the creation of charitable purpose trusts? The only basis for such a reduction, it is submitted, is that if removal of the exclusivity requirement made enforcement of charitable purpose trusts by the Crown less effective, then settlors, knowing the prospect of enforcement is reduced, might be less inclined to create charitable purpose trusts. Crown enforcement of charitable purpose trusts in Canada appears, however, to be virtually non-existent, so removal of an exclusivity requirement is likely to have little or no effect on the prospect of enforcement. It is unlikely that settlors, even if they knew about how charitable purpose trusts are enforced, would become less inclined to create charitable purpose trusts.

Perhaps more robust Crown enforcement, together with an exclusivity requirement, would encourage greater allocations of property to charitable purpose trusts. The settlor, however, can achieve the effect of an exclusivity requirement simply by settling funds in trust on exclusively charitable purposes. The exclusivity requirement may, however, frustrate settlor intentions. Consider, for instance, the effect of the exclusivity rule under the current law with the distinction between charitable and non-charitable trusts and with an exclusivity requirement for charitable purpose trusts. A trust draftsperson might inform the settlor that a trust of mixed charitable and non-charitable purposes runs the risk that it will either be devoted exclusively to charitable purposes or held void with a resulting trust to the settlor (or her or his estate). The trust draftsperson might suggest creating two separate trusts, one for the charitable purposes and one for the non-charitable purposes. This, however, might not be entirely satisfactory for a settlor who wants to give the trustee the discretion to allocate funds between the settlor's intended non-charitable and charitable purposes. The settlor might instead indicate that the funds should be directed exclusively to the non-charitable purposes. The trust draftsperson should then advise the settlor that such a trust would be invalid, or, in the Canadian wait-and-see perpetuity legislation jurisdictions, that while such a gift would be valid, it would operate only as a power and only for a maximum of twenty-one years. The settlor might respond by directing all the funds to the charitable purposes even

though that is not quite what the settlor wants and in this way perhaps the exclusivity requirement could lead to more funds being directed to charitable purpose trusts. The settlor may, however, conclude that the simplest thing to do is to just give everything to her or his children or other persons or not give at all. Either way it frustrates the settlor's intention and it is hard to imagine that frustrating settlor intentions encourages more settling of property on charitable purpose trusts.

ii. Increase or No Significant Effect

If dispensing with the exclusivity requirement would not reduce the creation of charitable purpose trusts, it leaves only the other two outcomes. Either dispensing with the exclusivity requirement increases the creation of charitable purpose trusts, which many would likely consider desirable, or it would have no significant effect in which case it would make sense to dispense with the exclusivity requirement and avoid the problems it can create.[167] In the end, it is submitted, all an exclusivity requirement seems likely to do is frustrate settlors who would like to give trustees discretion to allocate funds between charitable and non-charitable purposes. Such frustration may cause some settlors to abandon some or all of their intended charitable and non-charitable purposes. Settlors who are unaware, or are not advised of, the pitfalls of mixed charitable and non-charitable purposes are likely to have their intentions defeated.

2. Dispensing With an Exclusivity Requirement in a Regime of Private vs. Statutorily Identified Public Purpose Trusts

If the charitable versus non-charitable purpose trust distinction was replaced by a distinction between statutorily identified public purpose trusts, as well as express trusts for persons, that could be enforced by the Crown and those that could not, the reason for an exclusivity rule would remain. In other words, the Crown arguably could not enforce public purpose trusts effectively if they were mixed with private purpose

167. On the problems the exclusivity requirement can create see Part II.C.2 above, notes 78-81 and the accompanying text.

trusts and the trustee had the discretion to allocate trust funds among the various purposes. The trustee might simply point to his discretion to justify the expenditure of trust funds on private purposes to the exclusion of public purposes. If the settlor articulated purposes that were at least in part public purposes, then arguably there is a public interest in enforcement and the trustee should not be able to so easily avoid the expenditure of trust funds on public purposes by arguing that he has discretion to allocate funds between either the public or private purposes.

The same arguments that apply to dispensing with the exclusivity requirement in the context of current charitable and non-charitable purpose trusts would, however, apply in the context of a regime that had enforceable private trusts for persons or purposes together with statutorily identified express trusts (whether for persons or purposes) enforceable by the Crown. Assuming it is considered desirable to encourage donations of funds on trusts that provide public benefits, what would be the effect of an exclusivity requirement on the creation of such trusts? It seems unlikely, for the reasons noted above with respect to charitable purpose trusts, that it would encourage the creation of such trusts. If it would discourage them, then presumably it makes sense not to have an exclusivity requirement. If the effect of an exclusivity requirement would be insignificant, then it would seem to make sense not to have such a requirement since it would invite the same difficulties currently created by the exclusivity requirement for charitable purpose trusts.

B. Public Benefit

While a court determination of public benefit, as discussed in Part VII above, would inevitably remain, it would be a public benefit assessment in the context of determining whether a particular express trust could be enforced by the Crown and this assessment would be the same whether the express trust was for persons or purposes. Whether the purpose was charitable or not would not be relevant in the context of interpreting the statute that identified trusts for which Crown enforcement was available.

C. Political Purposes

The question of whether a trust for political purposes should be subject

to potential enforcement by the Crown should, it is submitted, be a question taken up in the legislative and political process of deciding whether the Crown should be given a statutory power to enforce such trusts. The question of whether "political purpose" trusts should be valid or not could then be subject to public debate. There arguably are public benefits to trusts for political purposes in a democratic society. This question should be confronted in the drafting of a statute that identifies the types of trusts for which Crown enforcement is available.[168] Either the statute should permit Crown enforcement of such trusts or it should not. If the statute permitted Crown enforcement of political purpose trusts, the public benefit question that courts will inevitability have to assess should be addressed in the statute in a way that avoids having the court decide whether there is a benefit to advocating a change in the law. The statute might, for instance, deem there to be a benefit to political purpose trusts. The question for the court might then be limited to the question of whether there was a sufficient public quality to the trust in terms of the potential number of persons who might benefit.[169] The question being addressed by the statute and by a court interpreting the statute would not be whether the trust is valid or not but whether the trust could be enforced by the Crown.[170]

If it was decided through the political process, that political purpose trusts were to, by statutory enactment, be treated as valid trusts, it

168. In other words, the current approach with respect to charitable purpose trusts mixes two questions together – one being the availability of Crown enforcement (currently tied up in the question of whether the purpose is charitable and provides a public benefit) and the question of whether trusts for political purposes should be allowed (that has become linked to the public benefit question).

169. A political purpose trust that advocated for a change to a law that, for example, benefited only the settlor or the settlor's close relatives would arguably not have the requisite public quality to justify Crown enforcement.

170. The validity of the trust would be a separate question focusing on, for instance, the capacity of the settlor to create the trust, the three certainties, the constitution of the trust, its compliance with any necessary formalities such as under wills legislation or statute of frauds legislation, and whether the purported trust was illegal or contrary to public policy.

would be important to address the mechanism for Crown enforcement. Enforcement through an Attorney General who is a member of cabinet would arguably be less than ideal since the Attorney General may be reluctant to enforce political purpose trusts that were against the interests of the government. It may, therefore, be important to create a Crown enforcement body that operates reasonably independently from government.

D. Perpetuity Exception

1. Abrogation of the Rule Against Perpetuities

Another distinction between charitable purpose trusts and non-charitable purpose trusts is that charitable purpose trusts, subject to limited exceptions, can be of perpetual duration. Obsolete uses of trust property can be dealt with through court *cy-près* orders. There is a move to abrogation of rules against perpetuities with Manitoba, Saskatchewan and Nova Scotia having all abrogated them in recent years.[171] The Uniform Law Conference of Canada has recommended similar abrogation in other Canadian common law jurisdictions.[172] If rules against perpetuities are abrogated, it would remove the perpetuity distinction between non-charitable purpose trusts and charitable purpose trusts. If non-charitable purpose trusts were treated as valid trusts and the rule against perpetuities were abrogated it would leave the potential for obsolete non-charitable purpose trusts and some means to vary such trusts would be needed. *Cy-près* might be extended to non-charitable purpose trusts, but it might be better to reassess the variation of express trusts more generally addressing not only the question of varying non-charitable purpose trusts, but also addressing issues in varying charitable purpose trusts and express trusts

171. For Manitoba see the *Manitoba Perpetuities and Accumulations Act*, *supra* note 156, ss 2-3; for Saskatchewan see the *Saskatchewan Trustee Act*, *supra* note 156, s 58; and for Nova Scotia see the *Nova Scotia Perpetuities Act*, *supra* note 156, ss 2-3.

172. Currently expressed in *Uniform Trustee Act*, *supra* note 2, s 88.

for persons.[173]

The same considerations would apply if the charitable versus non-charitable purpose trust distinction was replaced with a distinction between private purpose trusts and statutorily identified public purpose trusts and purpose trusts were generally made enforceable. Removal of the rule against perpetuities would make distinguishing between such public and private purpose trusts on the basis of a rule against perpetuities no longer necessary. There would, however, be a need for a means to vary such trusts if, over time, it became impossible or impracticable to execute them.

2. Retention of the Rule Against Perpetuities

If the rule against perpetuities or some alternative rule of similar purpose was to be retained, one would need to consider whether a distinction should be made in the application of the rule to valid private purpose trusts and statutorily identified public purpose trusts. If private purpose trusts were treated as valid trusts or were even just enforced as powers, concerns about dead hand control of funds through private purpose trusts (or powers) might remain and some means of bringing an end to such trusts (or powers) might be considered necessary.

E. Certainty of Objects for Charitable Purpose Trusts

Certainty of objects is one of the three certainties required for the creation of an express trust. For an express trust for persons the objects are the persons who are to be the beneficiaries of the trust and the certainty required is certainty of beneficiaries. For a purpose trust the objects are the purposes of the trust and the certainty required is certainty of purposes. A qualification with respect to certainty of purposes is made for charitable purpose trusts. The certainty need only be that the settlor intended the trust property to be held for charitable purposes. If, however, the particular charitable purposes are not clearly expressed, the

173. Nova Scotia has taken a different approach to variation of trusts than other jurisdictions. See the Nova Scotia *Variation of Trusts Act*, RSNS 1989, c 486, as amended by SNS 2011, c 42, s 6.

court can provide a scheme giving the trustee some clarity as to how the trust should be administered. In a testamentary trust, subject to concerns for dependent relief, it arguably makes sense to treat a clearly expressed charitable intent, even though the charitable purposes themselves are vague, as a valid charitable purpose trust to avoid the return of the trust property to the estate and, therefore, to residuary legatees or intestate heirs (giving them a windfall gain) contrary to the clearly expressed charitable intent of the testator and taking into account the public benefit quality of charitable purposes.

It seems desirable to retain this particular aspect of the exceptional treatment of charitable purpose trusts. One might, however, ask whether this same treatment should be extended to other purpose trusts. If, for instance, a testator put aside property to be held on trust for "benevolent or philanthropic purposes" without further elaboration, would it not also be reasonable to allow the trust to take effect with the court setting out a scheme for the "benevolent or philanthropic" purposes rather than having the property revert to the estate to be distributed to the residuary legatees or intestate heirs (subject, perhaps, to concerns for the needs of dependents). Returning the property to the estate to be distributed to residuary legatees or intestate heirs would be contrary to the clearly expressed benevolent or philanthropic intentions of the testator and the property intended by the settlor to be used for "benevolent or philanthropic" purposes could, in accordance with those words, be directed to purposes that could provide a public benefit (unless some other language in the will constrained the scope of persons whom the testator intended to benefit). Such an approach, in addition to providing a public benefit, would arguably be more consistent with the settlor's intent.

If property is settled on trust for vaguely expressed private purposes, the question would be more difficult. One would be faced with the issue of either finding the trust invalid and having the property revert to the estate, and therefore to persons the settlor apparently did not intend the property to go to, or trying to provide the trustee with some direction as to the purposes to which the property should be directed knowing that whatever purpose the property was directed to, given the settlor's vague

expression of purposes, probably also did not fit the intent of the settlor. There would likely be some purpose trusts that would be close to the line in terms of falling within the statutorily identified public purposes and for these a possible resolution might be to conclude they do fall within the statutorily identified public purposes and therefore apply the approach suggested above and provide a public purpose scheme for the use of the property. Where the property is clearly directed to purposes outside the statutorily identified public purposes, it would appear to leave a choice between two outcomes neither of which satisfies the intent of the settlor (whatever that might have been). A rule that returns the property to the settlor might give the settlor, hopefully assisted by a knowledgeable draftsperson, an incentive to spell out the purposes in sufficient detail, such that the trustee, and a court, has a clear enough indication of the settlor's intent. Perhaps an alternative would be to give the intended trustee a power to use or appoint the property for purposes or persons (other than the trustee) for a statutorily set period of time.

F. Effect on the *Income Tax Act* Approach to Registered Charities

If Canadian provincial legislative provisions were to do away with the distinction between charitable and non-charitable purpose trusts, how would it affect the tax subsidy provisions under the *Income Tax Act*? In interpreting the meaning of "charitable purposes" and "charitable activities" courts in Canada have made reference to the meaning of "charitable purposes" in the common law of trusts. Courts interpreting those *Income Tax Act* provisions could continue to do so even though the continued development of that expression would no longer be developed by Canadian common law in the context of trust law. The development of the meaning of that expression has, for some years now, largely been done by courts interpreting provisions in the *Income Tax Act*. Perhaps Parliament might then be inclined to give a renewed focus to identifying the types of purposes and activities for which donation tax credits should be provided. One might lament the absence of the contribution trust law court decisions might make to the meaning of charitable purposes under the *Income Tax Act* if the traditional approach to charitable purpose trusts

was no longer retained. The loss of such contributions might not, however, be that significant given the relative paucity of such cases. Further, one might question the benefit to be obtained from trust law decisions on the meaning of charitable purposes where the policy considerations are quite different from those at play in the context of providing an income tax credit that can lead to a potentially significant reduction in government tax revenues.

IX. Conclusion

Trust enforcement problems are not limited to problems arising from the absence of named beneficiaries, or a clearly described class of beneficiaries, in non-charitable purpose trusts. Enforcement of trusts will rarely, if ever, be perfect and the effectiveness of enforcement for any given trust will always be question of degree. While charitable purpose trusts are valid in spite of the absence of named beneficiaries or a clearly described class of beneficiaries on the basis of enforcement by the Crown as *parens patriae*, there will be gaps in enforcement, particularly in Canada where, with some qualification in the province of Ontario, there is no scheme for enforcement such as the English Charity Commission. Enforcement of express trusts for persons can also be weak, particularly in the context of discretionary trusts with a large number of beneficiaries.

While charitable purpose trusts are sometimes described as "public trusts," express trusts for persons and non-charitable purpose trusts can provide benefits for a broad community extending well beyond the settlor's family and personal friends. Express trusts for persons and non-charitable purpose trusts may well, in fact, provide benefits to a much broader range of persons than many charitable purpose trusts. Facilitating enforcement of all types of express trusts would, therefore, likely be conducive to encouraging philanthropy and ensuring the philanthropic use of funds directed to philanthropic purposes.

With provinces perhaps poised to enact the *Uniform Trustee Act* proposed by the Uniform Law Conference of Canada Law, it seems an appropriate time to consider extending aspects of enforcement for non-charitable purpose trusts, charitable purpose trusts and express trusts for persons. Enforcement could be facilitated by providing for the

enforcement of non-charitable purpose trusts by appointed enforcers that might include the settlor or the settlor's personal representatives. That same approach to enforcement could also allow for stronger enforcement of express trusts for persons, particularly where those express trusts for persons are discretionary trusts with a broad range of beneficiaries. Enforcement of all types of express trusts might also be expanded by allowing the court to grant standing to persons (other than named beneficiaries or the Crown) to enforce express trusts, perhaps with the settlor being allowed to opt in or out of granting the court such a power.

With a workable means of enforcement, non-charitable purpose trusts could be treated as legally valid so that the trust purposes become trustee duties and not just powers in the manner provided for in wait-and-see perpetuity legislation. Once non-charitable purpose trusts are treated as legally valid trusts there would arguably be no further need, at least from the standpoint of legal validity, to distinguish between charitable and non-charitable purpose trusts. While the distinction might be retained with respect to the question of whether the Crown (through the Attorney General or other nominee such as a Public Trustee) has standing to enforce the trust, this might be treated as a separate question with legislation describing the types of purpose trusts with respect to which the Crown would be given standing to enforce thereby making the question of Crown enforcement more directly a part of the political process.

If a statutory approach to Crown enforcement is taken, the question of public benefit can become part of the political process with the public benefit then presumed if the particular purpose falls within the statutorily identified purposes for which Crown enforcement is permitted. If public benefit was presumed for the statutory identified trusts for which Crown enforcement is permitted, the alleged problem that led to the invalidity of political purpose trusts is arguably no longer present. The validity of political purpose trusts might then be addressed as a question for the legislature and therefore more directly a part of the political process. If, as a result of such a political process, it is decided that political purpose trusts are to be treated as valid trusts, legislation might address as a separate question whether such trusts should be enforceable by the Crown. *Cy-*

près, or some modified version of it, could be extended to non-charitable purpose trusts made valid by providing for their enforcement as trusts. One might also extend the provision of administrative schemes for uncertain purposes beyond charitable purposes to the purposes statutorily deemed to be public purposes.

The exclusivity requirement should be dispensed with regardless of whether any other changes are made concerning the validity of non-charitable purpose trusts or expanding the range of means of enforcement for express trusts generally. The exclusivity requirement is not likely to have the effect of increasing the creation of charitable (or statutorily identified public) purpose trusts. It either discourages the creation of such trusts which would favour dispensing with it or it has no significant effect on the creation of such trusts in which case dispensing with it would avoid problems associated with it.

The Vulnerability Jurisdiction: Equity, *Parens Patriae*, and the Inherent Jurisdiction of the Court

Margaret Hall*

This paper describes how the English courts, in the "heroic act of judicial invention", have developed a distinct vulnerability jurisdiction, separate and apart from the ancient jurisdiction of parens patriae, through the exercise of the inherent jurisdiction of the court. This new jurisdiction provides a legal basis and mechanism for the disruption of exploitative relationship contexts. The objective of that disruption is not protection per se (the parens patriae objective), but the safeguarding of individual autonomy rights in situations where those rights cannot be effectively exercised without intervention. The paper concludes with a discussions of implications of the English "invention" in Canadian jurisprudence.

* Margaret Isabel Hall, LLB, LLM is an Associate Professor in the Faculty of Law, Thompson Rivers University, British Columbia, Canada. Margaret has published extensively in the area of law and aging, with a particular focus on mental capacity, legal responses to exploitation, critical law and aging, and vulnerability theory.

I. Introduction

This paper describes how the English courts, in a "heroic act of judicial invention",[1] have developed a distinct *vulnerability jurisdiction*, separate and apart from the ancient jurisdiction of *parens patriae*, through the exercise of the inherent jurisdiction of the court. This new jurisdiction provides a legal basis and mechanism for the disruption of exploitative relationship contexts. The objective of that disruption is not protection *per se* (the *parens patriae* objective), but the safeguarding of individual autonomy rights in situations where those rights cannot be effectively exercised *without* intervention. In doing so, the court is responding to relationship vulnerability, a particular quality of vulnerability that is not dependent on or derived from personal characteristics such as age, gender, or mental disability, although the relationship between these factors (together with others such as economic status) may intensify relationship

1. Sir James Munby, "Protecting the Rights of Vulnerable and Incapacitous Adults – The Role of the Courts: An Example of Judicial Law-making" (2014) 26 Child & Family Law Quarterly 64 at 77 [Munby].

vulnerability so as to justify intervention in a particular case.[2] This response is founded on the understanding that legal/public intervention is not the only possible source of autonomy restriction. In this way, the vulnerability jurisdiction is conceptually rooted in the doctrine of equitable fraud (in particular, the doctrine of undue influence), and the new jurisdiction is most coherently understood as an extension of the equitable doctrine (rather than the resurgence of a new *parens patriae*) in situations outside of the contractual/testamentary context and at the instigation of third parties (public or private).

The process of "judicial invention" through which the jurisdiction has developed has been lengthy and non-linear, generating confusion about its source and nature. In terms of both language and origin (the decision in the case of *In Re F*[3] ("*In Re F*"), a response to the disappearance of *parens patriae* with regard to mentally incapable adults in England), the new jurisdiction has been tangled up with the old to the extent that it has been described (mistakenly) as a rebirth and extension of *parens patriae*, "the invention … by the family judges of a full-blown welfare-based *parens patriae* jurisdiction in relation to incapacitated adults" *and* to other "vulnerable persons".[4]

One source of confusion has been the nebulous and ill-defined nature of the "inherent jurisdiction of the court" as distinct from *parens patriae* (which is occasionally described as "the inherent jurisdiction"). The distinction between the two is explained below. The language of "vulnerability", as used in the law generally and the cases discussed here in particular, is a further source of confusion. The new jurisdiction and *parens patriae* each enable public response to private vulnerability; vulnerability is not one idea, but several. Understanding the distinctions between these ideas is essential to understanding the nature of the new jurisdiction and how it differs from *parens patriae* in purpose and effect.

---

2. See MI Hall "Equity Theory: Responding to the Material Exploitation of the Vulnerable but Capable" in Israel Doron, ed, *Theories on Law and Ageing: The Jurisprudence of Elder Law* (Berlin: Springer Publications, 2008) at 107.
3. [1990] 2 AC 1 (HL) [*In Re F*].
4. Munby, *supra* note 1 at 77.

Parens patriae describes the state's responsibility to protect the members of identified "vulnerable populations": persons who are deemed incapable of protecting their own interests by reason of their particular personal characteristics. Children and mentally incapable adults are "vulnerable populations" of this kind and are, on this basis, the subjects of both *parens patriae* and specific legislation such as the *Mental Capacity Act 2005*[5] ("*Mental Capacity Act*") discussed below. The new vulnerability jurisdiction described in this paper responds to a more universal, mutable vulnerability that waxes and wanes in connection with personal and other contextual circumstances, including the "quality and quantity of resources we possess or can command".[6] The distinction is significant, providing a coherent theoretical and principled basis for the new jurisdiction *and* delimiting the kind and scope of intervention it justifies; not a capacious or amorphous power of protection (the *parens* paradigm), but a more specific intervention for the purpose of creating a space in which autonomy can be developed and exercised.

The first part of this paper describes in more detail the origin and nature of the *parens patriae* jurisdiction and the inherent jurisdiction of the court, respectively, together with a discussion of the distinctions between them. The second part of this paper describes the cases through which the new jurisdiction was developed prior to the *Mental Capacity Act* which filled the gap left by the removal of *parens patriae* with respect to

5. (UK), c 9.
6. Martha Albertson Fineman, "The Vulnerable Subject: Anchoring Equality in the Human Condition" (2008) 20 Yale Journal of Law and Feminism 1 (while "undeniably universal, human vulnerability is also particular: it is experienced uniquely by each of us and this experience is greatly influenced by the "quality and quantity of resources we possess or can command" at 8) and ("[v]ulnerability initially should be understood as arising from our embodiment, which carries with it the ever-present possibility of harm, injury, and misfortune from mildly adverse to catastrophically devastating events, whether accidental, intentional, or otherwise. Individuals can attempt to lessen the risk or mitigate the impact of such events, but they cannot eliminate their possibility. Understanding vulnerability begins with the realization that many such events are ultimately beyond human control" at 8).

mentally incompetent adults in England. A declaration of lawfulness on the basis of the common law doctrine of necessity is used by the English courts in these cases to establish best interests and to justify interventions for the purpose of protecting the rights and interests of incapable adults. The third part of this paper describes the development of a distinct vulnerability jurisdiction (*i.e.* not a replacement for *parens patriate*) through a series of cases decided after the *Mental Capacity Act* (which filled the *parens patriae* gap *vis a vis* incapable adults), culminating in *DL v A Local Authority*.[7] This new jurisdiction provides a basis on which the courts can respond to vulnerability arising through relationship contexts (in respect of which no legislation applies), as opposed to the "protection-needs" of vulnerable populations. The declaration of lawfulness in these later cases is rooted in the equitable doctrine of undue influence as opposed to the common law doctrine of necessity. This distinction is important and marks a decisive conceptual break from the earlier cases. The final part of this paper considers the implications of the English "invention" in the Canadian legal context.

II. *Parens Pariae* and the Inherent Jurisdiction of the Court

The *parens patriae* jurisdiction of superior courts, while occasionally referred to as an inherent jurisdiction, is very different in source and purpose from the inherent jurisdiction of said superior courts. The *parens patriae* jurisdiction originated in the personal authority and responsibility of the King. The inherent jurisdiction of the court, in contrast, has been described as the "essence" of the superior court — its "immanent attribute" and "very life-blood"; a "peculiar concept … so amorphous and ubiquitous and so pervasive in its operation that it seems to defy

7. [2012] EWCA Civ 253 [*Local Authority*].

the challenge to determine its quality and establish its limits".[8] Thomas Cromwell, writing extra-judicially, referred to the "inherent jurisdiction" as an "original jurisdiction in any matter unless jurisdiction is clearly taken away by statute",[9] inherited by the Canadian superior courts of general jurisdiction as the descendants of the Royal Courts of Justice.

A. *Parens Patriae*

Parens patriae refers to the state's authority and responsibility to protect the best interests of vulnerable persons (defined, for this purpose, as members of vulnerable populations). The source of the *parens patriae* jurisdiction was described by the Supreme Court of Canada in *Re Eve*[10] as "lost in the mists of antiquity", although the most probable theory was that Edward I had assumed the authority from the feudal lords "who would naturally take possession of the land of a tenant unable to perform his feudal duties".[11] Such persons were known as "lunatics" (persons who had become mentally disordered and so lost mental capacity) or "fools"

8. Jacob, below note 20; see also Keith Mason, "The Inherent Jurisdiction of the Court" (1983) 57 Australian Law Journal 449; MS Dockray, "The Inherent Jurisdiction to Regulate Civil Proceedings" (1997) 113 Law Quarterly Review 120 ("[t]here is no clear agreement on what [the inherent jurisdiction] is, where it came from, which courts and tribunals have it and what it can be used for" at 120) [Dockray]; see also *Shreem Holdings Inc v Barr Picard*, 2014 ABQB 112 ("[j]ust as the existence of the inherent jurisdiction of superior courts is indisputable and certain, the theoretical basis and scope of it are debatable" at para 26) [*Shreem Holdings*].
9. TA Cromwell, "Aspects of Constitutional Judicial Review in Canada" (1994-1995) 46 South Carolina Law Review 1031; see *e.g. Dominion Canners Ltd v Costanza*, [1923] SCR 46 at para 61; *In re Sproule*, [1886] 12 SCR 140.
10. [1986] 2 SCR 308 [*Re Eve*].
11. *Ibid* at para 32, the case involved a mother's application requesting that the court consent to the sterlisation of her mentally incapable daughter as an exercise of its *parens patriae* jurisdiction. The mental health legislation in the province at the time did not provide for substitute consent to the procedure. The Court declined to exercise the jurisdiction on the basis that it had not been established that it was in Eve's best interests.

(persons who had never had mental capacity).[12] As described in *Re Eve*, the court's "wardship" jurisdiction in relation to children became merged or "assimilated" over time with this jurisdiction over "mental incompetents" to comprise the *parens patriae* jurisdiction in respect of both vulnerable population groups.[13] The jurisdiction "continues to this day" (so long as it has not been supplanted by legislation) and remains applicable in specific situations not contemplated by legislation.[14] So long as the *parens patriae* jurisdiction is exercised in the best interests of the individual:

> the situations under which it can be exercised are legion … and the categories under which the jurisdiction can be exercised are never closed … the jurisdiction is a carefully guarded one. The courts will not readily assume that it [the *parens patriae* jurisdiction] has been removed by legislation where a necessity arises to protect a person who cannot protect himself. … Simply put, the discretion is to do what is necessary for the protection of the person for whose benefit it is exercised.[15]

Sir James Munby, writing about *parens patriae* in connection with the development of the new jurisdiction in England, located the origins of *parens patriae* in the prerogative powers of the medieval kings to take "responsibility for those without the capacity to look after themselves".[16] In Munby's account, the ancient power was "put on a statutory footing" with the creation of the Court of Wards and Liveries in 1540 which had jurisdiction over both children and the mentally incapable.[17] That court was abolished in 1646 and following the Restoration the jurisdictions

12. Munby, *supra* note 1 at 66.
13. *Re Eve*, *supra* note 10 at paras 40, 42 ("Lord Somers resembled the jurisdiction over infants, to the care which the Court takes with respect to lunatics, and supposed that the jurisdiction devolved on the Crown, in the same way" at para 42).
14. See *Beson v Director of Child Welfare* (*Nfld*), [1982] 2 SCR 716; *Re Eve*, *supra* note 10 ("[e]ven where there is legislation in the area, the courts will continue to use the *parens patriae* jurisdiction to deal with uncontemplated situations where it appears necessary to do so for the protection of those who fall within its ambit" at para 42).
15. *Re Eve*, *supra* note 10 at paras 74-75 (the jurisdiction was therefore "founded on necessity, namely the need to act for the protection of those who cannot care for themselves" at para 75).
16. Munby, *supra* note 1 at 66.
17. *Ibid*.

were separated: the *parens patriae* jurisdiction with relation to "infants" returned to the Chancery; the Crown's *parens patriae* power with relation to persons of "unsound mind" was assigned to specific individuals (initially to the Lord Chancellor). In 1956, the power was assigned by warrant to the Lord Chancellor and to the judges of the Court of Chancery. That warrant was revoked in England in 1960, on the coming into force of the *Mental Health Act 1959*,[18] effectively removing the *parens patriae* jurisdiction in respect of incompetent persons in England and creating the "gap" at issue in *In Re F* (a gap that was subsequently filled by legislation). In contrast, the *parens patriae* jurisdiction with respect to incapable adults, as discussed in *Re Eve*, was not removed in this way (or "swept away" in the language of the House of Lords in *In Re F*) in Canada.

B. The Inherent Jurisdiction of the Court

"Just as the existence of the inherent jurisdiction of superior courts is indisputable and certain, the theoretical basis and scope of it are debatable".[19]

The inherent jurisdiction was described by Master IH Jacob as "a residual source of powers, which the court may draw upon as necessary whenever it is just or equitable to do so, and in particular to ensure the observance of due process of law, to prevent improper vexation or oppression, to do justice between the parties and to secure a fair trial

18. (UK), 7 & 8 Eliz II, c 72.
19. *Shreem Holdings*, *supra* note 8 at para 26; see also Dockray, *supra* note 8 ("[t]here is no clear agreement on what it [the inherent jurisdiction] is, where it came from, which courts and tribunals have it and what it can be used for" at 20).

between them".[20] These powers are derived

> not from any statute or rule of law, but from the very nature of the court as a superior court of law … This description has been criticised as being 'metaphysical' … but I think nevertheless that it is apt to describe the quality of this jurisdiction. For the essential character of a superior court of law necessarily involves that it should be invested with a power to maintain its authority and to prevent its process being obstructed and abused. Such a power is intrinsic in a superior court; it is its very life-blood, its very essence, its immanent attribute. Without such a power, the court would have form but would lack substance. The jurisdiction which is inherent in a superior court of law is that which enables it to fulfil itself as a court of law.[21]

The Supreme Court of Canada, drawing on Master Jacob's definition, has described the inherent jurisdiction as both "inexhaustibly" various and as a "narrow core" of powers. Justice Binnie, quoting Master Jacob, observed in *R v Caron*[22] ("*Caron*") that the inherent jurisdiction may be invoked in "an apparently inexhaustible variety of circumstances and may be exercised in different ways … even in respect of matters which

20. IH Jacob, "The Inherent Jurisdiction of the Court" (1970) 23 Current Legal Problems 23 at 51 [*Jacob*] cited in *R v Caron*, 2011 SCC 5 at para 24 [*Caron*] (Jacob's article has been cited on nine separate occasions by the Supreme Court of Canada: *Société des Acadiens du Nouveau-Brunswick Inc v Association of Parents for Fairness in Education, Grand Falls District 50 Branch*), [1986] 1 SCR 549 at para 95 per Justice Wilson (granting leave to appeal to a non-party); *BCGEU v British Columbia* (*Attorney General*), [1988] 2 SCR 214 at para 49 (issuing injunction on the court's own motion to guarantee access to court facilities); *R v Morales*, [1992] 3 SCR 711 at para 87 per Justice Gonthier (discretion regarding bail); *R v Hinse*, [1995] 4 SCR 597 at para 21 per Chief Justice Lamer (stay of criminal proceedings for abuse of process); *MacMillan Bloedel v Simpson*, [1995] 4 SCR 725 at paras 29-31 per Lamer CJ (punishing for contempt out of court) [*MacMillan*]; *R v Rose*, [1998] 3 SCR 262 at paras 64, 131 per Justice L'Heureux-Dubé and Justices Cory, Iacobucci and Bastarache, respectively (discretion to grant a right of reply in a criminal trial); *R v Cunningham*, [2010] 1 SCR 331 at para 18 (authority to refuse defence counsel's request to withdraw); *Ontario v Criminal Lawyers Association of Ontario*, 2013 SCC 43 ("[t]he inherent jurisdiction of the court is limited by institutional roles and capacities" at para 24) [*Criminal Lawyers*]).
21. Jacob, *supra* note 20 at 27 cited in *MacMillan*, *ibid* at para 20.
22. *Caron*, *supra* note 20.

are regulated by statute or by rule of court, so long as it [the jurisdiction] can do so without contravening any statutory provision".[23] A "categories approach" to the inherent jurisdiction, he concluded, was therefore inappropriate although this "very plenitude" required that the "inherent jurisdiction be exercised sparingly and with caution".[24]

Chief Justice Lamer, in *MacMillan Bloedel Ltd v Simpson*[25] ("*MacMillan*") described a "core" or "inherent" jurisdiction that is beyond the reach of Parliament and the provincial legislatures in the absence of constitutional amendment".[26] This "core", he concluded, was made up of the court's "essential and immanent attributes" (quoting Master Jacob) and therefore to "[r]emove any part of [it] emasculates the court, making it something other than a superior court".[27] The content of that "core" is described in the case as comprising "those powers which are essential to the administration of justice and the maintenance of the rule of law".[28] Justice Karakatsanis, giving the reasons for the majority in *Ontario v Criminal Lawyers Association of Ontario*,[29] described the "core" as "a very narrow one which includes only critically important jurisdictions which are essential to the existence of a superior court of inherent jurisdiction and to the preservation of its foundational role within our legal system".[30]

---

23. Jacob, *supra* note 20 at 23-24 cited in *Caron*, *ibid* at paras 29, 32.
24. *Caron*, *ibid* at para 30.
25. *MacMillan*, *supra* note 20 (giving reasons for the majority).
26. *Ibid* at para 8 (referring to the reasons given by Chief Justice McEachern in the British Columbia Supreme Court).
27. *Ibid* ("[t]he full range of powers which comprise the inherent jurisdiction of a superior court are, together, its 'essential character' or 'immanent attribute'. To remove any part of this core emasculates the court, making it something other than a superior court" at para 30).
28. *Ibid* ("[i]t is unnecessary in this case to enumerate the precise powers which compose inherent jurisdiction, as the power to punish for contempt *ex facie* [at issue in the *Simpson* case] is obviously within that jurisdiction. The power to punish for all forms of contempt is one of the defining features of superior courts" at para 38).
29. *Criminal Lawyers*, *supra* note 20.
30. *Ibid* at para 19, per Justice Karakatsanis (giving the reasons for the majority, quoting from *Reference re Amendments to the Residential Tenancies Act (NS)*, [1996] 1 SCR 186 at para 56 per Lamer CJ).

At once "inexhaustibly various" and "very narrow", William H Charles concluded that "one might have thought" that "to attempt a definition … of such a mysterious and unruly concept … would involve a mission impossible".[31] Rosara Joseph has suggested that "the inherent jurisdiction of the High Court was better understood as being comprised of a number of separate jurisdictions, which have developed piecemeal and mostly in isolation" (rather than an "amorphous single source of jurisdiction").[32] These "jurisdictions" are identified by Joseph as including "*parens patriae*, punishment for contempt of court, judicial review, bail and jurisdiction over officers of the Court" and the "more shadowy category of inherent jurisdiction: the Court's jurisdiction to revisit its own 'null' decisions".[33]

The apparently contradictory and "unruly" nature of the inherent jurisdiction may be resolved if the jurisdiction is understood as a single source from which two different *kinds* of powers may flow, powers which may subsequently be developed through the case law to comprise distinct jurisdictions[34] including those described by Joseph (with the exception of *parens patriae*, the origins of which lie in a very different source as described above). The first kind of powers flowing from the inherent jurisdiction source are those powers "essential to the administration of justice and the maintenance of the rule of law" as described in *MacMillan*.[35] These "core" powers, pertaining to the self-governance functions of the court, are essential to the court's identity and as such are constitutionally protected. The inherent jurisdiction, as a residual source of power, may *also* be drawn upon "as necessary whenever it is just or equitable to do so" (to quote Master Jacob) to generate *non*-core exercises of power "as

31. William H Charles, "Inherent Jurisdiction and Its Application in Nova Scotia Courts: Metaphysical, Historical or Pragmatic?" (2010) 33 Dalhousie Law Journal 63 at 66.
32. Rosara Joseph, "Inherent Jurisdiction and Inherent Powers in New Zealand" (2005) 11 Canterbury Law Review at 220.
33. *Ibid*.
34. This approach is consistent with Master Jacob's description of the inherent jurisdiction as a "residual source of powers" and with Justice Binnie's rejection of the kind of "categories-approach" suggested by Joseph.
35. *MacMillan*, *supra* note 20 at para 38.

and when the need arises".[36] These non-core powers may subsequently be reduced or even replaced by legislation without "emasculating" the court; in such a case the actions of the legislature will remove the "necessity" for the inherent jurisdiction to be used in a particular way. It is therefore right for the power (derived from the inherent jurisdiction source) to recede. Through the development of this second kind of power, the inherent jurisdiction acts as a "great safety net" to be expanded or retracted in connection with legislation and through which the superior court exercises its "metaphysical" function: to prevent "improper vexation or oppression" and to effect justice and equity between the parties.

The first set of cases discussed below show the courts drawing on the "source" of the inherent jurisdiction to apply the common law doctrine of necessity in a series of increasingly "new" circumstances using the declaration as a means of doing so. The second set of cases discussed below show the courts developing that power further as a means of implementing the principles of equitable fraud outside of the traditional transactional and testamentary contexts. Crucial in both sets of cases is the existence of a justice gap caused by the absence of legislation; as demonstrated in these cases, the inherent jurisdiction provides the authority and the means (the declaration) on the basis of which such gaps *must* be filled by the courts, drawing on the principles of equity and the common law to do so. The power drawn from the inherent jurisdiction must retreat where the gap has been filled by legislation, as in the kinds of circumstances at issue in the necessity cases, discussed below.

II. The Cases: *In Re F* and After

A. *In Re F*

In the case of *In Re F*, the House of Lords interpreted and applied the common law doctrine of necessity to fill a gap left by the disappearance of *parens patriae* in relation to mentally incapable adults. The court drew on the inherent jurisdiction of the court, as a residual source of power, to enable the principled exercise of that doctrine through the mechanism of

36. Dockray, *supra* note 8 at 124.

the declaration.

In Re F concerned a "momentous and irrevocable"[37] medical decision with significant public policy implications: whether a sterilisation procedure could be lawfully performed on a mentally disabled woman, "F". F was unable to consent to the procedure by reason of her disability, nor could anyone else consent on F's behalf (the applicable mental health legislation did not provide for substitute consent to the procedure).[38] The medial professionals involved in F's care, together with F's mother, agreed that pregnancy and birth would be a "disaster" and "catastrophic" for F; and F (through her mother) asked the court either to provide consent to the procedure or to make a declaration that the procedure could be lawfully performed without consent.[39] The *parens patriae* jurisdiction of the court applying to mentally incompetent adults (which would have provided a basis on which the court could consent to the operation) had been "swept away"[40] by legislative reform some years before, leaving no basis on which the court could give consent to the procedure.

The House of Lords, like the Court of Appeal before it,[41] found that the procedure proposed (being in the best interests of F) was justified on the basis of the common law doctrine of *necessity*, which essentially provided a policy-based rationale for dispensing with the requirement of consent where an interference that would otherwise comprise a trespass (to person or property) is "justified *summa necessitate*, by the

37. See *Re S* (*Hospital Patient: Court's Jurisdiction*), [1996] Fam 1 (CA (Civ) (Eng)) at para 4.
38. *In Re F*, *supra* note 3 at 22 citing *Collins v Wilcock*, [1984] 1 WLR 1172 (QB (Eng))(without consent, medical treatment (like any other interference with the body of another) would constitute a trespass to the person: the tort of battery).
39. *In Re F, supra* note 3 at 2-3.
40. *In Re B* (*A Minor*)*(Wardship: Sterilisation*), [1988] AC 199 (HL)(Lord Brandon noted that no difficulty would arise regarding the Court's current jurisdiction to consent to the procedure if F were a minor suffering from a comparable mental disability, in which case the court would exercise its wardship jurisdiction to make a decision based on the best interests of the minor).
41. *Re F* (*Sterilization: Mental Patient*), [1989] 2 FLR 376 (HL).

immediate urgency of the occasion, and a due regard for the public safety or convenience".[42] Necessity provided a justification for medical treatment without consent in situations of emergency, as where a surgeon amputated the limb of an unconscious passenger to free him from the wreckage of a railway accident. F, as a person with a disability, could not coherently be regarded as existing in a "permanent state of emergency", but a "clear and logical connection" *did* exist between the position of a person unable to consent by reason of emergency and a person unable to consent by reason of lasting mental incapacity.[43] In both cases, disallowing medical treatment that was in a patient's best interests would effectively deprive that person of the care that he or she would receive if able to give consent; that deprivation was more meaningful to the individual than the corresponding abrogation of the right to non-interference protected by the doctrine of trespass.[44] It was the court's obligation to fill the gap left by the legislature's removal of *parens patriae*. Using the common law (the "great safety net which lies behind all statute law")[45] to "fill gaps, if and in so far as these gaps have to be filled in the interests of society as a whole" was described by Justice Donaldson in the Court of Appeal as "one of the most important duties of judges".[46]

Lord Goff of Chiveley distinguished between situations of true emergency (the unconscious passenger in a railway accident) and situations involving a "permanent or semi-permanent state of affairs".[47] Necessity justified intervention in both situations, but the permissible scope of intervention was different in each. The medical intervenor in a true emergency situation could lawfully do no more than what was reasonably required in the best interests of the patient until the patient regained the ability to consent. This limitation had no rational basis where the "state of affairs" precluding consent was "permanent or semi-permanent": a person in this situation will likely *never* be able to consent

42. *Morey v Fitzgerald* (1884), 56 Vt 487 (Sup Ct (Vermont)) at 489.
43. *In Re F*, *supra* note 3 at 17.
44. *Mallette v Shulman* (1990), 72 OR (2d) 417 (CA).
45. *In Re F*, *supra* note 3 at 13.
46. *Ibid*.
47. *Ibid* at 25.

or, if so, will only be able to do so after a lengthy and indeterminate period of time. In the meantime:

> the need to care for [such a patient] is obvious and the doctor must then act in the best interests of his patient, just as if he had received his patient's consent to do so. Were this not so, much useful treatment and care could, in theory at least, be denied to the unfortunate.[48]

Unlike the emergency situation, "humdrum" or "routine" care, including "simple care such as dressing and undressing and putting to bed",[49] would also be justified on the basis of necessity (with no requirement of legal authorisation) where the inability to consent was a permanent or semi-permanent state.

A majority in the House of Lords, as in the Court of Appeal, found that the sterilisation procedure was justified on the basis of necessity and therefore lawful without the approval of the court (in the same way that approval was not required before emergency treatment could be lawfully given). Nevertheless, the special nature of the procedure, involving potentially competing interests (between F, her mother, and the physicians) and the fundamental personal rights of F engaged, made the involvement of the court *desirable* (and also practicable, as no emergency "on the spot" medical decision making was required). The Court of Appeal had concluded that a new rule of court, requiring a determination by the court that a procedure of this kind was in the best interests of the patient, was needed. A mere declaration that the operation would be lawful would "change nothing" and merely declare that "had a course of action been taken without resort to the court, it would have been lawful anyway".[50] This was inadequate:

> [i]n the context of the most sensitive and potentially controversial forms of treatment the public interest requires that the courts should give express

48. *Ibid* at 26.
49. *Ibid* ("[w]hen the state of affairs is permanent, or semi-permanent, action properly taken to preserve the life, health or well-being of the assisted person may well transcend such measures as surgical operation or substantial medical treatment and may extend to include such humdrum matters as routine medical or dental treatment, even simple care such as dressing and undressing and putting to bed" at 26).
50. *Ibid* at 9.

> approval before the treatment is carried out and thereby provide an independent and broad based third opinion.[51]

In the meantime (pending formulation of this new rule) the court was "fortunately" able to draw on its "inherent jurisdiction to regulate its own proceedings", which meant that approval of the court was required before the sterilisation could proceed.[52]

The House of Lords held that the court had no jurisdiction to create the new rule proposed by the Court of Appeal, as this would effectively replicate the *parens patriae* jurisdiction with respect to mentally incapable adults that the legislature had removed:

> [i]f [the *parens patriae* jurisdiction], or something comparable to it, is to be re-created, then it must be for the legislature and not for the courts to do the re-creating. Rules of Court can only, as a matter of law … prescribe the practice and procedure to be followed by the court when it is exercising a jurisdiction which already exists. They cannot confer jurisdiction, and, if they purported to do so, they would be *ultra vires*.[53]

A declaration could not be *required* in a situation of this kind,[54] but it was "open to the court under its inherent jurisdiction to make a declaration that a proposed operation was in a patient's best interests" and in the current case it was "highly desirable"[55] that such a declaration should be sought by those caring for the woman. A declaration would, as Lord Donaldson noted, "change nothing" in the sense that it could not "make lawful that which would otherwise be unlawful".[56] A declaration would establish by judicial process, however, "whether the proposed operation is in the best interests of the patient and therefore lawful, or not in the patient's best interests and therefore unlawful".[57] In order to make

---

51. *Ibid* at 13.
52. *Ibid* at 10.
53. *Ibid* at 12, per Lord Brandon of Oakwood.
54. *Ibid* ("[t]he rule [pertaining to the court's power to make declarations] does no more than say that there is no procedural objection to an action being brought for a declaration whether any other kind of relief is asked for or available or not" at 13).
55. *Ibid* at 5.
56. *Ibid* at 23.
57. *Ibid* at 13, per Lord Brandon of Oakwood.

a declaration of lawfulness (because it is in the patient's best interests), the court would be obliged to make an inquiry and "reasoned decision" *about* those best interests, "substantially the same" process as if the court's approval were required through a new rule.[58] In effect, the mechanism of the declaration would provide the "independent and broad based third opinion" sought by the Court of Appeal through the new rule. "If the old *parens patriae* jurisdiction were still available … there would be no difficulty", Lord Brandon noted:

> [but] having regard to the present limitations on the jurisdiction of the court, by which I mean its inability to exercise the *parens patriae* jurisdiction with respect to adults of unsound mind, the procedure by way of declaration is, in principle, an appropriate and satisfactory procedure to be used in a case of this kind".[59]

Lord Goff concluded that there seemed "little, if any, practical difference between seeking the court's approval under the *parens patriae* jurisdiction and seeking a declaration as to the lawfulness of the operation".[60]

In the opinion of Lord Griffiths, the involvement of the court in these circumstances was not only desirable, but should be required, not by a rule of court, but by the doctrine of necessity itself. "The law ought to be that [medical providers] must obtain the approval of the court before they sterilise a woman incapable of giving consent. I believe that it is open to your Lordships to develop a common law rule to this effect".[61] The common law had proved sufficiently flexible in the past to develop public interest based exceptions to the general rule "that the individual is the master of his own fate" by placing "constraints on the harm that people may consent to being inflicted on their own bodies".[62] "The time has now come", Lord Griffith concluded, "for a further development to forbid, again in the public interest, the sterilisation of a woman with healthy reproductive organs who, either through mental incompetence

58. *Ibid*.
59. *Ibid* at 14.
60. *Ibid* at 32.
61. *Ibid* at 19-20.
62. *Ibid* at 20 citing *Attorney General's Reference* (*No. 6 of 1980*), [1981] QB 715 (CA (Crim)(Eng)); *Rex v Donovan*, [1934] 2 KB 498 (CA (Crim) (Eng)).

or youth, is incapable of giving her fully informed consent unless such an operation has been enquired into and sanctioned by the High Court".[63] As "second best" to a new common law rule, Lord Griffith accepted the declaration procedure described by Lords Brandon and Goff.

B. After *In Re F*: Developing the Inherent Jurisdiction

A series of cases following *In Re F* applied the interpretation of necessity developed in that case (itself an extension through analogy of the necessity justification in emergency settings to non-emergency medical treatment for persons unable to consent by reason of mental disability) to justify both medical and non-medical interventions. In both medical and non-medical settings, intervention could be justified on the basis of necessity only if in the best interests of the individual; outside of the medical context this has been interpreted as a requirement that intervention is necessary for the purpose of safeguarding the *rights* of an individual who, by reason of mental incapability, is incapable of doing so herself. The mechanism of the declaration was used in both contexts to enable a "third opinion" on the question of whether the intervention proposed was in the best interests of the individual concerned and, therefore, lawful.

Sir Stephen Brown, in a case involving a mentally incapable person (to whom the old *parens patriae* jurisdiction would have applied) and "special category"[64] medical treatment, described the inherent jurisdiction "discovered" in *In Re F* as a "patrimonial" jurisdiction, "not strictly *parens patriae* but similar in all practical respects to it".[65] In *In Re S (Adult Patient: Sterilisation)*[66] (another case, like *In Re F*, involving a mentally incapable patient and a proposed sterilisation procedure), Lord Thorpe referred to the relationship between this "patrimonial jurisdiction" and *parens patriae* as:

> a distinction without a difference … By which I mean that the parens patriae jurisdiction is only the term of art for the wardship jurisdiction which is alternatively described as the inherent jurisdiction. That which is patrimonial

63. *In Re F*, *supra* note 3 at 20.
64. Involving competing interests, fundamental rights and the public interest.
65. *Re G* (*Adult Patient: Publicity*), [1995] 2 FLR 528 (Fam (Eng)) at 530.
66. [2001] Fam 15 (CA (Civ)(Eng)).

> is that which is inherited from the ancestral past. It therefore follows that whilst the decision in *Re F* signposted the inadvertent loss of the *parens patriae* jurisdiction in relation to incompetent adults, the alternative jurisdiction which it established, the declaratory decree, was to be exercised upon the same basis, namely that relief would be granted if the welfare of the patient required it and equally refused if the welfare of the patient did not.[67]

The distinction between the necessity-based declaration and the old *parens patriae* emerges with greater clarity in a subsequent series of cases concerning the rights and best interests of incapable individuals outside of the medical context.

In Re C (Adult Patient)[68] ("*Access: Jurisdiction*") concerned a situation where one parent was restricting the access of another to their mentally disabled adult child (who was herself unable to consent to or to refuse the restriction). Justice Eastham found that the child had a common law right to freedom of association and that the conduct of the parent was in violation of that right. A declaration in such a case could be granted. It would not work to *make* the restriction of access illegal, but simply recognise it as such (because it was in violation of the adult child's right). *In Re S (Hospital Patient: Court's Jurisdiction)*[69] concerned a patient ("S") who had become mentally incapable following a stroke. S was currently being cared for in a hospital in England, where he had lived for many years, but his estranged wife now wished to move S to Norway. S's long term English mistress sought a declaration from the court, on the basis of its inherent jurisdiction, that it would be unlawful to remove S from England. This case raised the question of who was entitled to bring an application for a declaration on the basis of the rights of a mentally incapable person who was unable to consent to the intervention.[70] The mistress in *In Re S*, unlike the parent in *In Re C*, did not have a recognised legal basis on which to seek the declaration with regard to the rights of S. The court held that the jurisdiction could be invoked by any party whose past or present relation with the incapable person gave him a genuine and legitimate interest in obtaining a decision (and not a "stranger" or

67. *Ibid* at 29-30.
68. [1994] 1 FCR 705 (Fam (Eng)) [*In Re C*].
69. [1996] Fam 1 (CA (Civ)(Eng)).
70. Pending a determination of S's best interest.

"officious busybody") and that the matter in question — S's residence — was one in respect of which a declaration of lawfulness could be made.[71] In *Cambridgeshire County Council v R and Others*,[72] an application for a declaration was brought by a local authority (as a body with a legitimate interest in the rights of "R"). The application was not successful for two reasons. First, it had not been established that R was incapable of making her own decision about the proposed intervention (and therefore the doctrine of necessity did not apply);[73] second, it had not been established what rights of R, if any, were in need of safeguarding.[74] R was a 20 year old woman with a learning disability who had been taken into care at the age of 10. R's father had been sentenced to four years' imprisonment after admitting to serious sexual offences committed against R when she was a child; the other members of R's family, including her mother, had always denied that the offences took place. At the time the application was brought, R was living in supported accommodation provided by the local authority. The authority was now worried that R's mother was trying to persuade R to leave her current housing and return to live with her family — a move the authority believed would have very negative consequences for R. The authority asked the court, on the basis of necessity and the inherent jurisdiction recognized in *In Re F*, to make a declaration that the authority could lawfully prevent R's family from removing or attempting to remove R from her present accommodation and from contacting R without the authority's consent.

The local authority had maintained that R was not capable of making this decision, proposing that the following test of decision-making capability was appropriate in this case:

i. if unsupervised contact would be damaging to R's welfare,
ii. the court should consider the intention likely to be held by a person of proper understanding in respect of it, and
iii. if such a person would be likely to object to it, then

71. *In Re C*, *supra* note 68 at 2.
72. [1994] 2 FCR 973 (Fam (Eng)).
73. *Ibid* at 975.
74. *Ibid*.

iv. (on the assumption that the right not to have contact is a right protected by the common law) the law should afford a person who is not legally competent the same protection as it would afford the legally competent.[75]

Rejecting this test, Lady Hale observed:

> [t]hat it provides no help in deciding who is or is not legally competent and comes dangerously close to asserting that someone who decides to do things which others consider are not in their best interests is for that very reason incompetent. That has never been the law in this country. The test of competence in other areas has always been the capacity to understand the nature and effect of the transaction or other action proposed.[76]

Furthermore, Lady Hale described the declaration sought as one which would effectively transform a lawful activity (R's communication with her family members) into an unlawful one, and not a "mere" declaration of the lawfulness or unlawfulness of the activity in question:

> [i]t is necessary to ask what legal right there is for R to be protected against the actions which the local authority seeks to control or prohibit in this case, and also what legal right the authority has to be appointed in effect as her protector. It is access, or freedom of association, rather than harassment, or freedom from association, which is protected under English law. … Far from supporting a legal right, the declarations sought would interfere with one, and in circumstances in which it has not been argued before me that a legal justification for doing so exists.[77]

Six years later, a similar situation was considered in *Re F (Adult: Court's Jurisdiction)*.[78] *Re F* concerned a young girl ("T"), now 18, who was described as having an intellectual age of 5 to 8 years old. T had been placed in local authority accommodation for persons with mental disabilities at the age of 16 with the consent of her mother. Prior to that, T's home life with her parents was described as abusive and neglectful (such that the local authority eventually placed T's seven younger siblings in care). The mother had subsequently withdrawn her consent to T's accommodation placement; T had also expressed a desire to live with her mother. When T turned 18 the local authority had succeeded in

75. *Ibid* at 977.
76. *Ibid* at 975-76.
77. *Ibid* at 976-77.
78. [2000] EWCA Civ 192 [*Re F*].

obtaining an order for guardianship, but that order was overturned by the Court of Appeal[79] on the basis that recent legislative amendments had "radically restricted" the "categories of people who could be received into guardianship" excluding persons in the position of T:

> [g]uardianship cannot now be used for clients who suffer from any form or arrested or incomplete development of the mind unless it is associated with "abnormally aggressive" or "seriously irresponsible" conduct. Unless the meaning of these words is distorted, the vast majority of those with a learning disability (mental handicap) will be excluded from guardianship. The benign side of the guardianship coin was nowhere in evidence in the new legislation. The present state of the statute books therefore reflects a single-minded view of personal guardianship as a method of restricting civil rights and liberties rather than as a method of enhancing them.[80]

Applying this restrictive construction of the legislation to T, her desire to return to the family home was not "seriously irresponsible" in the sense required, and the guardianship order was overturned.

The local authority now sought a declaration on the basis of the inherent jurisdiction of the court that it could lawfully restrict T's contact with her natural family (principally her mother) and that T should remain in the local authority accommodation. The court found that T was *not* capable of making the decision of whether to have contact with her family (unlike R in the *Cambridgeshire County Council* case) and that doing so would be deleterious to her rights (which T was unable to protect herself). Lord Sedley opined that "T is so unable to judge what is in her own best interests that no humane society could leave her adrift and at risk simply because she has reached the age of 18".[81] T's situation was, in this sense, analogous to that of the young woman in *In Re F*: unable (in the opinion of the court) to make the crucial decision herself with no-one able to consent on her behalf (the guardianship order in respect of T having been overturned).

As in *In Re F*, the court found that the common law doctrine of

79. *Re F* (*Mental Health Act: Guardianship*), [2000] 1 FCR 11 (CA (Civ) (Eng)).

80. *Ibid* at 17, citing UK, Law Commission, *Mental Incapacity* (Law Commission Report 231)(London: Her Majesty's Stationary Office, 1995) at para 2.2.1.

81. *Re F, supra* note 78 at 48.

necessity was the basis on which the court could make the declaration that was being sought. "If there is no recourse to the doctrine of necessity, the court has no jurisdiction to make any declarations to enable the local authority to … regulate future arrangements for T".[82] In *In Re F*, necessity "filled the gap" left by the disappearance of *parens patriae* regarding the decision in question without legislative replacement. This case raised the question of whether an analogous gap was created by the legislature's "radical restriction" of guardianship legislation (prior to which the local authority could have acted as T's guardian) *or* whether the legislature had intended to create a law-free space for individual autonomous choice (a space to remain un-filled rather than a gap). Would the court, in making the declaration requested, be "assuming an inherent power to restore what parliament had removed" through its "deliberate and wholesale curtailment" of guardianship?

The court concluded that the reform of guardianship legislation had created

> [a]n obvious gap in the framework for care of mentally incapacitated adults. If the court cannot act and the local authority is right [regarding the abusive home environment] this vulnerable young woman would be left at serious risk with no recourse to protection, other than the future possibility of the criminal law. This is a serious injustice to T who has rights which she is, herself, unable to protect. … [quoting Lord Donaldson at the Court of Appeal in *In Re F*] The common law is the great safety net which lies behind all statute law and is capable of filling gaps in that law, if and inso far as those gaps have to be filled in the interests of society.[83]

The restriction of the legislation, coupled with T's inability to protect her own interests, meant that, without the intervention sought by the local authority, T would be effectively deprived of rights to which she would otherwise be entitled (either because T could have protected her rights independently or because a guardian could have done so on T's behalf). Lord Justice Thorpe concluded his reasons by cautioning that his judgment should *not* be understood as "restoring more or less the *parens patriae* jurisdiction, albeit re-labelled",[84] referring to Lord Goff's

82. *Ibid* at 39.
83. *Ibid* at 41-42.
84. *Ibid* at 47.

discussion in *In Re F* of the "wider range" of care justified by necessity in cases involving persons with permanent or semi-permanent mental disorder (the implication being that the contact restrictions were an extension of the kind of non-medical "hum-drum" or everyday care decisions Lord Goff described).

In Re A Local Authority (A Restraint on Publication)[85] concerned a group of individuals characterised as "vulnerable" and "adults under a disability". The mental capability of these individuals, or lack thereof, is never established, although they are treated by the court as mentally incapable persons to whom the *In Re F* jurisdiction applies (and to whom the *parens patriae* jurisdiction would have applied prior to its demise).[86] This case is an important turning point in the development of the jurisdiction through the case law. Unlike the previous cases discussed, the doctrine of necessity and declaration mechanism are used here to justify a protective order — an injunction — and not merely to declare the lawfulness (or unlawfulness) of a proposed course of action.

The "adults under a disability" whose rights were in issue in this case had all lived in a particular foster home as children and had returned to live in the home as adults. The foster home had been the subject of an inquiry carried out by the local authority, and the public solicitor representing these "vulnerable" adults now sought a ban on publication of the inquiry report on the grounds that the ensuing media scrutiny would cause upset and stress.[87] Dame Butler-Sloss agreed,[88] noting that,

85. [2003] EWHC 2746 (Fam).
86. *Ibid* at paras 86-97.
87. *Ibid* at paras 41-42.
88. *Ibid* ("[t]hey have now returned to live at the Home. They have had consideration and distressing disruption of their lived and are, as set out in the Report, vulnerable. A period of peace, stability and a chance to settle down again after the very real upset of their lives is threatened by the likely intense media cover if this Report is published. They are all under some disability but not such, as far as I know, as to prevent possibly all of them, but certainly at least 4 of them, from understanding the impact of press and other media intrusion. That intrusion would affect their daily lives and would be very likely to be disruptive, distressing and contrary to the need for them to settle back in the Home" at para 98).

following the decision in *In Re F*:

> the circumstances within which a court will exercise the inherent jurisdiction through the common law doctrine of necessity are not restricted to granting declarations in medical issues. It is a flexible remedy and adaptable to ensure the protection of a person who is under a disability. It has been extended to questions of residence and contact. Until there is legislation passed which will protect and oversee the welfare of those under a permanent disability the courts have a duty to continue, as Lord Donaldson said in *Re F (Mental Patient: Sterilisation)*, to use the common law as the great safety net to fill gaps where it is clearly necessary to do so.[89]

The new question in this case was whether the inherent jurisdiction could be exercised for the purposes of making a positive order that would ban publication of the report (justified on the basis of necessity):

> [i]n the previous cases about adults under a disability, the issues have been the lawfulness of the proposed course of action and considerations as to their best interests. That cannot be the correct approach in the present case. The application of the inherent jurisdiction would seem more appropriately to be treated as the exercise of a protective jurisdiction rather than a custodial jurisdiction.[90]

Dame Butler-Sloss granted an injunction preventing the authority from publishing the report.

The "flexible" remedy described in *In Re A Local Authority* was developed further in *Re G (An Adult)*,[91] in which the court relaxed (if it did not yet abandon) the requirement of mental incapability which, through analogy to the inability to consent in emergency medical situations, had provided the conceptual basis for the application of the common law doctrine of necessity from *In Re F* onwards.

Re G concerned a young woman ("G") who was not mentally incapable at the time the application was brought, although she had been incapable in the past by reason of her mental illness. G's condition had stabilised under psychiatric care and for the last ten years she had been residing in supportive housing provided by the local authority. The local authority now sought a declaration that it could lawfully restrict contact

89. *Ibid* at para 96.
90. *Ibid* at para 97.
91. [2004] EWHC 2222 (Fam) [*Re G*].

between G and her father.[92] G's father had been violent towards G and her mother in the past and the local authority was concerned about the strongly negative impact that contact with her father had on G. Following a previous period of contact with her father, G's condition regressed to the point that she became mentally incapable. Justice Bennett agreed with the experts involved in G's care that "[i]f the restrictions [on G's contact with her father] were lifted, G's mental health would deteriorate to such an extent that she would again become incapacitated … Such a reversion would be disastrous for G".[93]

Justice Bennett dismissed as "unattractive" the proposition that the court's jurisdiction would be "entirely dependent on the shifting sands of whether or not G did, or did not, have the requisite mental capacity at a particular time".[94] The doctrine of necessity therefore applied here, as in *In Re F*, to justify the restrictions the authority sought to impose. Quoting extensively from *In Re F*, Bennett J agreed with Lord Justice Sedley in that case that the doctrine of necessity was not restricted to "medical and similar emergencies".[95] "The concept of necessity has its role to play in all branches of our law of obligations — in contract … in tort … in restitution … and in our criminal law. It is therefore a concept of great importance".[96] The common law doctrine of necessity justified the intervention sought by the local authority (restricting G's contact with her father) for the purpose of safeguarding G's rights — in this case, her right to not be deprived of her mental capability:

> If the declarations sought are in G's best interests, the court, by intervening, far from depriving G of her right to make decisions…will be ensuring that G's now stable and improving mental health is sustained, that G has the best possible chance of continuing to be mentally capable, and of ensuring a quality of life that [she had previously been] unable to enjoy.[97]

---

92. *Ibid* at para 1.
93. *Ibid* at para 86.
94. *Ibid* at para 91.
95. *Ibid* at para 102.
96. *Re G*, *supra* note 91 referring to the statement of Lord Goff of Chieveley in *R v Bournewood Community and Mental Health Trust*, [1999] AC 458 (HL).
97. *Ibid* at para 104.

> [...]
>
> In my judgement, the common law demands that … the court act by investigating, and if it is in G's best interests, making the declarations sought … the 'focal point of the inquiry must be the situation which … has led to the application for declaratory relief under the inherent jurisdiction'.[98]

IV. *Re SK* to *DL v A Local Authority*: Development of the Vulnerability Jurisdiction

A series of cases decided after *Re G* apply the *In Re F* declaration (*i.e.* a declaration used as a means of determining the substantive lawfulness of a proposed intervention and thereby requiring a determination of whether the intervention is in the best interests of the individual) in situations involving the rights of individuals who, while identified as "vulnerable", are indisputably mentally capable. The source of vulnerability described in each case is an oppressive or exploitative relationship context from which the individual cannot separate herself[99] and by reason of which she cannot safeguard her rights independently. *Re G* can be seen to form a bridge between these later cases and the cases coming before it. The ultimate objective of the intervention proposed by the local authority in *Re G* was the prevention of G's regression to her earlier state of mental incapacity; the source of the threat both to G's physical wellbeing (the resurgence of her mental illness) and to her rights (her right to mental capability and therefore autonomy) was her relationship with her abusive father; the disruption of that relationship was the immediate objective of the proposed intervention.

The missing element of mental incapability in these cases changes the legal (common law and equity) basis on which the declaration is made in these cases and, thereby, the kind of intervention that can be justified. The medical intervention proposed in *In Re F* was lawful on the basis of necessity *because* F was mentally incapable and for that reason unable to consent, with no one else able to consent on her behalf (establishing, through analogy to the emergency cases, the common law justification of

98. *Ibid* at para 112.

99. All of the cases discussed here involve women.

necessity *if* the intervention could be shown to be in F's best interests). The cases discussed in the previous section, whether within or outside of the medical context, retain the fundamental premise of that analysis: in situations where a person's mental incapacity makes consent impossible, and where substitute consent is not available, necessity justifies an intervention that is in the person's best interest (and to which he or she would be entitled if able to consent).

The cases discussed in this section could be explained as a further development-through-analogy of the *In Re F* necessity analysis, extending the justification of necessity to situations where an oppressive and/or exploitative relationship context deprives the individual of her ability to consent in a way that is analogous to the incapacity of the unconscious train-wreck survivor (the original emergency exception). The analogy to emergency medical treatment is stretched thin, however, in a way that is inconsistent with the situation of necessity within the wider framework of tort law as a limited *exception* to the rule that the individual is, in the words of Lord Griffiths, the "master of her fate". It has also been suggested that the characterisation of a new "vulnerable persons" category as equivalent to unconscious victims or permanently consent-disabled persons (as described in By Lord Goff in *In Re F*) reduces the "vulnerable" to a bundle of faulty personal characteristics or risk factors and deprives individuals so labelled of the free choice to which they would otherwise be entitled.[100]

I propose that it is more coherent (and less problematic in terms of autonomy) to understand these cases as marking a decisive conceptual break from the earlier necessity based cases following *In Re F*. In the case of a mentally *capable* adult in a non-emergency situation, the doctrine of necessity (with regards to intervention without consent) does not coherently apply. The "lawful" basis of the interventions sought in the cases involving vulnerable but capable persons is more correctly understood as the equitable doctrine of undue influence and the principles underlying it; through the mechanism of the declaration, drawing on the inherent

100. Michael C Dunn, Isabel CH Clare & Anthony J Holland, "To Empower or to Protect? Constructing the 'Vulnerable Adult' in English Law and Public Policy" (2008) 28:2 Legal Studies 234 at 241.

jurisdiction of the court, these cases show the court giving effect to the principles underlying the doctrine in novel situations[101] (as the court used the mechanism of the declaration to interpret and give effect to the doctrine of necessity in the novel situation presented by *In Re F*). The use of the declaration as a mechanism through which the court may exercise its inherent jurisdiction in the way described in *In Re F* and subsequent cases is not limited to situations involving the common law doctrine of necessity; rather, the necessity cases provided the context in which this particular means of exercising the inherent jurisdiction originated, but to which it is not contained.

The language of "capacity" in the vulnerable-but-capable cases is potentially confusing but the mere word "capacity", like vulnerability, has more than a single meaning; the more sensible meaning of "capacity" in the context of these cases relates to the nature of free choice as explained by the doctrine of undue influence. That idea is not the same as the cognitive ability described as "mental capacity" in the necessity cases. The fact that the declaration cases applying the doctrine of necessity outside of the medical context (discussed in the previous section) also concern relationships of oppression/exploitation, is another source of confusion. The objective of the proposed interventions in the vulnerable-but-capable cases is crucially different, however, in a way that is consistent with the distinct basis of their "lawfulness": not the protection of rights and interests (as in the necessity-based non-medical cases) but the facilitation of free choice through the disruption of oppressive/exploitative relationship context.

A. *Re SK*

With the exception of *Re G* (involving fluctuating mental capacity), *Re SK* (*Proposed Plaintiff*) *(An Adult by way of her Litigation Friend*),[102] is the first case in which the court, drawing on its inherent jurisdiction, declared

101. The doctrine of undue influence developed in the transactional context, including gifts (and wills), although its theoretical framework has also been applied to consent in other contexts; see *Norberg v Wynrib*, [1992] 2 SCR 226.

102. [2004] EWHC 3202 (Fam).

as lawful a proposed non-medical intervention regarding an individual who, while characterised as vulnerable, was clearly mentally capable.

Re SK involved a female British citizen ("SK") who, consular officials suspected, was being kept in Bangaldesh for the purposes of a forced marriage. A solicitor in the Community Liaison Unit at the Foreign and Commonwealth Office brought an application asking that the court make an order, on the basis of its inherent jurisdiction, for the purpose of "protect[ing] and secur[ing] the well-being and best interests of SK and to ensure that she may freely express her wishes concerning her country and place of residence and concerning her marital status".[103] The order requested would require SK's family to "assist and allow" SK to visit the British High Commission and be interviewed alone. It did not cause SK to undergo a marriage ceremony and did not "threaten, intimidate … harass" or use violence towards SK.[104]

There was no doubt that SK was mentally capable; nor had she ever been mentally incapable. Nevertheless, Justice Singer concluded that, if the "gravely disquieting" information received by the consular offices in Bangladesh and London proved to be substantially well-founded,

> [t]here would be serious cause for concern about her capacity to control her own life and destiny at the moment. This notwithstanding that she is an adult and is emancipated, at least in terms of English law, and should not be the subject of duress or force or be deprived of the ability to make her own decisions.[105]
>
> If, in truth, she were forced to marry or if, in truth, that is the outcome which she may contemplate and fear, then steps taken in furtherance of those ends would be a series of acts to which she did not consent. Indeed her very capacity to consent would have been overborne by fear, duress or threat. If therefore she has been through or faces the prospect of going through a ceremony of marriage with which she is, in fact, not in agreement it would be a voidable marriage, but nevertheless one which might engender irreparable and severe physical and emotional consequences for its victim.[106]

The court concluded that "the inherent jurisdiction now, like wardship has been, is a sufficiently flexible remedy to evolve in accordance with

103. *Ibid* at annex.
104. *Ibid* at paras 2-4.
105. *Ibid* at para 3.
106. *Ibid* at para 4.

social needs and social values" and granted the order sought.[107]

A similar situation was considered one year after *Re SK* in *Re SA (Vulnerable Adult with Capacity: Marriage)*.[108] That case concerned a British woman age 18 ("SA") whose family was of Pakistani Muslim origin. SA was profoundly deaf and communicated through British Sign Language, which neither of her parents understood; SA's communication with her family was therefore extremely limited as she could not understand, lip-read or sign in Punjabi or Urdu (the main languages spoken in the family). SA also had significant visual loss in one eye and had been assessed as having the intellectual level of a 13 or 14-year-old child. The local authority was worried that the family of SA planned to arrange, or possibly even force SA into a marriage in Pakistan. SA had expressed that she was happy to have an arranged marriage but would want to approve her parents' choice of husband for her. She also wanted any future husband to speak English and to come and live in the UK; she did not want to go live in Pakistan.[109]

SA had recently been assessed by a Forensic Psychologist (working for the local authority) as having the mental capacity required for marriage. The assessor also noted that if SA married a person who could not communicate with her, or if she was moved to an environment where she was entirely surrounded by people who could not communicate with her, it was very likely that SA would become extremely distressed and isolated, posing a significant risk to her future well-being and mental health.[110] On this basis, the local authority sought an order similar to the

107. *Ibid* at para 8.
108. [2005] EWHC 2942 (Fam).
109. *Ibid*.
110. *Ibid* at para 15.

order sought and granted in *Re SK*.[111]

Sir James Munby, giving judgment in the case, explained the basis for exercising the inherent jurisdiction "rediscovered" in *In Re F* in this case.[112] While it had always been recognised that the "jurisdiction is exercisable in relation to any adult who is for the time being, and whether permanently or merely temporarily, either disabled by mental incapacity from making his own decision or, although not mentally incapacitated, unable to communicate his decision"[113] the immediate question was "whether the jurisdiction extends further".[114] Surveying the case law Sir Munby concluded that, "[i]n my judgment, it does. I must now explain why":[115]

> [i]n the light of these authorities it can be seen that the inherent jurisdiction is no longer correctly to be understood as confined to cases where a vulnerable adult is disabled by mental incapacity from making his own decision about the matter in hand and cases where an adult, although not mentally incapacitated, is unable to communicate his decision. The jurisdiction, in my judgment, extends to a wider class of vulnerable adults.[116]
>
> It would be unwise, and indeed inappropriate, for me even to attempt to define who might fall into this group in relation to whom the court can properly exercise its inherent jurisdiction. I disavow any such intention. It suffices for present purposes to say that, in my judgment, the authorities to which I have referred demonstrate that the inherent jurisdiction can be exercised in relation to a vulnerable adult who, even if not incapacitated by mental disorder or

111. *Ibid* (the order made in *Re SA* prohibited her family from threatening, intimidating or harassing SA; using violence on SA; or preventing SA from communicating alone with her solicitor. SA's family was also prohibited from applying for any travel documents for SA; removing or attempting to remove SA from the jurisdiction of England and Wales; and from causing, making arrangements for, or permitting SA to be married without her express written consent. The order also provided for undertakings from a groom that he will return to live in England if SA wished, and that, if SA were to remain in Bangladesh after marriage, a visit with an official from the British High commission would be arranged for the purpose of establishing her free consent to remain).
112. *Ibid* at para 46.
113. *Ibid*.
114. *Ibid* at para 48.
115. *Ibid*.
116. *Ibid* at para 76.

> mental illness, is, or is reasonably believed to be, either (i) under constraint or (ii) subject to coercion or undue influence or (iii) for some other reason deprived of the capacity to make the relevant decision, or disabled from making a free choice, or incapacitated or disabled from giving or expressing a real and genuine consent.[117]

This inherent jurisdiction would apply to all persons whose "capacity" (as described by Sir Munby) has been impaired in one of the senses, and for one of the reasons, referred to in the passage above but was "not confined to those who are vulnerable adults, however that expression is understood, nor is a vulnerable adult amenable as such to the jurisdiction".[118]

> The significance in this context of the concept of a vulnerable adult is pragmatic and evidential: it is simply that an adult who is vulnerable is more likely to fall into the category of the incapacitated in relation to whom the inherent jurisdiction is exercisable than an adult who is not vulnerable. So it is likely to be easier to persuade the court that there is a case calling for investigation where the adult is apparently vulnerable than where the adult is not on the face of it vulnerable. That is all.[119]

Sir Munby's reference to the "concept of the vulnerable adult" seems in relation to members of vulnerable populations, who are only more likely to be "incapacitated or disabled from giving or expressing a real and genuine consent".

Granting the order sought ("designed to provide a practical solution to the concerns raised by the local authority and other professionals and, very importantly, to reflect what SA herself wants and expects from her husband"),[120] Sir Munby stated that:

> [b]y taking this course, far from depriving SA of her right to make decisions I am ensuring, as best I can, that she has the best possible chance of future happiness. I am taking these steps to protect, support and enhance SA's capacity to control her own life and destiny in the way she would wish.[121]

The analysis developed in *Re SK* and *Re SA* was applied outside of the arranged marriage context in the case of *A Local Authority v A*,[122] regarding

117. *Ibid* at para 77.
118. *Ibid* at para 83.
119. *Ibid*.
120. *Ibid* at para 27.
121. *Ibid* at para 133.
122. [2010] EWHC 1549 (Fam).

a "vulnerable" 29 year old woman with severe learning difficulties and an assessed IQ of 53 ("A") who was found to be incapable of making decisions in relation to contraception, "not to a small extent due to her husband's negative influence on her decision-making capacity".[123] Before her marriage A had given birth to two children, both of whom had been removed from her care after birth. The local authority was concerned that her husband was putting her under pressure to refuse contraception, and provided evidence that A had complained that her husband had hit her and that she did not wish to have a baby.[124] A is *not* described in the case as "mentally incapable"; rather, the court ascribes A's inability to make her *own* decision regarding contraception to the totality of her "cognitive limitations", "social impairment", and "personal characteristics, associated with both her learning disability and her personality, in connection with her 'ambivalence (including mixed feelings and confusion) about her husband and the pressure he seems to place on her to have a family'".[125] In the opinion of a consultant retained by the local authority, the "pressure" experienced by A from her husband was contributed to "by Mrs. A's personal characteristics" and by "Mr. A's personal characteristics, including a suspicious and hostile stance in relation to support services, leading to his giving Mrs. A mixed messages about what is in her interests, thereby 'confusing her' more and therefore incapacitating her further".[126] On the basis of the "completely unequal dynamic in the relationship between Mr. and Mrs. A" the court concluded that "her decision not to continue taking contraception is not the product of her free will"[127] and that "[w]here such circumstances pertain … the court has a wide inherent jurisdiction to prevent conduct by the dominant party which coerces or unduly influences the vulnerable party from making free decisions".[128] Regarding the question of whether the *Mental Capacity Act*, a comprehensive legislated scheme regarding mentally incapable

123. *Ibid* at paras 36-38.
124. *Ibid* at paras 18, 32, 34.
125. *Ibid* at para 51.
126. *Ibid*.
127. *Ibid* at para 73.
128. *Ibid* at para 79.

adults, had removed the need/justification for exercising the inherent jurisdiction the court in regards to an "incapacitated person" (incapacity here referring to A's inability to make her own decisions due to the matrix of factors described above), the court concluded that the "wide inherent jurisdiction" of the court applied to an incapacitated person in the same way as to a person *with* capacity, "except that the aim of providing him or her with relief from the coercion is first to gain capacity and, if achieved, then to enable him to reach a free decision".[129]

B. *DL v A Local Authority*

The case of *DL v A Local Authority* provides the most complete articulation of the equitable doctrine of undue influence as providing the doctrinal justification for the exercise of the inherent jurisdiction in the vulnerable-but-capable cases.

In that case, a local authority sought and was granted an injunction (on the basis of the inherent jurisdiction of the court to make a declaration that the injunction was lawful) against a 55 year old son ("DL") for the purpose of "regulating" his controlling, threatening and coercive conduct towards his (mentally capable) 85 year old mother. The mother, wishing to preserve her relationship with her son, did not want any proceedings taken against him.[130]

The decision was appealed on the question of "whether, despite the extensive territory now occupied by the [*Mental Capacity Act 2005*], a jurisdictional hinterland exists outside its borders to deal with cases of 'vulnerable adults' who fall outside that Act and which are determined

129. *Ibid* at paras 79-80 (in the event, as Mr. A had given assurances to the court that he would not block A's communication with professionals who could advise her in an "ability-appropriate way" about contraception, the court decided not to make an order restricting contact or conduct but to rely on Mr. A "to honour his assurances to the court … in a spirit of co-operation in trying to enable A to reach contraceptive capacity" at para 80); see also *Local Authority X v MM, KM*, [2007] EWHC 2003 (Fam); *Re A (Male Sterilisation)*, [2000] 1 FLR 549 (CA (Civ)(Eng)); *LBL v RYJ and VJ*, [2010] EWHC 2665 (Fam) [*RYJ and VJ*].

130. *Local Authority*, *supra* note 7 at para 8.

under the inherent jurisdiction".[131] The appeal was dismissed. The court defined the scope and purpose of the existing "jurisdictional hinterland" as limited "to facilitat[ing] the process of unencumbered decision-making" rather than "imposing a decision upon [a person] whether as to welfare or finance".[132] The court continued:

> I do not accept that the jurisdiction … is extensive and all-encompassing, or one which may threaten the autonomy of every adult in the country. It is … targeted solely at those adults whose ability to make decisions for themselves has been compromised by matters other than those covered by the *Mental Capacity Act 2005*. I, like Munby J before me in Re SA, am determined not to offer a definition so as to limit or constrict the group of 'vulnerable adults' for whose benefit this jurisdiction may be deployed … The appellant's submissions rightly place a premium upon an individual's autonomy to make his own decisions. However this point, rather than being one against the existence of the inherent jurisdiction in these cases, is in my view a strong argument in favour of it. The jurisdiction … is in part aimed at enhancing or liberating the autonomy of a vulnerable adult whose autonomy has been compromised by a reason other than mental incapacity.[133]

In the circumstances of this case, the conclusion that the "inherent jurisdiction remains available for use in cases that fall outside the [*Mental Capacity Act 2005*]"[134] was further justified on a "sound and strong public policy basis"[135] the "sadly all too easy to contemplate … existence of elder abuse" although "the use of the term 'elder' in that label may inadvertently limit it to a particular age group whereas, as the cases demonstrate, the will of a vulnerable adult of any age may, in certain circumstances, be overborne".[136]

> Where the facts justify it, such individuals require and deserve the protection of the authorities and the law so that they may regain the very autonomy that the appellant rightly prizes. The young woman in *Re G* who would, as Bennett J described, lose her mental capacity if she were once again exposed to the unbridled and adverse influence of her father is a striking example of precisely

131. *Ibid* at para 1, per Lord Justice McFarlane.
132. *Ibid* at para 32 quoting from the decision of Justice Macur in *RYJ and VJ*, *supra* note 130 at para 62; see also *Westminster City Council v C*, [2008] EWCA Civ 198.
133. *Local Authority*, *supra* note 7 at paras 53, 54.
134. *Ibid* at para 63, per Justice David.
135. *Ibid*.
136. *Ibid*.

this point.[137]

DL had put forward the argument that the legislature had considered incorporating undue influence into the ambit of the *Mental Capacity Act* and decided against it as an "immensely complex" exercise in drafting" that would require "significant safeguards to avoid unnecessary intervention".[138] DL's contention was that the legislature had explicitly excluded undue influence induced "incapacity" and it was not for the court to reintroduce it by other means.[139] The omission was not a gap, but a deliberate space. That argument was unsuccessful; the legislature's inability to codify undue influence confirms the court's ongoing responsibility *vis a vis* identifying and responding to this particular source of harm. The non-legislatibility of undue influence reflects the varied nature of undue influence itself, requiring in each case a "meticulous examination of the facts".[140] The nature of undue influence, and the particular inequity with which the doctrine is concerned, requires the kind of judicial response that is provided through the exercise of the inherent jurisdiction of the court.

V. Implications for Canadian Law

The interpretation and application of a distinct undue influence based exercise of the inherent jurisdiction in the English courts (the vulnerability jurisdiction described in *DL v A Local Authority*) is complicated in Canadian law by two factors. The first of these is the confusion, referred to above, between the new jurisdiction and *parens patriae*. Canada, never having excised the traditional *parens patriae* jurisdiction with regards to incapable adults (referred to in *Re Eve* as "a carefully guarded"),[141] has no need for a "new" or revived *parens patriae*. The second complicating

137. *Ibid*.
138. *Ibid* at para 37, citing a joint committee report of both the Houses of Parliament considering the draft Mental Health Bill.
139. *Ibid* at para 39.
140. *National Westminster Bank v Morgan*, [1985] 1 AC 686 at 709 (HL).
141. *Re* Eve, *supra* note 10 ("[t]he courts will not readily assume that it has been removed by legislation where a necessity arises to protect a person who cannot protect himself" at para 75).

factor is Canadian judicial interpretation of the inherent jurisdiction of the court, which has focused overwhelmingly on the "core" jurisdiction that is protected by section 96 of the Canadian *Constitution Act, 1982*.[142]

The interaction between these two factors can be seen in the decision of the British Columbia Supreme Court in *Temoin v Martin*[143] ("*Temoin*"), considered the question of whether the inherent jurisdiction could be invoked for the purpose of ordering a medical examination in connection with an application for committeeship pursuant to the *Patients Property Act*.[144] Declining to exercise the inherent jurisdiction for this purpose, Madam Justice Fisher described the "essential purpose" of the jurisdiction as

> to maintain and protect its [the court's] own adjudicative powers … by way of regulating the practice of the court and preventing abuse of its process. This is demonstrated by the kinds of cases in which inherent jurisdiction has been invoked: see, for example, *MacMillan* (contempt of court) and *Caron* (interim costs).[145]

Justice Fisher concluded that *parens patriae* was "more appropriate to the issues raised by the Petitioner"[146] but the *prima facie* incompetence required for the exercise of that jurisdiction had not been established (as *parens patriae* could not be exercised with respect to capable adults). The facts of *Temoin* are in fact suggestive of the more complicated form of impaired decision making capacity described in the case of *A v A Local Authority*,[147] an inter-section of relationship context, intellectual limitation, and personality (*Temoin* concerned an individual, M, who

142. Being Schedule B to the *Canada Act 1982* (UK), 1982, c 11; see *e.g.* *Criminal Lawyers*, *supra* note 20; see also *Reference re Amendments to the Residential Tenancies Act (NS)*, [1996] 1 SCR 186.
143. 2011 BCSC 1727, aff'd 2012 BCCA 250 [*Temoin*].
144. RSBC 1996, c 349; see *Temoin*, *supra* note 143 at para 1.
145. *Temoin*, *supra* note 143 at para 44 (cases in which the court has ordered a medical examination for the purposes of providing evidence and facilitating a fair trial include *Kujawa v Kujawa* (1990), 87 Sask R 101 (QB); *Hayman v Criddle*, 2010 SKQB 94; *Barnes (Litigation Guardian of) v London (City) Board of Education* (1994), 34 CPC 3d 51 (Ont Sup Ct J (Div Ct)).
146. *Temoin*, *supra* note 143 at para 45.
147. [2002] EWHC 18 (Fam).

was suffering from cognitive decline and also "significant pressure" from his second wife to change his will and to make other transactions).[148] Once defined as a case about the appropriate exercise of *parens patriae*, however, the focus shifted solely to the issue of M's bio-cognitive capacity and the case is replete with discussions of the various capacity tests which M had undergone and the scores or outcomes of those tests; in contrast to the English cases interpreting and applying the inherent jurisdiction, M's relationship context and its impact on his decision-making is invisible.

The inherent jurisdiction of the court has been drawn on to grant a common law restraining order in two Alberta cases: *RP v RV*[149] ("*RP*") and *ATC v NS*.[150] These cases suggest a broader interpretation of the jurisdiction[151] (beyond the "core") on the basis of its nature as "a residual source of powers, which the court may draw on as necessary whenever it is just or equitable to do so" (enabling "the judiciary to uphold, to protect and to fulfil the judicial function of administering justice according to law in a regular, orderly and effective manner").[152] Justice Hughes in *RP* granted a common law restraining order (in a family law context) describing the order as an *injunction*, "an order, historically of an equitable nature, restraining the person to whom it is directed from performing a specific act".[153] In doing so, Hughes J described the "common law jurisdiction to grant a restraining order" as

> flow[ing] from the inherent jurisdiction of provincial superior courts to hear any matter properly coming before it, in combination with the general power of those courts to grant injunctive relief as equitable remedy … [t]he discretionary power to grant all manner of injunctions is an equitable remedy

148. *Ibid* at paras 5-15.
149. 2012 ABQB 353. The applicant in that case sought a common law restraining order rather than a protection order pursuant to the *Protection Against Family Violence Act*, RSA 2000, c P-27, s 4.
150. 2014 ABQB 132 [*ATC*].
151. Although this is not explicitly stated; the implication is in the application, and the reasons given for it.
152. *ATC*, *supra* note 149 (Justice Hughes in *RP v RV*, *supra* note 148 referring to the passage from *Caron*, *supra* note 20 at para 17).
153. *ATC*, *supra* note 149 at para 15.

> that dates back to English law.[154]

RP was subsequently applied in the 2014 case of *ATC* Justice Lee concluding that the inherent jurisdiction of the court "must be more encompassing than its common law historical development and as well … go beyond its present statutory limits".[155] Granting a restraining order as an exercise of the inherent jurisdiction it was therefore necessary only for the court to determine:

> [t]hat the parties genuinely do not get along and are a threat to each other, not necessarily in terms of their personal safety or property damage, but also in terms of the damage that can be done to their reputations and lives … This harm is not physical harm involving one's personal safety, or damage or property, but still is serious emotional harm carried out through the internet, or caused by stalking and other harassing behavior … Accordingly … the practical solution to the problem is simply to use the Court's inherent jurisdiction in matters such as this to grant permanent mutual restraining orders in favour of each party against the other.[156]

The Alberta cases show the inherent jurisdiction being exercised outside of its "core" and for a purpose unconnected to the court's ability to control its own administration and operation; in these cases the jurisdiction is, indeed, being drawn upon to respond to a particular form of intensified vulnerability that is not recognized or provided for in legislation, drawing on the "great safety net" of the common law and equity to do so.

VI. Conclusion

Recognizing and articulating the "heroic judicial invention" of an "entirely novel jurisdiction" as an exercise of the inherent jurisdiction that is separate and distinct from *parens patriae* has the potential to facilitate future development of this "lusty child of equity" in the

---

154. *ATC*, *supra* note 149 at paras 16, 19 (the court's inherent jurisdiction and authority to grant equitable relief has been codified in the *Alberta Judicature Act*, RSA 2000, c. J-2, ss 8, 13(2) but did not derive from it).
155. *ATC*, *supra* note 149 at para 18.
156. *Ibid* at paras 18-19; see also *R v Burke*, 2012 NSSC 119. In this case the Nova Scotia Supreme Court invoked the inherent jurisdiction to grant a common law peace bond, while recognising that the jurisdiction should be used "sparingly".

Canadian courts.[157] This new vulnerability jurisdiction incorporates the idea of vulnerability as essential to the human condition, made more or less intense dependent on the interplay between social/relationship context and one's biological or embodied state of being (as opposed to the association of vulnerability with "vulnerable populations"). The objective of the "novel jurisdiction" is to provide a measured and coherent response to the vulnerability caused by relationships of oppression and exploitation which may then, as explained in *Re G*, increase physiological and/or cognitive resilience, and therefore autonomy. This is an exciting, and very modern, idea.

157. Munby, *supra* note 1. For Munby, of course, equity's newest child was a new *parens patriae* applying to vulnerable adults.

Charitable Trusts and Discrimination: Two Themes

Matthew Harding*

In this article, I consider two doctrinal themes available to judges in equity who must deal with what I call "discriminatory charitable trusts". In Part II, I concentrate on the theme of public policy. I review how this theme has been deployed in cases about discriminatory trusts, charitable and "private", before turning to some theoretical considerations that bear on the proper application of the public policy doctrine. In Part III, I turn to the theme of public benefit. I argue that, although the public benefit test applied in equity when working out whether a trust is for a charitable purpose is scarcely used in responding to discriminatory charitable trusts, it has considerable potential as a tool for judges seeking to respond in nuanced ways to such trusts. In Part IV, I conclude by offering some thoughts as to whether judges should opt for the theme of public policy or the theme of public benefit when deliberating about discriminatory charitable trusts.

* Professor of Law, University of Melbourne. Versions of this paper were presented at the Conference on Succession and Trusts Law at the University of the Western Cape on 17 April 2015 and at the Law of Charitable Trusts Colloquium at Université de Montréal on 8 May 2015. I learned a great deal from each of these events, and my sincere thanks are due to all those who participated in them. Very special thanks are due to François du Toit and Matthew Harrington for truly exceptional hospitality during my visits to Cape Town and Montreal.

I. Introduction

The problems associated with what, in this article, I will call "discriminatory charitable trusts" are not new to equity. However, in the twenty-first century they have assumed, and may be expected to continue to assume, greater importance than ever before. In various jurisdictions in which the political community has organised along broadly liberal lines, equitable responses to discriminatory charitable trusts now play out against a backdrop of human rights law, the constitutional expression of anti-discrimination norms, and a public culture in which tolerance of discrimination on grounds like race, sex, and religion is at its lowest point in human history. Such responses also take their place within legal and intellectual frameworks in which there is growing scepticism about the plausibility of distinctions between "public" and "private" spheres and "public" and "private" law. In utilising extant doctrinal themes to fashion just solutions to old problems, judges exercising equitable jurisdiction who must deal with discriminatory charitable trusts are presented with challenges that they have not traditionally faced.

With such challenges in view, this article will focus on the two main doctrinal themes available to judges in equity who are asked to determine whether discriminatory charitable trusts should be interfered with on account of their discriminatory character. The two themes are public policy and public benefit. In Part II of the article, I consider the theme of public policy. I review how this theme has been deployed in cases about discriminatory trusts, charitable and "private", before turning to some theoretical considerations that bear on the proper application of the public policy doctrine in such cases. I conclude that more work must be done if the theme of public policy is to be rendered appropriately sensitive to normative considerations underpinning it in cases about

discriminatory charitable trusts. In Part III, I turn to the theme of public benefit. I argue that, although the public benefit test applied in equity when working out whether a trust is for a charitable purpose is, perhaps surprisingly, scarcely used in responding to discriminatory charitable trusts, it has considerable potential as a tool for judges seeking to respond in nuanced ways to such trusts. In Part IV, I conclude by offering some thoughts as to whether judges should opt for the theme of public policy or the theme of public benefit when deliberating about discriminatory charitable trusts.

At the outset, two points of clarification are in order. First, when I refer to "discriminatory" trusts, I have in mind trusts the terms of which explicitly mete out unfavourable treatment to some class of persons based on the fact that the class shares an element or elements of human identity. However, I make no attempt to describe or explain the circumstances in which such discrimination ought to be of moral or legal concern.[1] Instead, I rely on what I take to be the intuitive proposition that at least some instances of such discrimination ought to be of both moral and legal concern, and I assume that this proposition is sufficient to animate my arguments in this paper. Secondly, when I refer to discriminatory "charitable" trusts, I describe such trusts as charitable in a provisional sense only; I mean to refer to the fact that those trusts are for purposes that are charitable purposes except for the fact that they entail discrimination. Thus, I leave open the possibility that such trusts, once the discriminatory character of their purposes is brought into view, might turn out not to be charitable all things considered. Moreover, I do not enter into debates about whether the appropriate response to discriminatory charitable trusts is to declare them invalid or to vary their terms *cy-près*: instead, I pose questions at a higher level of generality about whether such trusts ought to be "interfered" with in one or another way.

1. I do attempt such arguments elsewhere: see Matthew Harding, *Charity Law and the Liberal State* (Cambridge: Cambridge University Press, 2014) ch 7 [Harding, *Charity Law*].

II. Public Policy

The most prominent theme in equitable responses to discriminatory charitable trusts is the theme of public policy. In cases about discriminatory charitable trusts where judges deploy this theme in their reasoning, they typically seek to ascertain whether or not the trust in question offends the doctrine according to which dispositions may be interfered with, including struck down altogether, on grounds of public policy. An investigation into public policy in relation to a discriminatory charitable trust usually entails some assessment of the extent to which equality norms inform public policy, along with an effort to balance such norms against the freedom of disposition of the settlor of the trust. In this part, I consider what the case law reveals about how judges work with the theme of public policy when dealing with discriminatory charitable trusts, before undertaking a critical analysis of that case law.

Cases in which judges have been asked to interfere with discriminatory charitable trusts on public policy grounds are not numerous. Nonetheless, such cases tend to support the proposition that, traditionally at least, judges have been reluctant to invoke the public policy doctrine against discriminatory charitable trusts. In England, although judges have shown themselves willing to order *cy-près* variation of discriminatory charitable trusts,[2] they have never made such orders explicitly on public policy grounds. For example, in *Re Lysaght*, a testamentary gift was made to the Royal College of Surgeons for the purpose of providing scholarships to medical students; its terms discriminated against Roman Catholic and Jewish students.[3] Justice Buckley said that it would be "going much too far" to say that the trust was contrary to public policy,[4] but he nonetheless approved a *cy-près* scheme excising the discriminatory terms because the testatrix's intention was that the trust be administered by the College, and the College would not accept the gift unless the trust was rendered non-

2. See also *In Re Harding*, [2007] EWHC 3 (Ch)(where a discriminatory charitable trust was varied on statutory grounds).
3. [1966] Ch 191 (Eng) [*Re Lysaght*].
4. *Ibid* at 206.

discriminatory in relation to religion.[5] In Australia too, a recent decision of the Supreme Court of New South Wales also reveals judicial reluctance to deploy the public policy tool against a discriminatory charitable trust. In *Kay v South Eastern Sydney Area Health Service*[6] ("*Kay*"), a testamentary gift for the treatment of "White babies" was upheld,[7] Chief Justice Young in equity saying only that "generally speaking testators can be as capricious as they like and … if they wish to benefit a charity in respect of, or [sic] even of, a discriminatory group, they are at liberty to do so".[8] These cases reveal judges balancing freedom of disposition and equality in the setting of the public policy doctrine by assuming that freedom of disposition outweighs equality.

The traditional reluctance of judges to interfere with discriminatory charitable trusts on public policy grounds has been accompanied by an even more pronounced judicial reluctance to invoke the public policy doctrine against discriminatory "private" trusts. There is, as is well known, a long tradition of judges interfering with dispositions for reasons of public policy.[9] At the same time, there is a long, if poorly understood, tradition of judges recognising equality norms within the public policy doctrine.[10] Nonetheless, in cases of "private" trusts, viz., trusts whose objects are persons identified by name or ascertainable by reference to a described class, courts have traditionally refused to invoke public policy to respond

5. *Ibid* at 209. See also *Re Dominion Students' Hall Trust*, [1947] 1 Ch 183 (Eng); *Re Meres' Will Trusts* (1957)(Ch (Eng)), cited in "Law Report, May 3", *The Times* (3 May 1957) 10.
6. [2003] NSWSC 292 (Austl).
7. *Ibid* at para 2 (in the will, the testatrix had underlined the word "White" twice).
8. *Ibid* at para 18. Compare *Home for Incurables of Baltimore City v University of Maryland Medical System Corporation*, 797 A (2d) 746 (Md Ct App 2002 (US))(and thanks to Evelyn Brody for bringing that case to my attention).
9. I note relevant cases in Matthew Harding, "Some Arguments against Discriminatory Gifts and Trusts" (2011) 31:2 Oxford Journal of Legal Studies 303 at 310-11 [Harding, "Some Arguments"].
10. See Peter Benson, "Equality of Opportunity and Private Law" in Daniel Friedman & Daphne Barak-Erez, eds, *Human Rights in Private Law* (Portland: Hart Publishing, 2001) 201 at 209.

to any discrimination entailed in the terms of the trusts in question.[11] In withholding the public policy tool in such cases, courts have resolved a perceived competition between freedom of disposition and equality by finding that freedom of disposition prevails. A good illustration of this is *Blathwayt v Baron Cawley*.[12] There, a testamentary disposition in terms that discriminated against Roman Catholics was upheld notwithstanding its discriminatory character. In upholding the disposition, members of the House of Lords noted the equality interest of the class affected by the discrimination, and acknowledged that equality norms informed public policy in England. However, their Lordships thought that, on the facts of the case, any such equality norms were clearly outweighed by the testator's freedom of disposition.[13] For Lord Wilberforce, "neither by express provision nor by implication has private selection yet become a matter of public policy".[14]

The traditional judicial tolerance towards discriminatory charitable and "private" trusts has been unsettled in recent decades in two jurisdictions: Canada and South Africa. In Canada, the key case, decided in 1990, is *Canada Trust Co v Ontario (Human Rights Commission)*[15] ("*Canada Trust*"). The trust in question, the Leonard Foundation, was settled by a prominent Canadian in the 1920s to fund educational scholarships; the recitals and provisions of the trust deed made clear, in unmistakably bigoted terms, that the scholarships were not to be awarded except to white Protestants of British nationality or "parentage". The trust deed also stipulated that no more than a quarter of available funds should be paid each year to female candidates, and it contained terms that discriminated against other candidates on grounds of parental

11. Judges have been more willing to interfere with such trusts where their discriminatory terms take the form of uncertain conditions: see Harding, "Some Arguments", *supra* note 9 at 307-10 for a discussion of relevant cases.
12. [1976] AC 397 (HL) [*Blathwayt*]. I discuss other illustrative cases in Harding, "Some Arguments", *supra* note 9 at 304-305.
13. *Blathwayt*, *supra* note 11 at 425-26, per Lord Wilberforce; 429, per Lord Cross; 441, per Lord Edmund-Davies.
14. *Ibid* at 426.
15. (1990), 69 DLR (4th) 321 (Ont CA) [*Canada Trust*].

occupation.[16] In response to public pressure, the trustee of the Leonard Foundation applied for judicial directions as to the validity of the trust. At first instance, the trust was found not to offend public policy,[17] but the Ontario Court of Appeal ordered a *cy-près* scheme excising the terms of the trust deed that discriminated on grounds of race, sex, nationality, and religion.[18] In that decision, all the members of the Court thought that the terms of the trust offended public policy and should be varied on that basis.

A majority of the Court framed the relevant question for decision as one that demanded a balancing of freedom of disposition and equality via the public policy doctrine. Justice Robins noted the significance of freedom of disposition in Canadian law,[19] but he went on to state that equality norms inform Canadian public policy in important ways, in light of the diverse character of Canadian society and the nation's constitutional commitments.[20] For Robins JA:

> [t]he settlor's freedom to dispose of his property through the creation of a charitable trust fashioned along these lines must give way to current principles of public policy under which all races and religions are to be treated on a footing of equality and accorded equal regard and equal respect.[21]

In this and other passages[22] Robins JA seemed to say that equality may outweigh freedom of disposition not only in the case of a discriminatory charitable trust, such as the Leonard Foundation, but also in the case

16. The relevant provisions of the trust deed are set out *ibid* at 326-29. Eligibility to participate in the management of the trust was also restricted based on race, nationality and religion.
17. *Canada Trust Co v Ontario (Human Rights Commission)*(1987), 61 OR (2d) 75 (H Ct J).
18. The provisions that discriminated on grounds of parental occupation were left undisturbed.
19. *Canada Trust*, *supra* note 15 at 334.
20. *Ibid* at 334-35.
21. *Ibid* at 335.
22. See also *ibid* ("[t]he freedom of an owner of property to dispose of his or her property as he or she chooses is an important social interest that has long been recognised in our society and is firmly rooted in our law … That interest must, however, be limited in the case of this trust by public policy considerations" at 334).

of a discriminatory "private" trust. The majority judgment in *Canada Trust* may thus be read as a radical departure from the traditional judicial reluctance to invoke public policy against discriminatory "private" trusts, and there is evidence that Canadian courts have read the judgment in this way since *Canada Trust* was decided.[23]

In his concurring judgment in *Canada Trust*, Justice Tarnopolsky joined with the majority in conceiving of the question for decision as one that required a balancing of freedom of disposition and equality via the public policy doctrine. Like the majority, Tarnopolsky JA ruled that Canadian public policy entailed equality norms;[24] like the majority, he also acknowledged the importance of freedom of disposition to the law:[25]

> [i]n this case the court must, as it does in so many areas of law, engage in a balancing process. Important as it is to permit individuals to dispose of their property as they see fit, it cannot be an absolute right. The law imposes restrictions on freedom of both contract and testamentary disposition.[26]

However, while Tarnopolsky JA agreed with the majority that a balancing exercise was necessary, he seems to have disagreed with the majority on the right way to balance freedom of disposition and equality in cases of discriminatory "private" trusts. He stated that:

> [t]his decision does not affect private, family trusts. By that I mean that it does not affect testamentary dispositions or outright gifts that are not also charitable trusts. Historically, charitable trusts have received special protection … This preferential treatment is justified on the ground that charitable trusts are dedicated to the benefit of the community … It is the public nature of charitable trusts which attracts the requirement that they conform to the public policy against discrimination.[27]

---

23. See *McCorkill v Streed*, 2014 NBQB 148; *Spence v BMO Trust Company*, 2015 ONSC 615. Perhaps ironically, Canadian courts have proven less willing to interfere with discriminatory *charitable* trusts on public policy grounds in the years since *Canada Trust* was decided: see *Re Ramsden Estate* (1996), 139 DLR (4th) 746 (PESC (TD)) [*Re Ramsden Estate*]; *University of Victoria v British Columbia (AG)*, 2000 BCSC 445 [*University of Victoria*]; *Re The Esther G Castanera Scholarship Fund*, 2015 MBQB 28 [*Castanera*]. I discuss the latter three cases below.
24. *Canada Trust, supra* note 15 at 348-52.
25. *Ibid* at 353.
26. *Ibid*.
27. *Ibid*.

For Tarnopolsky JA, it was because charitable trusts are in some relevant sense "public" trusts that freedom of disposition and equality in cases of discriminatory charitable trusts are to be balanced in favour of equality.[28]

Also at odds with the traditional judicial reluctance to invoke public policy against discriminatory trusts is recent South African jurisprudence. While these South African developments are not, strictly speaking, developments in equity — South Africa has no tradition of equity — they are of obvious relevance for those jurisdictions where discriminatory trusts fall to be considered by judges exercising equitable jurisdiction. Prior to South Africa's current constitutional settlement, South African courts were loath to interfere with such trusts on public policy, or "*boni mores*", grounds;[29] as François du Toit points out in his important work on the subject, there are reasons to think that historically freedom of disposition has been prized in South Africa even more than in the common law world.[30] However, all that changed with the coming into effect of the *Constitution of the Republic of South Africa, 1996*.[31] According to section 9(4) of the Constitution of South Africa, constitutionally protected equality rights may be enforced by citizens against each other

28. See further Harding, *Charity Law*, *supra* note 1 at 215-16.
29. Although see *Ex parte President of the Conference of the Methodist Church of Southern Africa NO: In re William Marsh Will Trust*, [1993] 2 SA 697 (Cape Prov Div)(decided under s 13 of the *Trust Property Control Act, 1988* (SA) No 57 of 1988 [*Trust Property Control Act*]). The section "empowers a court to vary any trust provision where such provision occasions consequences which, in the opinion of the court, the trust founder failed to contemplate or foresee … and such provision is, *inter alia*, in conflict with the public interest": François du Toit, "Constitutionalism, Public Policy and Discriminatory Testamentary Bequests; A Good Fit between Common Law and Civil Law in South Africa's Mixed Jurisdiction" (2012) 27 Tulane European and Civil Law Forum 97 at 111 [du Toit, "A Good Fit"]. The inquiry into "public interest" for the purposes of s 13 is similar to the inquiry under the public policy doctrine.
30. du Toit, "A Good Fit", *ibid* at 114-16.
31. No 108 of 1996 [*Constitution of South Africa*].

as well as against the state;[32] this provision led commentators, including du Toit, to argue in the years following the new constitutional settlement that some previously uncontroversial discriminatory dispositions would no longer survive the scrutiny of South African courts.[33] And those predictions have indeed been borne out in the post-1996 jurisprudence; South African courts are now willing to interfere with discriminatory trusts on public policy grounds.[34] At the same time though, freedom of disposition, at least in a testamentary setting, has been placed on a constitutional footing as well, a matter to which I return below.

In South Africa, the first post-1996 case to deal with a discriminatory trust was *Minister for Education v Syfrets Trust Ltd NO*[35] ("*Minister for Education*"). There, the terms of a testamentary trust for the purpose of funding educational scholarships discriminated against non-Europeans, Jews and women. Justice Griesel of the High Court of South Africa was asked to order that the offending provisions of the will be deleted, so that the trust could be administered in a non-discriminatory fashion. Justice Griesel granted the order. Rather than dealing with the case as one requiring an enforcement of the equality rights set out in section 9(4) of the Constitution,[36] Griesel J applied the public policy doctrine.[37] He spelled out the ways in which the equality norms enshrined in the Constitution now informed and gave content to public policy in South

32. See also *Promotion of Equality and Prevention of Unfair Discrimination Act* (SA), Act 4 of 2000.
33. See François du Toit, "The Constitutionally Bound Dead Hand? The Impact of Constitutional Rights and Principles on Freedom of Testation in South African Law" (2001) 12:2 Stellenbosch Law Review 222, and the sources cited therein.
34. And recently, it seems, via direct application of the equality provisions of the Constitution: see Fatima Schroeder, "Whites-only Bursaries to be Scrapped", *Iol News* (25 April 2015), online: iol news <www.iol.co.za/news/south-africa>. I am grateful to Marius de Waal for alerting me to this case; written reasons for the decision had not been published when this article went to press.
35. *Minister for Education v Syfrets Trust Ltd NO*, [2006] ZAWCHC 65 (SA) [*Minister for Education*].
36. As he had been invited to do: *ibid* at para 9.
37. *Ibid* at para 16.

Africa.[38] Justice Griesel was of the view that those public policy-informing equality norms outweighed the freedom of disposition of the testator, in light of the fundamental nature of the commitment of the South African polity to equality in the Constitution.[39] This approach was also taken in *Curators Ad Litem to Certain Beneficiaries of Emma Smith Educational Fund v The University of KwaZulu-Natal*,[40] in which the Supreme Court of Appeal appealed to "the fundamental values of our Constitution and the constitutional imperative to move away from our racially divided past" in upholding an order varying the terms of another discriminatory trust for funding educational scholarships.[41]

Another recent South African case to deal with a discriminatory trust is *BoE Trust Limited NO*[42] ("*BoE Trust*"). The terms of a testamentary trust, once again for the purpose of funding educational scholarships, discriminated against students who were not "White". The testatrix had also made provision in the will that, "[i]n the event that it should become impossible for my trustee[s] to carry out the terms of the trust", the income of the trust should be paid to certain named charities.[43] The terms of the educational trust required that professors from four named universities participate in its management; the universities in question refused to participate as long as the trust contained a racially discriminatory provision.[44] The trustees therefore sought an order deleting the provision in question. The Supreme Court of Appeal refused to grant the order. In doing so, the Court ruled that freedom of testation is the subject of a right protected by section 25(1) of the South African Constitution,[45] a proposition that Griesel J had assumed to be correct in

38. *Ibid* at paras 23-32.
39. *Ibid* at paras 39-46. Also relevant was that the trust was to be administered by a public body, viz., a university, *ibid* at para 45.
40. [2010] ZASCA 136 (SA).
41. *Ibid* at para 42. The order had been made under section 13 of the *Trust Property Control Act*, *supra* note 29.
42. [2012] ZASCA 147 (SA) [*BoE Trust*].
43. *Ibid* at para 3.
44. *Ibid* at paras 3, 7-9.
45. *Ibid* at paras 26-27.

Minister for Education but had not ruled on.[46] Significantly, the Court also linked freedom of testation to " … the founding constitutional principle of human dignity. The right to dignity allows the living, and the dying, the peace of mind of knowing that their last wishes would be respected after they have passed away".[47]

Immediately prior to affirming that the right to freedom of testation is protected by the Constitution, the Court issued this statement:

> [t]he giving of the bursaries as [the testatrix] had intended had become impossible as a result of the universities' stance. Must the alternative provided in the will be given effect to? Does [the testatrix's] right to dispose of her assets as she saw fit, whether we agree with her exercise of that right or not, require a court to see at least whether there is a way in which to interpret her will so as that it does not offend public policy?[48]

Having posed those questions, the Court answered them by giving effect to the provision of the will that was conditional on the educational trust being impossible to carry out.[49] Rather than resolving the competition between freedom of disposition and equality that was raised by the educational trust, the Court, thanks to the provisions of the will, was able to sidestep that competition altogether. In one sense, for the Court to have given effect to the terms of the will in that way seems unremarkable; at the same time though, there are reasons to worry about the Court's emphasis on freedom of testation, and I return to these shortly.

As the case law shows, the public policy doctrine is a useful tool for judges who seek to balance the demands of freedom of disposition and equality in responding to discriminatory trusts. The doctrine is an especially useful tool for judges in jurisdictions where equality norms figure in the Constitution, as the South African jurisprudence amply demonstrates; through the doctrine, judges may draw on constitutional equality norms in giving content to equity according to orthodox methods of judicial reasoning.[50] That said, invocation of the theme of public

46. *Minister for Education*, *supra* note 35 at para 18.
47. *BoE Trust*, *supra* note 42 at para 27.
48. *Ibid* at para 25.
49. *Ibid* at paras 30-31.
50. For fuller discussion, see Harding, "Some Arguments", *supra* note 9 at 310-16.

policy in cases of discriminatory trusts raises questions that demand careful scrutiny, and these questions have not, to date, received sufficient answers in the case law. Perhaps the most pressing of these questions might be stated as follows: does public policy demand that freedom of disposition and equality be balanced in respect of discriminatory charitable trusts differently than in respect of discriminatory "private" trusts? According to the traditional view, the answer is no: public policy demands that freedom of disposition should prevail in respect of both types of trust. According to the majority in *Canada Trust* and, it would seem, recent South African jurisprudence, the answer is also no, but in a different way: public policy demands that equality norms prevail over freedom of disposition in respect of both charitable and "private" trusts that discriminate.[51] According to Tarnopolsky JA in *Canada Trust*, on the other hand, the answer is yes: charitable trusts, because they are in some relevant sense "public" trusts, are susceptible to equality norms in ways that "private" trusts are not. Which view is to be preferred?

One theoretical effort to address this question is to be found in the work of Lorraine Weinrib and Ernest Weinrib.[52] The Weinribs begin with the proposition that certain values underpin and animate the entire legal order, including private law, of which the law of trusts and equity more generally are a part.[53] The Weinribs, writing in a Canadian setting, locate these values in the written constitution,[54] but nothing in the Weinribs' analysis precludes the possibility that the values might emerge from judge-made law. The important point is that the values are foundational and affect private law. In the Weinribs' view, in a jurisdiction where

51. The South African cases discussed above were about discriminatory charitable trusts. However, their reasoning seems clearly applicable to discriminatory "private" trusts as well.
52. Lorraine E Weinrib & Ernest J Weinrib, "Constitutional Values and Private Law in Canada" in Daniel Friedman & Daphne Barak-Erez, eds, *Human Rights and Private Law* (Portland: Hart Publishing, 2001) 43 [Weinrib & Weinrib, "Constitutional Values"].
53. *Ibid* at 50-51.
54. This is in keeping with Canadian jurisprudence: see *Retail, Wholesale and Department Store Union, Local 580 v Dolphin Delivery Ltd*, [1986] 2 SCR 573; *Hill v Church of Scientology*, [1995] 2 SCR 1130.

such foundational values include equality norms, those equality norms affect private law, and private law must be appropriately sensitive to them. At the same time though, the Weinribs point out that private law must be sensitive to freedom of disposition, and it may be assumed that they regard this freedom also as a fundamental value underpinning the legal system.[55] Again, this value might have recognition in a written constitution — as is the case in South Africa with regard to freedom of testation, as we saw earlier — or it might emerge from judge-made law; again, the important point is that the value of freedom of disposition is fundamental and affects private law. In a jurisdiction where equality norms and freedom of disposition constitute fundamental values of the legal system, the Weinribs appear to think that the appropriate legal response to discriminatory trusts demands that the right balance be struck between those values.[56]

The Weinribs argue that such a balancing exercise must be carried out according to what they call a "proportionality" principle;[57] they describe the principle in the following passage:

> [a] central aspect of one normative principle is granted priority over a comparatively more marginal aspect of another. Relevant to this exercise would be a comparison of the principles that favour the two parties, in which one asks whether the triumph of the plaintiff's principle would impact more heavily on the defendant's than the triumph of the defendant's principle would impact on the plaintiff's.[58]

According to the Weinribs, when applied to discriminatory "private" trusts, this principle demands that freedom of disposition prevail. The Weinribs argue that for such a trust to be interfered with owing to its discriminatory character would be to disturb freedom of disposition in a disproportionate way; after all, they say, freedom of disposition in a "private" trust is precisely the freedom to choose between different possible beneficiaries in ways that discriminate between them.[59] For the Weinribs, matters are otherwise in the case of a discriminatory charitable

55. Weinrib & Weinrib, "Constitutional Values", *supra* note 52 at 68.
56. *Ibid* at 57-59.
57. *Ibid* at 57.
58. *Ibid* at 58.
59. *Ibid* at 68.

trust. In such a case, freedom of disposition is exercised to give effect to a public benefit purpose, and, to that extent, to subject that exercise of dispositive freedom to equality norms that operate in the public sphere is consistent with the choices entailed in it.[60]

If the Weinribs' view of how freedom of disposition and equality should be balanced in cases of discriminatory trusts is sound, then Tarnopolsky JA seems to have been on the right track when, in *Canada Trust*, he drew a distinction between charitable and "private" trusts that discriminate and found that only trusts of the former type should be interfered with on public policy grounds. But embedded in the Weinribs' view is a contestable understanding of the value of freedom of disposition that should be exposed and scrutinised before the view is accepted as sound. The Weinribs seem to assume that freedom of disposition is valuable because and to the extent that the disponor chooses the objects of her disposition, whether those objects be persons or purposes. On this view, the meanings and consequences of the disponor's choice, along with the identity and character of objects of that choice — what may compendiously, if somewhat tendentiously, be called the expressive and teleological aspects of the choice — are of no relevance to understanding the sense in which the freedom to make the choice is valuable. It follows that such expressive and teleological aspects of disponors' choices are irrelevant to understanding the value of freedom of disposition in cases of discriminatory trusts. Among the irrelevant considerations might be that the expression of a disponor's choice demeans some identity-based group, or that a historically disadvantaged class is placed at a relative disadvantage by the choice in question.

The Weinribs' assumption that the value of freedom of disposition inheres in choice itself is consistent with a particular view of the moral practice of private law, a view that Ernest Weinrib has explored more fully in his other work.[61] According to that view, expressive and teleological considerations ought to play no role in private law. Rather, private law should be confined to the public enforcement of the demands of right

60. *Ibid*.

61. Ernest J Weinrib, *The Idea of Private Law*, 2d (Oxford: Oxford University Press, 2012) [Weinrib, *Private Law*].

understood according to the moral theory of Immanuel Kant;[62] from a Kantian perspective, the demands of right are demands that a legal framework enable each person to act freely and purposively in the world consistent with the free and purposive action of each other person.[63] The demands of right are therefore agnostic as to expressive and teleological dimensions of free and purposive choice; it is free and purposive choice itself that matters. And it follows that, within the moral practice of private law, choice may be treated as valuable in itself, irrespective of its expressive or teleological dimensions, so long as it is consistent with the demands of Kantian right.[64] Thus the legal system may, indeed should, constrain choices that dominate or coerce others, but there is no moral requirement that the legal system constrain choices that discriminate against others on grounds relating to their identity.

The Weinribs' assumption that the value of freedom of disposition inheres in choice itself is not self-evident; neither is the Kantian account of the moral practice of private law, which animates that assumption. Thus, the Weinribs should provide some argument to support the proposition that private law is a moral practice confined to giving effect to Kantian right; and in the absence of such an argument, the Weinribs' contestable assumption about the value of freedom of disposition should not be accepted. It has been well explained elsewhere why the arguments that Ernest Weinrib provides for his interpretation of private law are

62. In their jointly authored piece, the Weinribs talk about the principle of "transactional equality"; this principle appears to refer to the demands of Kantian right: Weinrib & Weinrib, "Constitutional Values", *supra* note 52 at 58 and *passim*.

63. On the demands of Kantian right, see also Arthur Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy* (Cambridge: Harvard University Press, 2009).

64. I say "treated as valuable" because, in a sense, from the perspective of Kantian right questions of value are altogether irrelevant to the question whether the law should facilitate or constrain freedom of disposition: Weinrib, *Private Law*, *supra* note 61 at 109-13. From this perspective, to ask about the "value" of freedom of disposition is to ask the wrong question. Nonetheless, for ease of expression I refer in the text to freedom of disposition as a "value" even when referring to the Weinribs' ideas.

unconvincing.[65] Moreover, there are plausible alternative accounts of private law according to which private law should take an interest in the expressive and teleological dimensions of choices, even though those choices are consistent with the demands of Kantian right.[66] I therefore want to proceed by insisting that the value of freedom of disposition may depend in important ways on just the sorts of considerations — relating to the expressive and teleological dimensions of the choices entailed in exercises of that freedom — that the Weinribs rule out of play in their treatment of that value. In particular, some exercises of freedom of disposition might turn out to lack value in certain ways where they entail choices to settle trusts on discriminatory terms. And in extreme cases of such value-lacking exercises of dispositive freedom, a balancing of the values of freedom of disposition and equality might turn out to be inappropriate, because there might be nothing of value to place on the scales on the "freedom of disposition" side. In those extreme cases, it might be misleading to say that discriminatory trusts offend public policy in spite of the value of freedom of disposition; it might be more accurate to say that discriminatory trusts both offend public policy and are products of valueless exercises of dispositive freedom.

To illustrate these points about extreme cases, consider an example of a discriminatory disposition from beyond the law of trusts: the discriminatory restrictive covenant in the Canadian case of *Re Noble and Wolf*.[67] The covenant in question purported to prohibit the sale of land to "any person of the Jewish, Hebrew, Semitic, Negro or coloured race or blood".[68] Was the exercise of freedom of disposition entailed in binding a purchaser of land to this covenant a valuable one? The Weinribs seem

65. See *e.g.* Hanoch Dagan, "The Limited Autonomy of Private Law" (2008) 56:3 American Journal of Comparative Law 809.
66. For example, accounts informed by higher order teleological accounts of political morality: for two such higher order accounts see John Finnis, *Natural Law and Natural Rights*, 2d (Oxford: Oxford University Press, 2011); Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986).
67. [1949] OR 503 (CA) [*Re Noble and Wolf*].
68. *Ibid* at 513.

to suggest not in their discussion of the case.[69] And yet according to the Weinribs' own understanding of the value of freedom of disposition, it seems that a choice to dispose of land on the terms of the covenant in *Re Noble and Wolf* may be treated in law as valuable because it is consistent with the demands of Kantian right; to that extent, the choice may be weighed in the balance against equality norms in working out whether or not to enforce the covenant. Moreover, according to the Weinribs' understanding of the value of dispositive freedom, it seems that the balancing exercise should be determined by giving effect to the disponor's choice because any other result would interfere disproportionately with the value of freedom of disposition. This reflects in substance the position taken by the Ontario Court of Appeal, members of which appealed to the "sanctity" or "liberty" of contract in upholding the covenant in question.[70] But the Weinribs' reaction to the case seems the right one notwithstanding their theoretical commitments, once it is accepted that the expressive and teleological dimensions of choices may be taken into account in assessing their value. On this view, that the covenant expressed a bigoted contempt for certain religious and racial groups, and that it was directed at excluding members of those groups from living in a particular residential community, combine to suggest that the covenant was the product of a valueless exercise of dispositive freedom and one that merited no legal protection or concern.[71]

Extreme cases like *Re Noble and Wolf* show that where judges share the Weinribs' contestable assumption about the value of freedom of

69. Weinrib & Weinrib, "Constitutional Values", *supra* note 52 at 63.

70. *Re Noble and Wolf*, *supra* note 67 at 524 per Justice Henderson, 530 per Justice Hogg. The case went on appeal to the Supreme Court of Canada, where the covenant was struck down, but not on public policy grounds: *Noble v Alley*, [1951] SCR 64.

71. From this perspective, it is a matter for regret that in *Re Noble and Wolf* the Court departed from the earlier decision of Justice Mackay of the Ontario High Court in *Re Drummond Wren*, [1945] OR 778 (HC), striking down a similar discriminatory restrictive covenant on public policy grounds. In the law of Ontario today, discriminatory restrictive covenants are void under section 22 of the *Conveyancing and Law of Property Act*, RSO 1990, c C-34.

disposition, they may attribute weight to the value of freedom of disposition in the setting of the public policy doctrine even though that value should be given no weight in the circumstances. And, of particular relevance to this paper, the possibility of such extreme cases calls into question the willingness of courts to attribute weight to freedom of disposition as well as equality when applying the public policy doctrine in cases about discriminatory trusts. Courts have shown this willingness not only in jurisdictions like England and Australia where such trusts have been upheld, but also in jurisdictions like Canada and South Africa where such trusts have been interfered with on public policy grounds. For example, recall the emphasis on freedom of testation in the South African case of *BoE Trust* and the unquestioning association of that value with the fundamental constitutional value of human dignity. This emphasis seems at odds with the proposition that, in some cases, a bigoted and offensive exercise of freedom of testation might be utterly without value and deserving of no legal protection or concern. In such cases, a South African court might do better to assert that the exercise of testamentary freedom is valueless and invoke the public policy rule against it on that basis.

At this point, it must be noticed that extreme cases like *Re Noble and Wolf* are rare. To begin with, it is unusual for a discriminatory disposition to be discriminatory in a way that ought to be of concern to the law. As the Weinribs point out, freedom of disposition just is the freedom to discriminate in choosing objects of the disposition, and most of the time such discrimination — for example, the choice of one child over another as heir, or the choice of a "relief of poverty" purpose over an "advancement of education" one — is and ought to be unremarkable from a legal point of view.[72] Moreover, even in cases where dispositions do entail the sort of discrimination in which the law should take an interest, that discrimination is often plausibly construed as a means to bringing

72. Although, in the case of the choice of one child over another as heir, see *Spence, supra* note 23.

about a valuable state of affairs.[73] And nowhere is this clearer than in the case of discriminatory charitable trusts. Such trusts are usually oriented to a purpose of a type recognised as *prima facie* charitable in law, but the class of persons who stand to benefit directly from that purpose being carried out is restricted in some way. That discriminatory charitable trusts typically have this profile may be illustrated even by cases of such trusts that discriminate in an egregious way, such as the trusts in *Canada Trust* and *Kay.* The most plausible construction of the terms of those trusts is that they were for "advancement of education" and "advancement of health" purposes respectively, rather than for the purpose of discriminating against the classes affected by their discriminatory terms.[74]

Where discriminatory trusts are the products of exercises of dispositive freedom that are valuable to some degree even though they entail discrimination, there is something to weigh on the scales on the "freedom of disposition" side when assessing whether such trusts offend public policy. This leads us back to a principle approximating the Weinribs' principle of "proportionality", according to which freedom of disposition and equality are to be balanced against each other. But if we reject the Weinribs' contestable assumption about the value of freedom of disposition, the application of the principle of "proportionality"

73. Thus, the testamentary disposition in *Blathwayt* may – and I stress, *may* – have had more to do with the testator's loyal pride in the Protestantism of his aristocratic family than with bigoted prejudice against Roman Catholics: *Blathwayt*, *supra* note 12 at 426 per Lord Wilberforce, 429 per Lord Chelsea.

74. For a different view, see Adam Parachin, "Public Benefit, Discrimination and the Definition of Charity" in Kit Barker & Darryn Jensen, eds, *Private Law: Key Encounters with Public Law* (Cambridge: Cambridge University Press, 2013) 171 at 199-200 [Parachin, "Public Benefit"]. Of course, a trust might be created for the purpose of discriminating against some group; a trust for the purpose of "promoting the subordination of women" might be offered as an example. Such a trust could not be called "charitable" even in a provisional sense; it would clearly fail the public benefit test on any reasonable view of that test. But attempts to create trusts for the purpose of discrimination, as opposed to trusts for purposes that entail discrimination, are not known to the case law and may be assumed to be exceedingly rare.

must be sensitive to the fact that, in cases of discriminatory trusts, the proper balance between freedom of disposition and equality cannot be determined except with reference to the expressive and teleological dimensions of choices entailed in exercises of dispositive freedom. And this sensitivity is called for whether the balancing exercise is required in the setting of a charitable or a "private" trust. In a legal landscape in which the "proportionality" principle was applied in this way, a discriminatory charitable trust would not be susceptible to interference on public policy grounds solely because it was in some sense "public"; it would be susceptible to such interference to the extent that the exercise of dispositive freedom that brought it into existence lacked value. And a discriminatory "private" trust would be susceptible to interference for the same reason.

If the value of freedom of disposition is a function of expressive and teleological dimensions of the choices entailed in it, then there is work to be done developing the public policy doctrine so that it is more sensitive to that value. In *Canada Trust*, the majority barely engaged with normative considerations bearing on the value of dispositive freedom, simply stating that freedom of disposition was a value, that it must "give way" to equality norms on the facts of the case, and that to assert this was to "expatiate the obvious".[75] In particular, no real effort was made to explain why equality norms should prevail in *Canada Trust* despite the fact that the Leonard Foundation was a trust for the advancement of education and, to that extent, the product of a valuable exercise of dispositive freedom.[76] And judges in subsequent Canadian cases about discriminatory trusts for the advancement of education have also shown little appetite to engage with the value of freedom of disposition.[77] In *Re Ramsden Estate*, a case about a trust to provide scholarships to Protestant students, *Canada Trust*

75. *Canada Trust*, *supra* note 15 at 321, 334.
76. See *ibid* at 333, per Robins JA (for the view that the trust was for the advancement of education). The criticism in the text may also be made of the decision of the US Supreme Court in *Bob Jones University v US*, 461 US 574 (1983) [*Bob Jones*], and indeed was made by Chief Justice Rehnquist in his dissenting judgment in that case.
77. See the discussion in Parachin, "Public Benefit", *supra* note 74 at 178-79.

was distinguished on the basis that the discrimination in *Canada Trust* was "blatant".[78] There was no attempt to investigate whether the value of freedom of disposition was to be understood differently in the two cases because of the character of the discrimination in each. Similarly, in *University of Victoria v British Columbia (AG)*[79] ("*University of Victoria*"), a case about a trust for funding scholarships for Roman Catholic students, *Canada Trust* was distinguished because the discrimination in *Canada Trust* was thought to be "offensive" whereas the discrimination in the case at hand was not;[80] but again the court declined to consider whether and why the value of freedom of disposition might be affected in some way by the fact that such freedom is exercised in an "offensive" fashion.

Re The Esther G Castanera Scholarship Fund[81] ("*Castanera*") is the most recent Canadian case raising a discriminatory charitable trust. In that case, a testamentary trust to provide financial support to women graduates of a particular high school undertaking tertiary studies in the sciences was in view.[82] Unlike the judges in *Re Ramsden Estate* and *University of Victoria*, in *Castanera*, Justice Dewar of the Court of Queen's Bench of Manitoba undertook a careful analysis of normative considerations bearing on the public policy doctrine as it applies to discriminatory charitable trusts. In this regard, *Castanera* represents an improvement on the earlier cases. That said, Dewar J seems to have thought that the motivations underpinning a discriminatory charitable trust are of normative significance in applying the public policy doctrine; thus, he drew attention to the bigoted motivations underpinning the discriminatory trust in *Canada Trust*, and he contrasted the benign motivations of the testatrix in the case at hand.[83] But Dewar J seems to have been searching in the wrong place for normative considerations bearing on freedom of disposition in cases of discriminatory charitable trusts; after all, as a general rule motivations are irrelevant to questions

78. *Re Ramsden Estate*, *supra* note 23 at para 13.
79. *University of Victoria*, *supra* note 23.
80. *Ibid* at para 25.
81. *Castanera*, *supra* note 23.
82. *Ibid*.
83. *Ibid* at para 37.

relating to the construction and validity of trusts. Arguably, Canadian courts have some way to go before they identify those expressive and teleological considerations that truly bear on discriminatory charitable trusts.

III. Public Benefit

I ended Part II by suggesting that judges may need to develop the public policy doctrine in cases of discriminatory trusts so as to respond better to normative considerations that bear on the value of freedom of disposition. However, in cases of discriminatory *charitable* trusts, this need may not be an urgent one. In such cases, judges have at hand another theme — arising from equity's historical jurisdiction over charity — that might prove capable of delivering nuanced responses without at the same time demanding that judges reflect on the value of freedom of disposition. This is the theme of public benefit. It is, of course, well established in equity that a charitable trust must have a dominant or primary purpose that is, in some sense, a public benefit purpose. The jurisprudence on this public benefit requirement is large and complex, and in contemporary equity it may fairly be said that public benefit is the central organising idea in cases about charity.[84] Thus, it is strange that the theme of public benefit is hardly ever deployed in judicial responses to discriminatory charitable trusts; as I discussed above, in the relatively few cases about such trusts, invocation of the theme of public policy is the norm. Equally, the theme of public benefit seems not to be prominent, even if it is present, in cases about charities that engage in discrimination in the provision of services.[85] In this part, I will not speculate as to why the public benefit

84. The most comprehensive contemporary treatment of the public benefit test is Jonathan Garton, *Public Benefit in Charity Law* (Oxford: Oxford University Press, 2013) [Garton, *Public Benefit*].

85. Thus, there are references to public benefit in *Bob Jones*, *supra* note 76; *Catholic Care (Diocese of Leeds) v Charity Commission for England and Wales*, [2010] EWHC 520 (Ch); and *St Margaret's Children and Family Care Society v Office of the Scottish Charity Regulator* (2014), App 02/13 (Scottish Charity Appeals Panel) [*St Margaret's*] (but ultimately each of these cases can be understood as having been decided on grounds other than the public benefit test).

test has not been utilised by judges dealing with discriminatory charitable trusts. Rather, my aim is to offer some thoughts as to how the theme of public benefit might be developed by judges so as to respond to such trusts. I argue that the public benefit test contains resources on which judges may build in reasoning about public benefits and detriments associated with discriminatory charitable trusts. In particular, I suggest that public benefit jurisprudence might be developed to enable judges, in appropriate cases, to find that discriminatory charitable trusts are not of public benefit and should be interfered with on that basis.

The public benefit test has two components.[86] First, in order to satisfy the test, a purpose must stand to benefit a class of persons that is sufficiently "public" in character. This component of the test is directed against trusts that stand to benefit classes of persons whose relations to each other take the form of family, employment or associational ties and are accordingly viewed in charity law as "private";[87] latterly, in some jurisdictions, it has also come to be directed against trusts that exclude the poor.[88] Secondly, if a purpose is to satisfy the public benefit test, it must, if carried out, stand to benefit people, as opposed to causing detriment to people or having no discernible welfare implications one way or the other. In thinking about how a discriminatory charitable trust might be dealt with under the public benefit test, the first point to note is that, except in unusual cases, such a trust is likely to satisfy the "public" component of the test. As Adam Parachin has insightfully pointed out, trusts that stand to benefit only persons who share — or lack — some common element of human identity are not typically impugned by the "public" component of the public benefit test.[89] And if they are so impugned, it is likely to be because the identity-based class that they stand to benefit is also restricted

86. For fuller discussion of the two components of the public benefit test, see Harding, *Charity Law*, *supra* note 1 at 13-30.

87. See *e.g. In Re Compton*, [1945] Ch 123 (CA (Eng)); *Oppenheim v Tobacco Securities Trust Co Ltd*, [1951] AC 297 (HL).

88. *R (Independent Schools Council) v Charity Commission for England and Wales*, [2011] UKUT 421 (Upper Tribunal Tax and Chancery Chamber) [*Independent Schools*]; *Charities and Trustee Investment (Scotland) Act*, ASP 2005, c 10, s 8(2)(b).

89. Parachin, "Public Benefit", *supra* note 74 at 182-94.

based on family, employment or associational ties. Moreover, according to Parachin, judicial dicta to the effect that a purpose is "public" only where the class that stands to benefit from the purpose somehow reflects the character of the purpose should be read as applying only to purposes that meet the description "general public utility"; those dicta play no role in helping to explain why discriminatory charitable trusts might fail the public benefit test in cases about purposes that are not of "general public utility".[90]

Parachin convincingly argues that if discriminatory charitable trusts are to fall foul of the public benefit test on account of their discriminatory character, it is in most cases likely to be because in some sense their purposes fail the "benefit" component of the public benefit test.[91] However, the application of the "benefit" component in cases about discriminatory charitable trusts is complicated by the fact that such trusts are typically for purposes of types recognised as *prima facie* charitable in law. In the decided cases, discriminatory charitable trusts have usually been for the advancement of education, a type of purpose recognised as *prima facie* charitable since the time of the *Statute of Elizabeth*[92] and one of the four heads of charity articulated by Lord Macnaghten in *Commissioners for Special Purposes of Income Tax v Pemsel*.[93] It is not unreasonable to assert that the legal expression of a type of purpose as *prima facie* charitable indicates that there are strong reasons for thinking that the type of purpose in question stands to benefit people in well recognised and accepted ways.[94] This thought would seem to underpin the longstanding judicial practice of assuming or even presuming the benefit of purposes within certain heads of charity, a practice that continues in some jurisdictions even if it

90. *Ibid* at 186-93, discussing dicta of Lord Simonds in *Inland Revenue Commissioners v Baddeley*, [1955] AC 572 (HL) at 592.
91. Parachin, "Public Benefit", *supra* note 74 at 194-204.
92. *Statute of Charitable Uses*, 1601 (UK), 43 Eliz I, c 4.
93. [1891] AC 531 (HL) at 583 [*Pemsel*].
94. See Mary Synge's description of judicial practice as informed by this proposition: Mary Synge, *The 'New' Public Benefit Requirement: Making Sense of Charity Law?* (Portland: Hart Publishing, 2015) at 22-23.

is prohibited by statute in others.[95] Thus, where a trust is for a purpose of a type recognised as *prima facie* charitable in law — for example, the advancement of education — there are strong reasons to think that the trust satisfies the "benefit" component of the public benefit test, even if it entails discrimination. How, then, might a judge make a finding that a discriminatory charitable trust lacks benefit in these circumstances?

Public benefit jurisprudence points to the answer to this question. That jurisprudence shows that, where appropriate, judges working with the "benefit" component of the public benefit test are prepared to shift their attention away from the general character of a purpose under scrutiny and towards the likely social effects of carrying out that particular purpose.[96] Often this emphasis on likely social effects will entail some engagement with evidence of such effects that the parties have put before the court. A preparedness to look to likely social effects may be discerned in some cases about purposes of types recognised as prima facie charitable in law, most obviously, purposes within the first three heads of charity articulated by Lord Macnaghten in *Pemsel*, viz., "relief of poverty", "advancement of education", and "advancement of religion".[97] For example, in In *Re Macduff*,[98] Lord Rigby recognised that the purpose of teaching the art of theft, although in a general sense within the description of "advancement of education" and to that extent prima facie charitable, would nonetheless have undesirable social effects and therefore fail the "benefit" component of the public benefit test.[99] Moreover, in cases arising under the fourth head of charity set out by Lord Macnaghten in *Pemsel,* "other purposes

95. Judges are prohibited from presuming the benefit of purposes of certain types in England and Wales: *Charities Act 2011* (UK), c 25, s 4(2); Scotland: *Charities and Trustee Investment (Scotland) Act* ASP 2005; and Northern Ireland: *Charities Act (Northern Ireland) 2008* (NI), c 12, s 3, as amended by *Charities Act (Northern Ireland) 2013* (NI), c 3, s 1.
96. Drawing on the language of Jonathan Garton's taxonomy of public benefit, this shift might be described as one away from questions of "conceptual" benefit and towards questions of "demonstrable" benefit: Garton, *Public Benefit*, *supra* note 84 at paras 2.16-2.17.
97. *Pemsel*, *supra* note 93 at 583.
98. [1896] 2 Ch 451 (CA (Eng))
99. *Ibid* at 474.

beneficial to the community",[100] judges have had no choice but to look to likely social effects when applying the public benefit test, because the fourth head does not refer to a type of purpose regarded as prima facie charitable in law beyond pointing judges in the direction of the public benefit test. The best illustration of judges grappling with likely social effects in a case arising under the fourth head is *National Anti-Vivisection Society v Inland Revenue Commissioners*.[101] There, the House of Lords was asked to determine whether the purposes of a society that campaigned for laws to prohibit scientists from performing experiments on live animals satisfied the public benefit test. Among the reasons for decision was that the "benefit" component of the public benefit test was not satisfied. In reaching this finding, the Law Lords considered evidence of likely social effects in the form of the frustration of advances in medical research on the one hand, and the cultivation of humane sentiments on the other, and they concluded that evidence of the former effects was more compelling than evidence of the latter.[102]

Thus, public benefit jurisprudence shows that judges may look to the likely social effects of carrying out the purposes of discriminatory charitable trusts when applying the "benefit" component of the public benefit test to such trusts. And they may do this even in cases where the purposes of the trusts in question are of types that are recognised as *prima facie* charitable in law. But now another question presents itself. This question may be better grasped in light of a crude taxonomy of the social effects that might be generated by carrying out a purpose. This taxonomy is made up of two categories. The first contains what might be called "direct" social effects. So, for example, a trust for the purpose of providing educational scholarships to "white Protestant boys" might, when carried out, give educational opportunities to children within that class and at the same time deny other children outside the class similar educational opportunities. Carrying out the purpose of the trust is, in a sense, an immediate cause of these outcomes: to this extent, the outcomes fall within the category of direct social effects. In the second

100. *Pemsel*, *supra* note 93 at 583.
101. [1948] AC 31 (HL) [*Anti-Vivisection*].
102. *Ibid* at 48-49, per Lord Wright.

category are what might be called "indirect" social effects. So, to return to the example, the purpose of providing educational scholarships to "white Protestant boys" might, depending on the circumstances in which it is carried out, express attitudes and beliefs about certain identity-based groups that demean members of those groups, and those expressive effects might be felt in public culture irrespective of the impact of carrying out the purpose on children whose educational opportunities are directly affected, for good or ill, by the discrimination entailed in the purpose.

The question that is illuminated by this crude taxonomy of social effects is this: to what extent should indirect social effects be taken into account, along with direct social effects, in the setting of an inquiry into whether some discriminatory charitable trust satisfies the "benefit" component of the public benefit test? This question about indirect social effects is of general importance in public benefit jurisprudence, but it is of particular importance in cases about discriminatory charitable trusts because, according to many philosophers, discrimination has considerable indirect social effects that help to explain why it ought to be of moral and legal concern.[103] Indeed, in certain cases, whether a discriminatory charitable trust satisfies or fails the public benefit test might turn on the extent to which the indirect social effects of discrimination are taken into account. For example, return once more to the example of a trust for providing educational scholarships to "white Protestant boys", and imagine that the evidence shows clearly that students excluded from this class suffer no loss of educational opportunity as a result of that exclusion, perhaps because they remain eligible for a range of other scholarships that are not discriminatory in the same way. In these circumstances, the "benefit" component of the public benefit test seems clearly satisfied if only direct social effects are taken into account, but the position seems far

103. For detailed discussion, see Harding, *Charity Law*, *supra* note 1 at 226-33.

more complex once indirect social effects are brought into view.[104]

There are indications in the public benefit jurisprudence that judges are sometimes reluctant to take into account indirect social effects when working with the "benefit" component of the public benefit test. For example, in *R (Independent Schools Council) v Charity Commission for England and Wales*[105] ("*Independent Schools*"), the Upper Tribunal (Tax and Chancery) was invited to find that the purposes of independent fee-charging schools were not of public benefit because they contributed to inequality and social division in British society.[106] The Tribunal noted the force in such arguments but declined to take such claimed indirect social effects into account in applying the public benefit test. For the Tribunal, the evidence presented on the question of indirect social effects was not substantial enough to influence inquiries into public benefit; moreover, to make findings on the basis of evidence of the indirect social effects of private schools would be to venture onto political territory that was properly the province of the legislature.[107] But if *Independent Schools* reveals judicial wariness about indirect social effects, there are also cases on public benefit that indicate a different judicial attitude. In *Re Resch*,[108] for example, the Privy Council accepted the proposition that a private hospital stood to benefit the public because the private hospital relieved

104. See *St Margaret's*, *supra* note 85 at 12, 22-29 (where the Panel was asked, *inter alia*, whether the purposes of an adoption agency that excluded same sex couples from its services satisfied the public benefit test. In finding that the public benefit test was satisfied, the Panel made no reference to the indirect social effects of discrimination. Instead, the Panel focused on direct social effects, noting that no same sex couples had ever sought the services offered by the agency in question and that same sex couples seeking adoption services had a range of other agencies to choose from). In contrast, see *Catholic Care (Diocese of Leeds) v The Charity Commission for England and Wales*, [2011] UKFTT B1 (First Tier Tribunal General Regulatory Chamber) at para 59 (where the Tribunal acknowledged the indirect social effects of discrimination against same sex couples in the provision of adoption services).
105. *Independent Schools*, *supra* note 88.
106. *Ibid* at para 29.
107. *Ibid* at paras 96-109.
108. [1969] 1 AC 514 (PC (Austl)).

public hospitals and therefore the state of the burden of health care.[109] And in *Neville Estates Ltd v Madden*[110] ("*Neville Estates*"), Justice Cross considered that the purposes of a synagogue closed to the public satisfied the public benefit test because worshippers would leave the synagogue and go out into the community, thereby generating benefit to the public (albeit in ways that were not well explained in the judgment).[111]

Of course, in *Re Resch* and *Neville Estates*, purposes that ostensibly failed the public benefit test were ultimately found to satisfy the test once indirect social effects were taken into account. In contrast, if indirect social effects are to be taken into account when applying the "benefit" component of the public benefit test to discriminatory charitable trusts, then a purpose that ostensibly satisfies the public benefit test may ultimately be found to fail the test after those indirect social effects are taken into account. But it is not the outcomes in *Re Resch* and *Neville Estates* that are of present interest: it is the method adopted in those cases. The cases show that public benefit jurisprudence contains resources for judges who are minded to take into account indirect social effects in working out whether discriminatory charitable trusts satisfy the "benefit" component of the public benefit test. In light of these resources, judges may apply the "benefit" component of the public benefit test in a way that is sensitive to benefits associated with a discriminatory charitable purpose (including, importantly, benefits indicated by the fact that the purpose is of a type recognised as *prima facie* charitable in law), but at the same time takes into account both direct and indirect adverse social effects associated with the discrimination entailed in the purpose. These latter effects may be described, using the language of public benefit jurisprudence, as detriments to be taken into account in working out whether the "benefit" component of the public benefit test is met.

Earlier, I argued that, in cases of discriminatory charitable trusts, judges applying the public policy doctrine must balance the value of freedom of disposition, which is properly viewed as a function of the expressive and teleological aspects of choices entailed in the exercise of

109. *Ibid*.
110. [1962] Ch 832 (Eng) [*Neville Estates*].
111. *Ibid*.

that freedom, against the value of equality. Now we are in a position to see that judges who sought to apply the public benefit test in such cases would also be required to undertake a balancing exercise. In the setting of the public benefit test, the balancing exercise would be one according to which benefits associated with the purpose of a discriminatory charitable trust must be weighed against the detriments associated with the discrimination entailed in carrying out that purpose. As with the treatment of indirect social effects, public benefit jurisprudence offers judges some guidance in undertaking such a balancing exercise. In *Anti-Vivisection*, for example, the House of Lords placed emphasis on tangible and material social effects in finding that detriments associated with the frustration of advances in medical research outweighed benefits associated with the cultivation of humane sentiments.[112] But such guidance can only take judges so far; after a point, public benefit jurisprudence is vague as to how benefits and detriments ought to be weighed. Indeed, as I have argued elsewhere, the most plausible reading of *Anti-Vivisection* is one according to which the Law Lords experienced the balancing exercise as demanding a political choice between incommensurable considerations of value.[113]

That an application of the public benefit test in cases of discriminatory charitable trusts may lead to judges having to make choices between incommensurables should not be counted as a shortcoming in the public benefit jurisprudence; rather, it should be seen as a necessary consequence of the nature of what judges are asked to balance when working with

112. *Anti-Vivisection*, *supra* note 101 at 48-49, per Lord Wright.

113. See Harding, *Charity Law*, *supra* note 1 at 25.

that jurisprudence.[114] A robust public benefit jurisprudence dealing with discriminatory charitable trusts would recognise this fact. However, at the same time such jurisprudence would seek to provide reasoned accounts of the various benefits and detriments associated with such trusts. And in light of these reasoned accounts, judges might be able to identify circumstances in which discriminatory charitable trusts either clearly fail or clearly satisfy the "benefit" component of the public benefit test notwithstanding the incommensurable character of the benefits and detriments associated with them.[115] For example, the benefits associated with the purpose of such a trust might be relatively modest but significant detriments might be associated with its discriminatory terms; equally, there might be evidence of substantial benefits associated with the purpose of a trust but only negligible evidence of detriments associated with the discrimination entailed in that purpose. Moreover, even in cases where it is unclear how to balance incommensurable benefits and detriments associated with a discriminatory charitable trust, reasoned accounts of those benefits and detriments may be expected at least to give judges a clearer sense of what is at stake in their choices.

---

114. See also Parachin, "Public Benefit", *supra* note 74 at 204-205 (Parachin suggests that the balancing of incommensurables that is required if the public benefit test is to be applied to discriminatory charitable trusts takes judges beyond "the traditionally conceived boundaries of the judicial realm". This may be true, but as cases like *Anti-Vivisection* show, it is in the very nature of the public benefit test that it requires judges to make political choices among incommensurables; for as long as judges must apply that test when working out whether purposes are charitable in law, they must venture beyond the "traditionally conceived boundaries" to which Parachin refers).

115. This may be possible where the incommensurability of such benefits and detriments is, in Timothy Endicott's language, "vague" rather than "radical". Vague incommensurables may be compared with each other; sometimes the outcome of this comparison will be clear, but in a range of cases it will not. Radical incommensurables cannot be compared with each other. See Timothy Endicott, "Proportionality and Incommensurability" (2012) Oxford Legal Studies Research Paper No 40.

IV. Conclusion

A public benefit jurisprudence that took into account both direct and indirect social effects, and that demanded a balancing exercise in cases where such social effects took the form of both benefits and detriments, would be a useful instrument for judges seeking to respond in nuanced ways to discriminatory charitable trusts. That said, as I argued earlier, the public policy doctrine, at least where it rests on a sound understanding of the value of freedom of disposition, is also a useful instrument for such judges. At the same time, both the public policy doctrine and the public benefit test demand that competing considerations be weighed against each other, and a balance struck between them, in the setting of discriminatory charitable trusts: in the case of the public policy doctrine, it is the values of freedom of disposition and equality; in the case of the public benefit test it is the various benefits and detriments associated with the purpose of a discriminatory charitable trust. Given that both the public policy doctrine and the public benefit test seem fit, or at least might be rendered fit, for dealing with discriminatory charitable trusts, and given that neither tool is able to avoid the difficulties associated with balancing competing considerations, it is worth asking whether anything would be gained if judges ceased reaching for the theme of public policy when deciding cases about discriminatory charitable trusts, and started reaching for the theme of public benefit in its place.

One possible advantage is that, by reaching for the theme of public benefit, judges might be able to respond to, and in appropriate cases interfere with, discriminatory charitable trusts without at the same time engaging with the value of freedom of disposition. Earlier, I argued that this value might be misunderstood as having nothing to do with the expressive or teleological aspects of choices that are entailed in the exercise of dispositive freedom. To the extent that judges misunderstand the value of freedom of disposition in this way, there is a risk that they will attribute weight to that value in circumstances where no or less weight ought to be attributed to it. Consequently, discriminatory trusts might be upheld in circumstances where they should rather be interfered with. But if judges turn to the theme of public benefit when dealing with discriminatory charitable trusts, then at least in cases about discriminatory

charitable trusts judges will not have to consider the value of freedom of disposition; after all, the public benefit test is uninterested in that value. To that extent, the risk will be minimised that a misunderstanding of the value of dispositive freedom might affect the treatment of discriminatory charitable trusts.

However, at the same time a different risk might be realised. If judges come to think that discriminatory charitable trusts are a matter for public benefit jurisprudence, and that discriminatory "private" trusts are a matter for the public policy doctrine, then they may assume, as Tarnopolsky JA did in *Canada Trust*, that the reason why discriminatory charitable trusts are subject to equality norms is that they are in some sense "public" trusts, and on the basis of this assumption they may conclude that discriminatory "private" trusts are not subject to equality norms because of their "private" character. Earlier I argued that there is no reason to think that dispositions of a "private" character are necessarily exempt from equality norms that operate in the public sphere: such norms are relevant to understanding and evaluating the expressive and teleological dimensions of the choices entailed in "private" dispositions just as they are relevant to understanding those dimensions of the choices entailed in "public" dispositions. In truth, both charitable *and* "private" trusts that discriminate are subject to equality norms because the value of exercises of freedom of disposition that underpin them is, in part, to be understood in light of such norms. In addition, discriminatory charitable trusts are subject to equality norms via the public benefit test. So judges exercising equity's jurisdiction over charity are spoilt for choice in deciding how to give legal expression to the equality norms that apply to discriminatory charitable trusts. Judges exercising equity's jurisdiction over "private" trusts do not enjoy this luxury. In these circumstances, far from eschewing the public policy doctrine in favour of the public benefit test, there may be a case for judges exercising equity's jurisdiction over charity to eschew the public benefit test and continue, as they have done to date, to invoke the theme of public policy when interfering with discriminatory charitable trusts. In this way, such judges may ensure that a legal basis is maintained for interfering with discriminatory trusts of all types, whether charitable or "private" in character.

Conscience as the Organising Concept of Equity

Alastair Hudson*

This article sets out a defence of the concept of equity based on conscience by tracing its development from the earliest cases, by establishing that a conscience is something objective and not subjective, and by demonstrating that the idea of conscience provides a coherent central principle for equitable doctrines. Equity is based on a methodology identified by Aristotle in his Ethics which seeks to mitigate the rigour of abstract rules, and also on the idea of conscience. Contrary to most of the assumptions made in the academic commentary on equity, a conscience is an objectively constituted phenomenon. This understanding of a conscience is a commonplace across our culture in sources as disparate as the work of Freud and Kant, in Shakespeare's King Lear, and in Walt Disney's Pinocchio. The conscience is the internal policeman which is planted in our minds by our interactions with the outside world. Consequently, when a court judges in the name of conscience, that court is holding up the individual's behaviour to an objective standard. This conceptualisation of conscience and of "unconscionability" is shown to be the common thread running through the law on dishonest assistance, secret trusts, bribery, proprietary estoppel, ownership of the home and so on. The centuries-old arguments about the efficacy of equity turn on this understanding of a conscience and they can be resolved by reference to it.

* Professor of Equity & Finance Law, University of Exeter National Teaching Fellow.

I. Introduction

There has only ever been one real argument about equity. It was running in the sixteenth century and it is still running today. Either one considers equity to be open-textured and just, ideally suited to a world of constant change and unexpected challenges; or one sees it as

an enemy of order in the law, especially at a time when society needs certainty in its rules. That argument might have begun as a jurisdictional dispute between medieval courts, it might then have morphed into a dispute about whether courts of equity should have open discretion or operate a system of precedent, and it might then have refocused on the viability of the idea of "conscience" for the last four hundred years, but the modern arguments about constructive trusts, restitution and so forth are in essence that same argument in different clothing.

This argument is said to have begun in the reign of Henry II when the Lord Chancellor acquired a jurisdiction beyond the Council to dispense judgments which began increasingly to disagree with the common law.[1] This fight for territory reached its peak when Lord Chief Justice Coke argued with Lord Chancellor Ellesmere about whether the judgments of equity should take priority over the common law. This resulted in the judgment in the *Earl of Oxford's Case*,[2] which set the foundations of an English equity which combined Aristotle's model of equity with the idea of conscience, both of which are considered below. Subsequently, the argument shifted to a suspicion of the discretion which the numerous courts of equity[3] deployed in the name of conscience. It was argued that the courts of equity were not bound by precedent at one time.[4] Richard Francis began his *Maxims of Equity*,[5] published in 1739, by confronting the assertion that decisions of courts of equity were "uncertain and precarious"[6] because they were "not … bound by any established Rules or Orders".[7] His answer to this charge was that conscience would not cause judges to act arbitrarily but rather that each of those judges would

1. See George Spence, *The Equitable Jurisdiction of the Court of Chancery* (London: Stevens and Norton, 1846) at 117.
2. (1615), 21 ER 485 (Ch) [*Earl of Oxford*].
3. There were several courts of equity: the Court of Chancery, the Court of Requests, the Court of Star Chamber and several other courts which dispensed equity.
4. Mitch 31 Hen Hen VI, Fitz Abr, *Subpena* pl 23 (as Fortescue CJ declared in 1452 "[w]e are to argue conscience here, not the law").
5. Richard Francis, *Maxims of Equity*, 2d (London: Lintot, 1739).
6. *Ibid* at 197.
7. *Ibid*.

"be guided by that infallible Monitor within his own Breast"[8] — that is, his conscience. Of course, to modern eyes this reads like the judges would be acting on the basis of their own, subjective consciences. In time, the argument shifted from this debate about unfettered discretion, which began to be resolved when Lord Chancellors like Ellesmere and Bacon began to consult precedents,[9] to a debate about the open-textured nature of the idea of "conscience".

This essay considers the meaning of the term "conscience" in the context of equity. Most of the discussions of the law in this area have focused on a consequentialist examination of the cases in which the words "conscience" or "unconscionable" have been mentioned. They conclude correctly, that no clear definition of "conscience" is ever given by the judges in the cases.[10] However, this leaves two stones unturned. First, the important understanding of a conscience as something that is "objectively constituted" and not entirely subjective, as most of the juristic discussions of conscience assume. Understanding that a conscience is something that is objective helps us to understand in turn that what the courts of equity are doing is to measure the defendant against an objective standard of acceptable behaviour. Second, the term "conscience" (and its technical corollary "unconscionable") does not need an *a priori* definition from the courts any more than the word "reasonable" requires such a dictionary-style definition from the common law courts. Rather, its meaning is to be derived from an analysis of the cases in which it has been used and by juristic discussion about what it should mean in the future.

This essay advances an understanding of an objective conscience and

8. *Ibid*.
9. Spence, *supra* note 1 at 416.
10. See *e.g.* the excellent surveys of the case law and the lack of any clear definitions in Martin Dixon, "Confining and Defining Proprietary Estoppel: The Role of Unconscionability" (2010) 30:3 Legal Studies 408; Nicholas Hopkins, "Conscience, Discretion and the Creation of Property Rights" (2006) 26:4 Legal Studies 475; Dennis R Klinck, "The Unexamined 'Conscience' of Contemporary Canadian Equity" (2001) 46:3 McGill Law Journal 571 [Klinck, "Unexamined Conscience"]; Margaret Halliwell, *Equity and Good Conscience*, 2d (London: Old Bailey Press, 2004).

demonstrates how that is viable in modern equity. We lawyers like to tell our stories and we like to build our models. This is the story of equity, the unmerited aggression it has attracted and the way in which a little wooden boy chose whether to grow up straight like the golden metwand his conscience required, or crooked like the timber that makes up most of humanity.

II. Roots

A. All That Glitters is Not Golden

Lord Coke expressed a typical desire for order in the law when he said that "all causes should be measured by the golden and straight metwand of the law, and not the incertain and crooked cord of discretion".[11] In other words, Coke considered that English law in the 17th century should prefer the golden metwand of the common law to the crooked cord of uncertainty that was offered by equity.[12]

A metwand is a measuring rod. Straight. Stiff. Accurate. Of course, it is useless at measuring anything which is not straight. There will always be a need for measuring things which are made in more intricate shapes. So, what sort of measure should we use if the metwand will not help us? Aristotle had an answer: "[a]n irregular object has a rule of irregular shape, like the leaden rule of Lesbian architecture: just as this rule is not

11. Edward Coke, *Fourth Institute of the Laws of England: Concerning the Jurisdiction of the Courts* (London: Law Booksellers, 1817) at 40-41; Edward Coke, *First Institute of the Laws of England: Commentary Upon Littleton* (London: S Brooke Paternoster Row, 1818)(Lord Coke referred to "the crooked cord of private opinion, which the vulgar call discretion" at 227b). Again, the invective in calling equity enthusiasts "vulgar". See also *Keighley's Case* (1609), 10 Co Rep 139a (KB (Eng))(Lord Coke suggested that discretion should be "limited and bound with the rule of reason and law").

12. Lord Hodson echoed that sentiment in *Pettitt v Pettitt*, [1970] AC 777 (HL) at 808 (the case which ironically spawned so many different approaches to the ownership of the home in England and Wales based on common intention).

rigid but is adapted to the shape of the stone …".[13] Aristotle used this example of the Lesbian architects' rule to explain how equity and formal legal rules interact. The Lesbian architects' rule was made of a malleable lead which would enable stonemasons to measure the intricate coppice work for which the island's workers were famous.

This leaden rule was more useful in particular circumstances than straight rules, just as equitable principles can be more useful than common law rules in specific circumstances. Aristotle used this metaphor as an explanation of why equity would sometimes be "superior" to formal codes of justice:

> equity, although just, is better than a kind of justice … it is a rectification of law in so far as law is defective on account of its generality. This in fact is also the reason why everything is not regulated by law [as opposed to equity]: it is because there are some cases that no law can be framed to cover, so that they require a special ordinance [or judgment]. An irregular object has a rule of irregular shape, like the leaden rule of Lesbian architecture: just as this rule is not rigid but is adapted to the shape of the stone, so the ordinance [or judgment] is framed to fit the circumstances.[14]

Aristotle argued for a qualified superiority of equity over strict systems of formal justice because equity can tailor a remedy to meet the needs of a particular case whereas formal rules can only legislate for the universal case. It is important to note that Aristotle was aware of the tension between an equity which could overrule legislation and the need for formal justice. I have argued that this should not be an argument for the eradication of equity[15] nor for rigid rule-making, but rather for a synthesis of these different types of law, as discussed below.[16] For Aristotle, the two systems

13. Aristotle, *Nicomachean Ethics*, translated by JAK Thomson (London: Penguin Classics, 1955) at 198 [Aristotle, *Nicomachean*].
14. *Ibid* at 200.
15. But see Peter Birks, "Trusts Raised to Avoid Unjust Enrichment: The *Westdeutsche* Case" (1996) 4 Restitution Law Review 3; Jack Beatson, *Use and Abuse of Unjust Enrichment* (Oxford: Clarendon, 1991); Peter Jaffey, *The Nature and Scope of Restitution* (Oxford: Hart, 2000)(who have argued for the extinguishment of equity in favour of the emerging confusions of English unjust enrichment thinking).
16. Alastair Hudson, *Great Debates in Equity & Trusts* (Basingstoke: Palgrave, 2014) at 250 [Hudson, *Great Debates*].

should work together to ensure that justice is achieved for the individual. If the legal system is predicated on the need for justice then there does not need to be any conflict between these two systems of law.[17]

Equity specialists simply have a different perspective from legal positivists on humanity and on the need to treat individuals (on some occasions) as being ends in themselves. Immanuel Kant said that "from the crooked timber of humanity no straight thing was ever made".[18] Human beings are quixotic, irrational creatures for all their supposed rationality and need for order. Consequently nothing entirely ordered comes from them. Anyone who has practised law would acknowledge this. Clients can be tearful, angry, deceitful, fragile, irrational and indeed sometimes crooked. Family lawyers have long recognised that they must use high-level principles to evaluate the needs of the human beings who come in front of them, and then frame the remedy to fit the circumstances.[19] Financial regulation has long since abandoned the hope that rigid rulebooks would deal with the non-stop movement in financial markets and instead have based their rulebooks on high-level principles which act as aids to interpretation for the more detailed regulations.[20]

---

17. *Cowper v Cowper* (1734), 2 P Wms 720 (Ch (Eng))(Sir Joseph Jekyll explained this mutual support between these areas of law when he held that"[t]he discretion which is exercised here, is to be governed by the rules of law and equity, which are not to oppose, but each in its turn to be subservient to the other; this discretion, in some cases, follows the law implicitly, in others, assists it, and advances the remedy, in others again, it relieves against the abuse, or allays the rigour of it" at 752).
18. Immanuel Kant, "Idea for a General History with a Cosmopolitan Purpose" (1894) 10 London Magazine 385 at 388 (originally published in 1784, in proposition 6).
19. *Miller v Miller*, *Macfarlane v Macfarlane*, [2006] UKHL 24 [*Miller*].
20. *British Bankers Association, R v The Financial Services Authority*, [2011] EWHC 999 (QB (Admin)); see Alastair Hudson, *The Law of Finance*, 2d (London: Sweet & Maxwell, 2013) at 9-33 [Hudson, *Finance*] (especially the discussion of the Financial Conduct Authority's Principles for Businesses Rulebook ("PRIN") which underpins the entire regulatory structure in the UK: "Principles for Business" Financial Conduct Authority, (April 2013) online: <http://www.fca.org.uk/static/documents/handbook-releases/high-level-standards136.pdf> [PRIN]).

(Those two contexts are considered in greater detail below). Different types of rule-making and dispute resolution are necessary in the modern world.

This is not to argue that orderly rule-making is unimportant; rather, it is to argue that a different form of dispute resolution will be appropriate in some circumstances. Consequently, Aristotle identified that no formal system of rules can hope to deal fairly with all of the issues which may be raised. The question is how well equity can fill that gap.

For this limited and entirely moral Aristotelian ambition, the existence of equity as a distinct stream of law in England has always enraged those common lawyers who see it as being subversive of the need for good order in legal rules. Its incremental development of principles and the supposed discretion of its judges was said to threaten the order which the common law sought to create. In giving judgment, Chief Justice Hale said that "[b]y the growth of equity on equity the heart of the common law is eaten out".[21] That is quite a charge. Indeed that sort of unnecessarily bitter invective continues in the modern discussion. Birks infamously described people who were prepared to allow conscience to underpin an area of law as being akin to the vile Nazi Heydrich, architect of the Final Solution and thereby the Holocaust, who expressed himself as being able to reconcile his evil with his conscience. While Birks may have claimed he was merely showing that all involved were falling into the same mistake as Heydrich, the metaphor is deeply unpleasant.[22] All of this invective is all the more remarkable given that the principal goal of equity specialists is to work sensitively so as to prevent unconscionable advantage being taken of claimants and so as to achieve just results in particular cases.

---

21. *Rosecarrick v Barton* (1672), 1 Cha Cas 217 (Ch (Eng)) at 219.
22. Birks, *supra* note 15 at 20. For a counter-blast to this unpleasant line of argument see Hudson, *Great Debates*, *supra* note 16 at 64; Steve Hedley, "The Taxonomy of Restitution", in Alastair Hudson, ed, *New Perspectives on Property Law, Obligations and Restitution* (London: Cavendish, 2004) at 151 [Hedley, "Taxonomy"] (both marvel at the use of the holocaust to cause such a weak debating point, as though arguing for a constructive trust were equivalent to arguing for genocide).

It has been suggested that judges like Lord Browne-Wilkinson have "discovered" the idea of conscience in England and Wales only in the mid-1990s. The response to that suggestion is threefold. First, those commentators simply failed to notice that judges like Vice-Chancellor Megarry in *Re Montagu's Settlement Trusts*,[23] Mr Justice Scott in *Attorney General v Guardian Newspapers Ltd (No 2)*[24] ("*Spycatcher*") and others always used the idea of "conscience" in their judgments. Second, English law was turned on its head by the open-textured juristic imagination of Lord Denning which brought so many equitable ideas with it. In the 1980s, the courts determinedly dismantled many of his ideas like estoppel licences, the new model constructive trust and the married woman's equity.[25] Consequently, that sort of discretionary talk became unfashionable.[26] When this demolition work was finished, it became possible for the conscionable brickwork of equity to be seen again. Third, in an increasingly complex world, we retreat into examinations of our values and of our principles. In essence, we return to our roots.

B. The Intellectual Roots of Equity

There are two intellectual roots to equity which are significant. The first is Aristotle's *Ethics*.[27] For Aristotle, equity is "superior to justice" in the sense that it allows decisions to be reached in individual cases

23. [1987] Ch 264 (Eng) [*Montagu's*].
24. [1990] AC 109 (HL) [*Spycatcher*].
25. For the attitude underpinning the former see *e.g. Eves v Eves*, [1975] 3 All ER 768 (CA)(where Lord Denning warned us that "[e]quity is not past the age of child-bearing" at 771); and for an example of the latter see cases like *Ashburn Anstalt v Arnold*, [1989] 1 Ch (CA (Eng))(which took those developments apart).
26. *Allen v Snyder*, [1977] 2 NSWLR 685 (CA (Austl))(the Australian courts were equally suspicious of it, considering that Denning's equitable developments were to be considered to be "a mutant from which further breeding should be discouraged" at 701).
27. Very similar accounts of equity as it related to Aristotle's four-tiered model of justice appear in Aristotle, *Nicomachean*, *supra* note 13; and in Aristotle, *Eudemian Ethics*, translated by Anthony Kenny (Oxford: Oxford University Press, 2011) at 71.

which correct "errors" made by legislators when creating general rules in legislation ("for all law is universal")[28] which did not anticipate individual injustices which might result in their application. Importantly, however, Aristotle explained that equity is also supportive of formal justice (such as statute) in essence because both streams of law are intended to generate justice. Ashburner in his *Principles of Equity* suggested that "[e]quity is a word which has been borrowed by law from morality, and which has acquired in law a strictly technical meaning";[29] and in the first footnote accompanying that opening remark that "[t]ext writers and judges in the time of Elizabeth, when they are dealing with the jurisdiction of the Court of Chancery, often define equity in terms translated from Aristotle",[30] which was an idea he took in turn from Spence and which is evident from judgments of the period.[31] Lord Ellesmere, the Lord Chancellor, held as follows in the *Earl of Oxford's Case*, in describing the purpose of the Court of Chancery: "[t]he cause why there is a Chancery is, for that men's actions are so divers and infinite, that it is impossible to make any general law which may aptly meet with every particular act, and not fail in some circumstances".[32] Self-evidently, his Lordship was echoing Aristotle's ideas very closely in that equity exists in part to meet circumstances in which the general rules of common law fail to provide just outcomes.

The second root of equity is the specifically conscience-based equity which was developed in England. The concept of conscience was established in this area by the time of the judgment of Lord Ellesmere in that same case when he explained that "[t]he office of the Chancellor is to correct men's consciences for frauds, breach of trusts, wrongs and oppressions, of what nature soever they be, and to soften and mollify the extremity of the law …".[33] Thus conscience is already considered to be

28. Aristotle, *Nicomachean*, *supra* note 13.
29. Denis Browne, *Ashburner's Principles of Equity*, 2d (London: William Clowes and Sons, 1933) at 3.
30. *Ibid*.
31. Spence, *supra* note 1 at 412.
32. *Earl of Oxford*, *supra* note 2.
33. *Ibid*.

at the heart of English equity. We shall consider the idea of conscience in greater detail below. We shall identify: first, the significance of its use by Lord Chancellors who were clerics tending the monarch's conscience; second, the significance of those clerics addressing the impure consciences of people who had committed frauds; third, the significance of the court "correcting" someone's conscience according to an objective idea of what that conscience ought to contain; and, fourth, the significance of the Court of Chancery interfering only when there have been unconscionable wrongs committed which the common law cannot prevent.

C. The Sublime Conscience

The Lord Chancellors (until the appointment of Lord Keeper Williams at the beginning of the 17th century) had all been ecclesiastics.[34] Therefore, the language of a sublime conscience in which they spoke, as bishops in the Christian church, who ministered and administered the monarch's conscience, came naturally to them when making judgment. They "ministered" in the sense of acting as priests and as de facto "prime minister" to the monarch. They "administered" in the sense of running the Chancery which issued writs on behalf of the Crown, and latterly running the Courts of Chancery and of Star Chamber.

Nevertheless, in all of this there remains a very important root for the modern, psychological notion of a conscience. The Lord Chancellor was often referred to as being the "keeper" of the monarch's conscience, which suggested that the conscience was something distinct from the monarch as a person.[35] As sovereign, the King or Queen had divinely-bestowed royal duties as well as a corporeal self. Therefore, the conscience that was activated through the Courts of Chancery was an expression of this monarchical power. When we consider the psychological understandings of a conscience today, we shall see that this separateness of the conscience from the conscious mind of the individual human being is central to its operation.

34. Geraint W Thomas, "James I, Equity and Lord Keeper John Williams" (1976) 91 English Historical Review 506.

35. Spence, *supra* note 1 (as Sir Christopher Hatton said, it is "the holy conscience of the Queen, for matter of equity" at 414).

Fascinatingly, Ashburner explained the operation of conscience in the 16th century as involving the court helping the defendant to purge themselves of a bad conscience because the conscience was the voice of God singing in the individual's head. So, it was said that "from the canon law equity derived its power of wringing a confession on oath from the conscience of the defendant".[36] Equity was not above using irons or whips at that time.

So, we can identify the roots of equity in this ancient soil. What this does not do is to take us to the heart of what a modern conscience might involve. It is to that task which we turn next.

III. Conscience and "Unconscionability"

A. The Argument

Whenever I read an account of equity or of conscience written by a legal academic, I assume that they will turn to define the word "conscience" in its colloquial or psychological sense. Aside from the excellent Young, Croft and Smith's *On Equity*[37] or my own work,[38] they never do. I am always surprised because it seems to me that many of their concerns about the term "conscience" being uncertain or amorphous could be resolved if they did so.

By way of an excellent example of a scholarly analysis of these concepts in the cases, Dixon[39] criticised the waxing and waning of the term "unconscionability" in the doctrine of proprietary estoppel in the wake of the decisions in *Yeoman's Row Management Ltd v Cobbe*[40] ("*Cobbe*")(in which Lord Scott appeared to limit the doctrine greatly) and

36. Browne, *supra* note 29.
37. Peter Young, Clyde Croft & Megan L Smith, *On Equity* (Pyrmont: Thomson Reuters, 2009) at 105.
38. See *e.g.* Alastair Hudson, *Equity & Trusts*, 8d (London: Routledge, 2014), s 37.2.
39. Martin Dixon, "Confining and Defining Proprietary Estoppel: the Role of Unconscionability" (2010) 30:3 Legal Studies 408.
40. [2008] UKHL 55.

Thorner v Major[41] (in which Lord Walker and Lord Rodger opened it out again) by demonstrating that the doctrine of unconscionability was never clearly defined. Dixon also identified Lord Walker's odd abandonment of the idea of unconscionability in *Cobbe* after having advanced it so clearly in *Jennings v Rice*.[42] His focus is not on defining what is meant by conscience. The same is true of Hopkins[43] and Klinck,[44] who both model the different approaches taken in different cases where judges have mentioned the concept of unconscionability.

What this commentary does not do is to define the meaning of conscience nor does it explain how such a concept underpins the important work that equity does. It is that project which is undertaken here.

In essence, the point which is advanced here is that a conscience is objectively constituted and that the courts are therefore judging what should have been in that conscience. As Lord Justice Chadwick held "[t]he enquiry is not whether the conscience of the party who has obtained the benefit … is affected in fact; the enquiry is whether, in the view of the court, it ought 'to be'.[45] This is the point: the courts are measuring what a person *ought to have* thought, what their conscience ought to have prompted them to do. The court is able to do that because the court stands for an objective statement of the values which should have been input to a person's conscience. Klinck argues that 'conscience' in equity is not co-extensive with moral conscience",[46] but one has to ask "why not?" If Klinck means that the legal model of conscience cannot be involved with every minor moral imperfection (such as untrue answers to the questions "have you washed your hands?" or "do you really love me?"), then that is sensible. The legal conscience of equity is to be reserved

41. [2009] UKHL 18.
42. [2002] EWCA Civ 159 [*Jennings*].
43. Nicholas Hopkins, "Conscience, Discretion and the Creation of Property Rights" (2006) 26:4 Legal Studies 475.
44. Klinck, "Unexamined Conscience", *supra* note 10.
45. *Jones v Morgan*, [2001] EWCA Civ 995 at para 35.
46. Dennis R Klinck, "The Nebulous Equitable Duty of Conscience" (2005) 31:1 Queen's Law Journal 206 at 212 [Klinck, "Nebulous"].

for imperfections which the courts consider to be sufficiently serious to warrant their attention. Otherwise, equity has always been wrapped up with general questions of morality framed in legal terms, as we shall see.

B. From "Unconscionability" to Conscience

The term "unconscionability" is a term of art. If we want to know what unconscionability means in a particular sense then we have only to examine the most recent case in the applicable jurisdiction on the most relevant equitable principle. For example, the concept of "unconscionability" in relation to liability for knowing receipt is defined in part by the judgment of Lord Justice Nourse, who found in *Bank of Credit & Commerce International v Akindele*[47] that "unconscionability" did not encompass the actions of a person who could not have known that his financial advisors were breaching their fiduciary duties by overpaying him on his investments.[48] Not investigating something which one could not have found out easily will not be considered to be unconscionable, as in *Re Montagu's Settlement*,[49] such that forgetting the terms of a trust prevented liability being imposed on a beneficiary who had absent-mindedly put the pick of the trust fund up for sale at auction. These cases help us to know what "unconscionability" means in this area. This may smack of casuistry; or, as moral philosophers would prefer to call it, consequentialism. Nevertheless, this is the way in which common law systems define terms of this sort.

Sometimes these consequentialist definitions are surprising. We know from the judgment of Lord Upjohn in *Boardman v Phipps*[50] that a defendant may be required to hold personal profits on constructive trust for the beneficiaries of a trust which he was advising as their solicitor, even though it was considered that he had a conscience "completely innocent in every way".[51] His actions were nevertheless considered by the majority of the House of Lords to have infringed a central requirement

47. [2000] 4 All ER 221 (CA).
48. *Ibid* at 238.
49. *Montagu's*, *supra* note 23.
50. [1967] 2 AC 46 (HL).
51. *Ibid* at 123.

of equity that a fiduciary must not benefit from a conflict of interest. That is, his actions were considered to be unconscionable in a technical sense because it is unconscionable for a fiduciary to permit a conflict of interest. This is not a general moral question, rather, it is a legal question. A general moral analysis might lead us to the supposition that a person who risked their own money, as a result of their own extensive research and skill, so as to make profits for themselves and other people should be entitled to keep those profits. By contrast, English equity finds that because that person was a fiduciary at the time, they are required to hold those profits on constructive trust. The unconscionability in the technical sense of that term is based on a fiduciary permitting a conflict of interest to occur.[52] Of course, another approach to the moral question might be to say: the defendant was a solicitor who had studied English trusts law and who therefore ought to have known that he was not entitled to take profits for himself from a trust to which he was giving legal advice without authorisation to do so, and therefore that a constructive trust was entirely appropriate.

The point is that the concept of what constitutes unconscionable behaviour in legal terms is defined in part by a line of authority, just as a section of hedge helps to define the perimeter of a field. The field is defined by the accumulation of those sections of plant that grow into a complete hedge. The shape of the field can move by the addition or removal of hedging material. In the same way, what is meant by the technical concept of unconscionability is defined by the shifting topography of case law.

So, if we want to know what unconscionability means in general terms then we have only to conduct a survey of the most recent binding

52. *Bray v Ford*, [1896] AC 44 (HL), per Lord Herschell ("[i]t is an inflexible rule of the court of equity that a person in a fiduciary position … is not allowed to put himself in a position where his interest and duty conflict" at 50).

precedents across that field covered by the equity textbooks.[53] It is not unknowable simply because the field is large. It is very knowable indeed. In many circumstances, it is possible to use external statements of appropriate behaviour (especially where there is a regulatory code binding on the defendant) to provide a yardstick for unconscionable behaviour in particular contexts.[54]

C. Defining "Conscience"

By contrast, the idea of a conscience is a different question. Many commentators choose not to think about a conscience at all or they fail to answer their own question. However, the conscience is not something which is entirely subjective. Rather, as a psychological phenomenon, it is experienced subjectively by the individual but in truth it is formulated from objective elements. This is self-evident simply from an examination of the literal meaning of the English word "conscience".

The etymology of the English word "conscience" is a combination of the word "con", meaning "with", and the word "science", meaning "knowledge". This idea of "knowledge with" comes from the ancient Greek "*suneidenai*" and refers to a person having "knowledge of oneself with oneself". In particular, the root of the Greek word suggested specifically "sharing knowledge of a defect held with oneself".[55] Here the individual has two separate components in their conscious mind: their conscious self and their conscience, which share knowledge of a defect together. Significantly, conscience is not simply subjective knowledge of oneself, but rather it recognises the existence of a conscious self and

53. This is a field in which there are many colossal textbooks (intended for practitioners and for students). See *e.g.* Klinck, "Unexamined Conscience", *supra* note 10; Klinck, "Nebulous", *supra* note 46 (Professor Klinck's encyclopaedic articles in the area illustrate quite how many cases there are in English and in Canadian law which have used the terms "conscience" and its derivatives).

54. Alastair Hudson, "The Synthesis of Public and Private in Finance Law", in Kit Barker & Darryn Jansen, eds, *Private Law: Key Encounters with Public Law* (Cambridge: Cambridge University Press, 2013) at 231-64.

55. Richard Sorabji, *Moral Conscience Through The Ages* (Oxford: Oxford University Press, 2014) at 12.

another self within one's mind. It is also suggested that a conscience at the time of the ancient Greeks pre-supposed the existence of a law which was being broken so that a defect would come to light.[56]

D. The Conscience as an Objective Phenomenon

1. Conscience in the Psychological Literature

The principal error into which most commentators fall when talking about conscience is in treating it as being an entirely subjective phenomenon which exists solely in the individual human's mind (so that there are seven billion entirely unique consciences in the world). In truth, the conscience exists outside the conscious mind in that there is no possibility of conscious control over it by the rational mind, as was suggested by the etymology of the word "conscience" above. Sigmund Freud explained the creation of the conscience as a psychological phenomenon in *Civilisation and Its Discontents*[57] in the following way:

> a portion of the ego [sets] itself up as the super-ego in opposition to the rest [of the psyche], and is now prepared, as "conscience", to exercise the same severe aggression against the ego that the latter would have liked to direct towards other individuals. The tension between the stern super-ego and the ego that is subject to it is what we call a "sense of guilt"; this manifests itself as a need for punishment.[58]

Thus, the conscience is assembled inside the mind with inputs from outside that individual mind. During infancy those messages come from parents and other family members; during childhood they also come from schoolteachers and others; and then during adulthood there is the legal system, the media and so forth all directing different inputs to the individual's mind.[59] Kant explains conscience in the following way:

> [e]very human being has a conscience and finds himself observed, threatened, and, in general, kept in awe (respect coupled with fear) by an internal judge; and this authority watching over the law in him is not something that he himself (voluntarily) makes, but something incorporated into his being. It

56. *Ibid*.
57. Sigmund Freud, *Civilisation and Its Discontents* (Harmondsworth: Penguin, 1930)
58. *Ibid* at 77.
59. Norbert Elias, *The Society of Individuals* (London: Continuum, 2001).

> follows him like his shadow … [60]

The Kantian model captures the same aggression that Freud identified: the conscience bites and it threatens.[61] The result is a conscience which contains an aggregation of individual responses to those social messages. The upshot is that our culture frequently presents the conscience as being something that is not only objectively constituted but which also exists entirely outside the conscious human mind.

2. The Ubiquity of the Externalised Conscience in Our Culture

Our literature and popular culture teem with examples of this sort of externalised conscience. In Shakespeare's *King Lear*, the Fool acts as conscience to the King and not simply as a capering amusement. Significantly he tells inconvenient truths to the King with the unbridled insolence of a conscience. The King's descent into madness comes when, importantly, he banishes his Fool. This was a frequent trope of European theatre: the capering fool as conscience, idiot savant and truth-teller.

Perhaps the clearest explanation of the external conscience comes from Walt Disney's *Pinocchio*. You may be familiar with the story but you may have forgotten how important the idea of conscience was to the Disney version of the story.[62] Pinocchio was a little wooden boy crafted by Geppetto who came magically to life thanks to a visiting Blue Fairy who wanted to reward Geppetto for his good works. However, the Blue Fairy realised that Pinocchio would lack a conscience and therefore could

60. Immanuel Kant, *Metaphysics of Morals* (Cambridge: Cambridge University Press, 1996) at 189.

61. The Anglo-Saxon expression the "Agonbyt of Inwyt" captures this: literally, the agonising bite of internal wit; or, a pang of conscience. The original metaphor of a "bite" is more visceral than the modern "pang" which suggests something slight.

62. The story of Pinocchio is based on Carlo Collodi's *Le Avventure di Pinocchio* first published in 1881 which was intended to be a warning to children. The character of Jiminy Cricket was a characteristic anthropomorphic development in the Disney version of the story which enlarged a cricket into a talking conscience, complete with top hat, umbrella and spats.

not be a real boy. Consequently, Jiminy Cricket is elevated by the Blue Fairy specifically to act as Pinocchio's conscience. It was the addition of Jiminy Cricket to his psyche that was essential in turning Pinocchio into "a real boy". Jiminy Cricket is a literal, external conscience.

Fascinatingly, the Fairy explains to Pinocchio at the outset that having a conscience, and thus proving himself to be worthy of becoming a real boy, requires that he be "brave, truthful and unselfish". The point is that the Fairy stipulates from the outset what the contents of Pinocchio's conscience should be. The contents of his conscience are dictated from outside himself, just like the psychological model of a conscience. Pinocchio does not decide for himself subjectively what is in good and bad conscience. Instead, Pinocchio misbehaves and his conscience (in the form of Jiminy Cricket) reproves him for it, and ultimately saves him from being transformed into a "jackass". As Pinocchio puts it: "[h]e's my conscience. He tells me what's right and wrong".[63]

The point is that the conscience in equity is an objectively constituted conscience: the question for the court is what a person's conscience *ought* to have told them to do.[64] The court is not asking the defendant what they personally claim to think is right or wrong. Instead, the court is asking what that person's conscience, formed by inter-action with that society, ought to have prompted them to do.

E. The Objective Conscience at Work: Dishonest Assistance

A good example of an objective form of unconscionability at work is the concept of "dishonesty" established in the Privy Council in *Royal Brunei Airlines v Tan*.[65] Lord Nicholls held that the concept of dishonesty was

63. This causes Lampwick to ask: "What? You mean you take orders from a grasshopper?".
64. The same point can be made by reference to the philosophy of aesthetics presented by Adorno whereby the appreciation of art is said to be objectively constituted (by the receipt of messages about what constitutes real art) and yet to be subjectively situated (in that one appreciated art within one's own mind): see Hudson, *Great Debates*, *supra* note 16 at 17.
65. [1995] 2 AC 378 (PC (Brunei)) [*Royal Brunei*].

not a subjective one but rather an assessment by the judge of what an honest person would have done in the circumstances: "acting dishonestly, or with a lack of probity, which is synonymous, means simply not acting as an honest person would in the circumstance. This is an objective standard".[66] So, a defendant was not dishonest because subjectively they believed themselves to be dishonest but rather because that person had failed to do what an objectively honest person would have done in the circumstances. As with conscience, society through the agency of the judge is deciding what an honest person should do.

The criminal law specialist, Lord Hutton, could not let go of the criminal standard of dishonesty in *Twinsectra Limited v Yardley*[67] in the House of Lords. Consequently, he held that dishonesty must involve both a failure to act as an honest person would have acted in the circumstances and also an appreciation on the defendant's part that honest people would have considered their behaviour to have been dishonest. It is this latter subjective element which characterises the criminal law approach. Importantly, it is precisely this latter element which was disavowed by the Privy Council in *Barlow Clowes International Ltd v Eurotrust International Ltd*[68] because it made it too easy for a defendant to argue that his personal moral code permitted him to do the thing which the common morality would have considered dishonest. In that case, the defendant turned a blind eye to the source of the funds which he was being asked to pay through his small fund in the Isle of Man while claiming that his personal moral code prevented him from cross-questioning a client as to the source of his funds. So, equity here does not allow subjective moral relativism. Rather, it deals only in objective standards.

This objective approach in the case law has not been without its problems. Most of those problems have come from a determination among the judges to see dishonesty as involving some level of subjectivity. Lord Nicholls held that "[t]he court will also have regard to personal attributes of the third party such as his experience and intelligence, and

66. *Ibid* at 381.
67. [2002] UKHL 12 [*Twinsectra*].
68. [2005] UKPC 37 (Isle of Man) at para 12.

the reason why he acted as he did".[69] This single sentence in a twelve page judgment which otherwise advances purely objective tests has been relied upon by numerous judges to justify taking into account a range of subjective factors about the defendant[70] ranging from their educational lack of attainment,[71] their lack of experience,[72] through to the stigma that would attach to their professional reputation if they were found to be dishonest.[73] By way of example, in *Markel International Insurance Company Ltd v Surety Guarantee Consultants Ltd*[74] Mr Justice Toulson spent several pages of his judgment examining the subjective situation of the defendant when supposedly applying an objective test. If the test had been entirely objective, then his Lordship would have been better directed to ignore those dozen pages of the defendant's personal history in favour of an assessment of what an honest person would have done in those circumstances.

The model suggested by Lord Nicholls demonstrates how an objective concept can work. The court has no interest in the defendant's personal beliefs, except to the extent that the circumstances in which the defendant found themselves shed light on the way in which an objectively honest person might behave. There is an issue about whether or not this concept of dishonesty actually equates to unconscionability at all. Lord Nicholls was at pains to point out that his test drew on the decision of Lord Selborne in *Barnes v Addy*[75] to the effect that the defendant must be dishonest and that it did not involve a general concept of unconscionability. And yet, this approach puts this area of law generally in line with other equitable doctrines which are concerned with the conscience of the defendant, whether that be as a result of theft, fraud, breach of fiduciary duty or, in this instance, dishonesty.

69. *Royal Brunei*, *supra* note 65 at 391.
70. *Starglade v Nash*, [2010] EWCA Civ 1314.
71. *Ibid*.
72. *Manolakaki v Constantinides*, [2004] EWHC 749 (Ch); *Markel International Insurance Co Ltd v Surety Guarantee Consultants Ltd*, [2008] EWHC 1135 (QB (Comm)).
73. *Twinsectra*, *supra* note 67 at 387 (per Lord Hutton).
74. [2008] EWHC 1135 (QB (Comm)).
75. (1874), LR 9 Ch App 244 (CA (Eng)).

E. Equity as a Methodology

Equity is a way of thinking as much as anything. For Aristotle it was a means of creating just outcomes from the application of limited formal rules. More generally, equity is one of those areas of law which uses high-level principles as moral precepts and aids to interpretation of more detailed rules. Thus, the idea of conscience supports the imposition of constructive trusts over people who acquire property from a trust by fraud both by reference to the general moral question of good conscience and as to the detailed case law, for example, on tracing property rights or the imposition of proprietary constructive trusts.[76] However, if we think of equity as being a methodology — using both general principles to guide decision-making and detailed rules developed as part of a doctrine of precedent — then we can identify different (even purer) forms of equity in other legal fields. Two examples are family law and finance law.

In UK finance law[77] there is a combination of financial regulation (created further to EU and UK statute by statutory regulatory bodies) and of ordinary substantive law (both case law and statutory principles relating to contract, tort, property, trust and so forth).[78] UK financial regulation is predicated on the Principles for Businesses Rulebook ("PRIN") which is used as a means of interpreting all of the thousands

76. *Westdeutsche Landesbank Girozentrale v Council of the London Borough of Islington*, [1996] AC 669 (HL) [*Westdeutsche*].

77. It is appropriate to talk of UK finance law because so many of its regulations are created as part of EU law and therefore the UK is the appropriate jurisdiction. However, when one considers the general law of contract, property and so forth, then there are different jurisdictions within the UK: England and Wales, Scotland, and Northern Ireland. In dealing with academic colleagues in Australia and in Canada, the effect of the presence of the European Union (like that of an enormous orbiting moon with a strong gravitational pull) is something which is difficult to explain. For example, there are few cases on unit trusts in the UK on the basis of trusts law because unit trusts are now regulated by the Financial Conduct Authority with its cheap, comparatively efficient Financial Services Ombudsman scheme.

78. See generally Hudson, *Finance*, *supra* note 20 ch 3 (for an explanation of the structure).

of more detailed regulations in the Financial Conduct Authority and Prudential Regulation Authority "Handbooks".[79] The first principle in PRIN is the requirement that every regulated person act with "integrity". This requirement has been used by the regulatory authorities to impose fines on financial institutions where their behaviour has fallen below the standards expected of them, for example in relation to fixing the important Libor interest rate. Interestingly, this general standard of integrity has been used as the justification for the fine and not more detailed regulations in other parts of the Handbook which had also been breached. This is a form of equity in that a general moral standard is being used to underpin and to apply detailed rules.

Similarly in family law there is a methodology of using high-level principles which are applied to individual cases in a way that is sensitive to context and which observes the rule of precedent after a fashion as a guide to the interpretation of those high-level principles. So, this equity-like methodology is a way of taking a high-level principle and applying it to the needs and circumstances of a particular family group in a way that is both principled and yet sensitive to context. By way of example, Baroness Hale and Lord Nicholls held in the joined appeals of *Miller v Miller* and *Macfarlane v Macfarlane* that applications for ancillary relief required the court to consider the needs of the family members, equal sharing of matrimonial property and family assets, and compensation.[80] This decision (in turn following the earlier decision of the House of Lords in *White v White*)[81] creates a means for lower courts to interpret the general principles applying to ancillary relief proceedings both by establishing general principles and by adding a gloss to earlier general principles. The way in which family law observes precedent is to see the law as being comprised of high-level principles with guidelines set out in the cases as to their application, as opposed to seeing the law as being

79. PRIN, *supra* note 20; the unfortunately named "Handbooks" contain both the regulatory "Rules" and the "Guidance" governing regulated persons: "Financial Conduct Authority Handbook" Financial Conduct Authority, online: < https://www.handbook.fca.org.uk>.

80. *Miller*, *supra* note 19.

81. [2001] 1 AC 596 (HL).

comprised of hard-and-fast rules.

The purest form of the use of equitable technique in English law may be family law. Section 1 of the *Children Act 1989*,[82] for example, provides that the welfare of the child is paramount. This statutory principle creates a high-level, central principle but one which is developed in the case law in a way that some moral philosophers would describe as being "consequentialist": that is, by developing the meaning of that principle through successive cases. That there is a central principle would be described by other moral philosophers as being part of a "deontological" method: that is, where a central moral principle is established which governs all decision-making. Common law systems do both things: they develop their principles consequentially from case-to-case, and they also observe rules and principles which were set out in earlier cases. Equity does not operate as a completely deontological project because its principles are developed through cases: for example, a rule concerning bribery arises only once a case on bribery appears, and so on. Nor is equity completely consequentialist because it has its general principles of "conscience" to guide it, and because the courts intermittently state or re-state its core principles.

These two methodologies — the consequentialist and the deontological — slip over one another like wet crabs in a basket in the equitable context. Equity uses high-level principles and precedents on the interpretation and meaning of those principles, and their application to individual cases, to achieve its goal of mitigating the rigour of the common law and correcting people's consciences. Equity is thus a methodology in itself.

E. Discretion in Equity

Much of the foregoing and on-going debate about equity revolves around a supposed binary division between "discretion" and "no discretion". It is supposed that "discretion" is bad because it permits judges to do whatever they want and in consequence it is presumed that the law is rendered unpredictable and chaotic. By contrast, "no discretion" assumes a series

82. (UK), c 41, s 1.

of carefully crafted taxonomic categories which present order and good sense: something which is not necessarily apparent in the English law of negligence, for example.

Of course, the courts do neither thing. Instead, the English courts have always applied a form of "weak discretion" in their use of equity. A strong discretion would involve the courts in making any decision that they saw fit. The model of strong discretion which was feared by Lord Eldon and judges of a similar ilk was a form of discretion in which the courts of equity were simply doing whatever they thought "right" in the abstract, perhaps by reference to general ethical principles expressed through the maxims of equity as drawn together by Francis and others.[83] This accompanied the debates about whether or not equity was governed by a doctrine of precedent at all. English courts have always been reluctant to do this. There are many examples of the English courts diluting any ostensibly strong discretion they may have. Where the courts were given the power by statute to grant injunctions whenever they considered it to be just and equitable, the Court of Appeal in *Jaggard v Sawyer*[84] set out limitations on the way in which those statutory powers would be used by the courts in practice, and each court continues to abide by those probanda.

A weaker form of discretion is still possible under the Aristotelian model: that is, the correction of the application of formal legal rules in particular cases in line with precedent and in line with clear principles setting out the way in which deviation from those formal legal rules is possible. A good example of this phenomenon is the law on secret trusts in England. The general principle is that a person may not knowingly act in a way which ought to have affected their conscience.[85] The case law precedent has clearly established that it is contrary to conscience, for example, for a person to agree to take property under a friend's will (on the understanding that they will apply that property for the maintenance of

83. Francis, *supra* note 5; Edmund Snell, *Equity* (reprint of the 1838 edition) (London: Sweet & Maxwell, 2000)(both used the approach of beginning with ancient equitable maxims and then linking them to decided cases).
84. [1995] 1 WLR 269 (CA (Civ)(Eng)).
85. *Westdeutsche*, *supra* note 76.

the testator's illegitimate child) and then to purport to take that property beneficially because no living person knows of the arrangement.[86] This line of cases, establishing the principle known as "secrets trusts", offends the provisions of section 9 of the *Wills Act 1837*[87] that a bequest in a will is only valid if it complies with the formalities set out in that statute, and moreover that a formally valid will may not be circumvented by parol evidence. However, the "friend" in this example would clearly be acting unethically in knowingly taking property not intended for them and would be causing harm to the child who was intended to be maintained by that property. Presumably the child's mother would be the one to propel the matter to court in practice. It is a clear example of Aristotle's equity that prompts a judge to prevent this unconscionable benefit being taken from the bequest: the statute did not anticipate this self-evident wrong and therefore the judge ought, on Aristotle's model, to correct the legislator's shortcomings in failing to anticipate that scenario. Assuming the matter comes up to proof, then an English judge would be acting in accordance with precedent in finding that there was a "secret trust" in existence. The approach taken in the case law is a weak discretion, even though it contravenes the *Wills Act 1837*, because it is an example of an unconscionable act which will invoke the creation of a trust which has been established by precedent and which prompts the judge to circumvent the statutory formalities when the authorities permit them to do so. Moreover, the judges only make their finding of a secret trust if the facts are clear and the requirements set out in earlier cases satisfied.

Another interesting example of discretion in English law is proprietary estoppel. At the time of writing, proprietary estoppel is the closest that English equity comes to a remedial constructive trust. Indeed, a part of the reason for there being less of a clamour for a remedial constructive trust in English law is twofold. First, the constructive trust is divided into well-established categories which operated institutionally in quite predictable ways. That is, the constructive trust comes into existence automatically

86. The case law in this area can be traced back at least to *Sellack v Harris* (1708), 2 Eq Ca Ab 46 (Eng) through *McCormick v Grogan* (1869), LR 4 HL 82.

87. (UK), c 26, s 9.

at the time when the defendant's conscience is said to have been affected by the matter which ought to have affected their conscience. As the House of Lords in *Westdeutsche Landesbank v Islington*[88] ("*Westdeutsche*") accepted, seemingly unanimously on this point, the requirements of the law on insolvency were such that this institutional clarity was a necessary facet of the English approach. Second, proprietary estoppel fulfils the need for a remedial doctrine. That proprietary estoppel is remedial is clear. In some cases, the court will award a purely personal remedy for an amount of money as in *Jennings v Rice*[89] where a payment of £200,000 was ordered; whereas in other cases the court will award a proprietary remedy as in *Pascoe v Turner*[90] and in *Re Basham,*[91] where the transfer of the freehold in a house was ordered; or the court may order a mixture of items of property and payments of money to compensate for detriment suffered as in *Gillett v Holt*[92] where the claimant received the freehold of a cottage, an identified field and a sum of money.

The principal limitation of proprietary estoppel is that it is available only where the claimant can demonstrate that there was a representation or assurance made to them on which they relied to their detriment.[93] The existence of this three stage test — which must be satisfied before the estoppel is made out — demonstrates the transformation of a potentially broad remedial discretion into a narrower, weaker discretion by restricting it to cases which satisfy those three elements.

What is apparent, however, is that there remains scope for remarkable flexibility in the remedies. That flexibility may be said to be taken to such extremes in some instances that it suggests a strong discretion. If that is so, there appears to be a division between a rational, needs-based equity and a genuinely creative equity. As an illustration of the needs-based equity, the Court of Appeal in *Baker v Baker*[94] held that an elderly relative

88. *Westdeutsche*, *supra* note 76.
89. *Jennings*, *supra* note 42.
90. [1979] 2 ALL ER 945 (CA (Civ)).
91. [1986] 1 WLR 1498 (Ch (Eng)) [*Basham*].
92. [2000] 2 All ER 289 (CA (Civ)).
93. *Basham*, *supra* note 91; *Thorner v Major*, [2009] UKHL 18.
94. [1993] 25 HLR 408 (CA (Civ)(Eng)).

who would have been entitled to an equitable interest in a co-owned home should be awarded sufficient money to pay for him to acquire an annuity which would fund the nursing and residential care that he would need for the rest of his life. This humane judgment recognised that the needs of the parties superseded the dictates of property law. This was, perhaps, an example of strong discretion being used within the confines of the law on proprietary estoppel to achieve the sort of result that would be familiar to a family lawyer. That is, an outcome focused on the needs of the parties beyond the detail of their property rights.

As an illustration of the latter, "creative" equity in *Stallion v Albert Stallion Holdings (Great Britain) Ltd*,[95] the court was faced with a situation in which Mr Stallion's first wife had been promised that she could occupy her former matrimonial home for the rest of her life provided that she did not contest her husband's divorce petition. Thereafter, the first wife, the second wife and Mr. Stallion himself had cohabited (apparently harmoniously) in the property in London's Waterloo district for some time before Mr. Stallion's death. It was held that the first wife had made out the estoppel and that by way of remedy the first wife and the second wife should continue to cohabit in that property. The suggestion that the former wife and the current wife of a man newly dead should live together by court order is at first surprising. The two women seem to have cohabited harmoniously before. The judge had seen the parties at first hand and heard all of the evidence. The first wife had been made a promise on which she had relied to her detriment in not contesting the divorce petition and the concomitant re-marriage. The judge clearly felt that the needs of all parties could best be met by their continued cohabitation. The resolution of the issues is reminiscent of that sort of creative equity which is the currency of the family courts.

95. [2009] EWHC 1950 (Ch).

IV. Equity: Confidences and Things of Conscience

A. The Obfuscation of the Equitable Roots of Confidence in English Law

Equity presents a way of thinking about legal disputes which is different from the common law. Equity has its own methodology — based historically on its maxims — of taking high-level principles and then, using precedent, applying them to specific factual situations. Typically, common lawyers are suspicious of this approach. What this suspicion can mean is that, in circumstances in which common law and equity overlap, the principles of equity are completely overlooked by common lawyers because the concepts used by equity have no meaning for them. A good example arises in relation to the English law on misuse of confidential information.[96]

In England and Wales, there has been much excitement in recent years about the development of a tort of misuse of private information.[97] What is interesting is that the discussion among the common lawyers about the supposed brave new world of confidentiality has entirely overlooked three things: first, that there continues to be a very important doctrine of breach of confidence in equity; second, that the principal remedy sought in cases of this sort is an injunction based on equitable principles; and, third, that the principles which have led to the new tort of misuse of private information were based on an equitable "obligation in conscience".

In the *Spycatcher*[98] litigation it was treated as settled law that the principle governing the action for breach of confidence was an equitable

96. See Alastair Hudson, "Equity, Confidentiality and the Nature of Property", in Helena Howe & Jonathan Griffiths, eds, *Concepts of Property in Intellectual Property Law* (Cambridge: Cambridge University Press, 2013) at 94-115.

97. *Campbell v MGN Ltd*, [2004] UKHL 22 at paras 13-14 (that is, the same Lord Nicholls who reorganised the principles of dishonest assistance in *Royal Brunei*, *supra* note 65).

98. *Spycatcher*, *supra* note 24.

"obligation in conscience" not to misuse confidential information.[99] What is interesting is that none of the common law commentators who have addressed this equitable root to the breach of confidence action[100] have identified what the equitable doctrine actually *is*. It is generally treated as being "obscure"[101] or not "firmly established".[102] This approach overlooks the continued use of equitable injunctions throughout the history of this doctrine, even in relation to the tort of misuse of private information today. It also overlooks the ubiquitous description of breach of confidence as being based on an "obligation of conscience" in equity in cases such as *Spycatcher* and *Douglas v Hello! Ltd*.[103] Nevertheless, common law commentators do not even mention the word "conscience" in their accounts, in spite of its continued use by the courts of England, Canada and Australia. As such, by overlooking the equitable doctrine, it has been suggested that the "modern action of confidence is of surprisingly recent vintage".[104] One is minded to say that it would appear to be of surprisingly recent vintage if one focuses exclusively on surprisingly recent cases and ignores the use of equitable concepts even then.

The reason for this omission is that conscience simply forms no part of the common law canon. Just as a military historian discussing the Spanish Civil War might overlook developments in Andalucian cuisine as being of no importance in their work, a common law specialist does not see the need to discuss the idea of "conscience". The point here is a simple one. Common lawyers are so antipathetic to the idea of conscience as a legal category that they ignore the word "conscience" and its derivatives completely. Whereas the concepts of conscience and unconscionability are central to equity, they are constantly overlooked when intellectual property lawyers and common lawyers discuss concepts like confidence.

---

99. *Ibid* at 281 (per Lord Goff) and at 255 (per Lord Keith).
100. See *e.g.* Tanya Aplin et al, *Gurry's Law of Confidence*, 2d (Oxford: Oxford University Press, 2012) [Aplin], Paul Stanley, *The Law of Confidentiality* (Oxford: Hart, 2008).
101. Aplin, *supra* note 100 at 12.
102. *Ibid* at 13.
103. [2007] UKHL 21.
104. Aplin, *supra* note 100 at 13.

B. The Continued Rude Health of the Equitable Doctrine of Breach of Confidence

That the equitable doctrine of confidence is still alive and well in England is beyond question. In the Supreme Court decision in *Vestergaard Frandsen A/S v Bestnet Europe Ltd*[105] in 2013, Lord Neuberger reminded us that the doctrine is equitable, that it is based on conscience, and that "in order for the conscience of the recipient to be affected, she must have agreed, or must know, that the information is confidential".[106] Thus, conscience-based equity remains central here. His lordship also quoted Mr Justice Megarry as having held in *Coco v AN Clark (Engineers) Ltd*[107] that "the equitable jurisdiction in cases of breach of confidence is ancient; confidence is the cousin of trust".[108] In 2011, the Court of Appeal in *Tchenguiz v Imerman*[109] made it clear that a "claim based on confidentiality is an equitable claim … [and that accordingly] … the normal equitable rules apply".[110] It is strange then that the common lawyers continue to overlook these principles. No mention of the word "conscience" even appears in their accounts of this area.

C. A Stronger Equity Outside England: the "Obligation in Conscience" in *Spycatcher*

Equity may well have had its roots in the oddities of English history but it seems to have grown deeper roots outside England. By way of example, the roots of the judgment of Scott J in *Spycatcher* were planted in large part in the judgments of Justice Mason in *Commonwealth of Australia v John Fairfax & Sons Ltd*[111] and of Justice Deane in *Moorgate Tobacco Ltd*

105. [2013] UKSC 31.
106. *Ibid* at para 23. Moreover, it is said that "confidence is the cousin of trust" in this context because it is part of the ancient equitable jurisdiction on breach of confidence at para 22.
107. [1969] RPC 41 (Ch (Eng)).
108. *Ibid* at 46.
109. [2010] EWCA Civ 908.
110. *Ibid* at para 74.
111. (1980), 147 CLR 39 (HCA).

v Philip Morris Ltd (No 2)[112] to the effect that there is a "general equitable jurisdiction" to grant relief in relation to breaches of confidence. As it was put by Deane J:

> [l]ike most heads of exclusively equitable jurisdiction, its rational basis does not lie in proprietary right. It lies in the notion of an obligation of conscience arising from the circumstances in or through which the information was communicated or obtained.[113]

In the New South Wales Court of Appeal, Chief Justice Street held that the appropriate form of remedy is in the largest sense an *in personam* remedy in that a court looks to the conscience of the individual in deciding which remedy or doctrine, if any, is appropriate.[114] These conceptualisations of the principles are in the grand tradition of equity and are signifiers of well-understood lines of precedent within equity.

V. Branching Out: Law and Morality

A. Conscience, Unjust Enrichment, and Politics

The judgment of Lord Browne-Wilkinson in *Westdeutsche*[115] reminded us that "conscience" is, and always has been, at the heart of trusts law in England and Wales as well as at the heart of equity. In particular, constructive trusts in the English law context are based on the idea that a proprietary constructive trust arises on an institutional basis by operation of law when the defendant has knowledge of some factor which ought to affect their conscience. The trust is therefore deemed to have come into existence at the moment when the defendant's conscience ought to have been affected: that is, at the moment of acquiring the necessary knowledge of the factor in question.

The debate which has ensued is due to a misunderstanding of what "conscience" involves. As outlined above, the commentators have chosen to treat the term "conscience" as being a purely technical term, and in consequence they have side-stepped the need to address complex

112. (1984), 156 CLR 414 (HCA).
113. *Ibid* at 437.
114. *Spycatcher*, *supra* note 24 at 152 (quoting Street CJ).
115. *Westdeutsche*, *supra* note 76.

questions as to the precise morality which might be capable of being enforced by the courts. This is the result of a postmodern turn in our jurisprudence: that is, we are reluctant to accept claims to "right" and "wrong" from anyone, even if (or possibly, especially if) they are a public official appointed to sit in judgment over us. At conferences, one may even hear trusts law professors describe this sort of traditional equitable approach as being "just not thinking". On that model, the only form of "thinking" is said to be the sort of taxonomy which has become popular among the English restitution school, borrowing a veneer of scientific rigour from the natural sciences in their use of the word "taxonomy". The problem is that the taxonomies produced by natural sciences are observations of the way in which the real world is actually ordered; whereas the taxonomies produced by legal positivists are ideological in nature in that they create structures of the way in which they want the world to be. As Nietzsche put it, the greatest artists in abstraction are in fact the people who create the categories.[116] That is, the purportedly value-free, apolitical rigour of taxonomic thinking actually conceals a deeply political, abstract project in law-making. That project is political in that judges are effectively being lobbied to change the law and in that the law is being remodelled to prefer the needs of commercial people over the rest of society by taking concepts and models from contract law in particular.

Beyond that sort of rigidity, equity permits principled decision-making and sensitive dispute resolution. The recent case law on bribery in England is a good example of this.

B. Bribery

A simple moral problem arises when a person receives a bribe: should that person be entitled to profit from that bribe? That problem is simple because the clear answer, assuming no supervening circumstances, is that no one should be able to profit from the receipt of a bribe. Yet, courts have made heavy weather of the legal conceptualisation of that same question.

116. Friedrich Nietzsche, *The Will To Power*, (reprint of the 1901 edition) (London: Doubleday, 2011) at 513.

Under English law, when a person receives a bribe and invests that bribe successfully, it is apparently quite difficult to know whether or not that person should be entitled to keep those successful investments which were made with the bribe. The argument based on *Lister & Co v Stubbs*[117] and on the decision of the Court of Appeal in *Sinclair Investments (UK) Ltd v Versailles Trade Finance Ltd*[118] ("*Sinclair*") is that the recipient of the bribe (the false fiduciary) only owes their beneficiary a debt equal to the amount of the bribe. There would only be a constructive trust if the fiduciary had misused trust property to acquire for themselves an unauthorised benefit, but not if the fiduciary had taken a bribe from some third party which formed no part of the trust property. If the fiduciary had merely received a bribe then they would be permitted to keep the property acquired with the bribe and any profits taken from that property.[119] This is simply morally wrong by any measure. A wrongdoer will have benefited from their wrongdoing.

The principal reason given for permitting the false fiduciary to keep the bribe is that in cases of insolvency (as in *Sinclair*), unsecured creditors are elevated to the status of secured creditors by being granted proprietary rights under a constructive trust in the bribe and in any property acquired with the bribe. This, of course, begs the question whether or not there is always an insolvency in cases of bribery. The Court of Appeal was gulled into changing the law on constructive trusts in all circumstances relating to bribes because of arguments relating to insolvencies.

It took the Federal Court of Australia in *Grimaldi v Chameleon Mining NL (No 2)*[120] to bring the English courts back to reason in this context. As the Federal Court pointed out, the receipt of a bribe is an unconscionable act which would not give rise to a constructive trust after *Sinclair*, and therefore bribery would be an anomaly in the category of

117. (1890), 45 Ch D 1 (CA (Eng)).
118. [2011] EWCA Civ 347.
119. In practice, accounting for the amount of this debt which is equal in amount to the bribe received may involve selling the property acquired with the bribe, but that does not amount to a proprietary right in that property for the beneficiaries.
120. [2012] FCAFC 6 (Austl).

unconscionable acts which give rise to constructive trusts in England and Wales.

The Supreme Court in *FHR European Ventures LLP v Mankarious*[121] ("*FHR European Ventures*") reversed the Court of Appeal's decision in *Sinclair*. In that case an agent had taken a secret commission from a vendor of land which meant that the agent agreed on behalf of its principal to acquire the land at a higher price than would otherwise have been required to be paid. It was held by Lord Neuberger, in imposing a constructive trust, that it was important to prevent the harm that is caused by bribery and by undisclosed commissions in commercial life. There is also a passing reference to the role of equity being to mitigate the rigour of the law; but there is no mention of conscience.

The fortunate outcome of this change in direction in the law was that the moral question was aligned with the legal analysis: a wrongdoer may not benefit from their wrongdoing, and therefore a false fiduciary may not resist a proprietary claim to the bribe and thus to any property acquired with that bribe. Nevertheless the precise reasoning of the Supreme Court in *FHR European Ventures* is not entirely satisfactory.

It is worthwhile taking a while to re-examine the judgment of Lord Templeman in *Attorney General for Hong Kong v Reid*[122] ("*Reid*") where his lordship was clear in predicating his judgment imposing a constructive trust over a bribe on the idea of conscience. There were two equitable roots in this judgment. First, the idea that the taking of a bribe is unconscionable. Second, the idea that equity looks upon as done that which ought to have been done and therefore, because the bribe should have been given to the beneficiary on its receipt, equity would treat property in the bribe as having passed to the beneficiary immediately on its receipt.[123] The basis for this constructive trust was that the receipt of the bribe was simply wrong. In that case, the receipt of a bribe by the Director of Public Prosecutions of Hong Kong was considered to be corrupt, in common with all bribery. Bribery was described by Lord

121. [2013] EWCA Civ 17 [*FHR European Ventures*].

122. [1994] 1 AC 324 (PC (NZL)) [*Reid*].

123. The beneficiary in that instance would be territory over which Reid was the Director of Public Prosecutions.

Templeman as being "an evil practice". A robust attitude to the moral questions looms large in that judgment. Moreover, the further principle established by Lord Templeman that the fiduciary must account for any reduction in the value of the property acquired with the bribe, as well as holding that property on constructive trust, suggests an element of retribution as well as restitution.

The principal difference — and it is a very important difference — between the judgment of Lord Templeman in *Reid*[124] and the judgment of Lord Neuberger in *FHR European Ventures*[125] is that Lord Templeman based his judgment on the idea of conscience whereas the Supreme Court in the latter case found a constructive trust, notionally in line with the judgment in *Reid*, but without mention of the concept of conscience. Without the idea of conscience, it is more difficult to understand why there is a constructive trust in relation to bribery. With the moral centre of a conscience-based equity restored, it becomes clear again why a constructive trust encompasses cases of bribery as well as cases of conflicts of interest, cases of theft, and so forth. By re-establishing conscience as the moral centre of this area of law, it becomes clear that the constructive trust is being imposed so as to prevent the immoral earning of a profit from a bribe. The wrong is the corruption and the breach of fiduciary duty. The constructive trust is imposed so that no benefit is taken from that wrong and because equity looks upon as done that which ought to have been done.

VI. Trusts of Homes

A comparative account of the different approaches to unconscionability would show how the Commonwealth turned its back on England in relation to a particularly important area of law — the ownership of the home — when English law adopted the "common intention constructive trust".[126] This form of constructive trust was artificial at its heart: the purported "common intention" is often supplied by the court[127] and

124. *Reid*, *supra* note 122.
125. *FHR European Ventures*, *supra* note 121.
126. *Gissing v Gissing*, [1971] AC 886 (HL).
127. *Oxley v Hiscock*, [2004] EWCA Civ 546.

therefore is not in fact the parties' own intention. The House of Lords in *Lloyds Bank v Rosset*,[128] motivated by the prospect of a slew of open-textured rule-making, introduced a two-tier form of constructive trust resonant of the common law of contract. Rights could only be acquired by express agreement coupled with (usually) financial detriment or by contribution to the purchase price or mortgage repayments. The more recent decision of the Supreme Court in *Jones v Kernott*[129] has acknowledged these difficulties and decided that, when a common intention cannot be found, then the court can look to what is "fair". In that regard, they have caught up with the law in New Zealand nearly twenty years after they had followed Lord Denning's lead.[130] The standard of "fairness" is undefined and it will take many years of judicial decision-making for us to see it take shape, to criticise that shape, and for it finally to adopt a recognisable form.

In Canada, the principle of unjust enrichment in ownership of the home — with its wilfully modern attitude to the rights of women and to the general avoidance of injustice in the ownership of the family home[131] — contrasts so starkly with the Oxford restitution school's rigid model of "unjust enrichment" as to be remarkable. English unjust enrichment has nothing to say about family law nor about the home. Its taxonomic focus is on areas abutting contract law and commercial law, together with its grids of numbered categories of claim.[132] While Oxford restitutionists call for the end of equity, they have nothing to say about the law on injunctions, ownership of the home, or those areas of law where high-level principles are used to guide decision-making. Their project, for all its sound and fury, is actually quite narrow. The "equity" they seek to excise is merely a branch of a much larger tree.

In Australia, where equity has always seemed to be at its strongest, the principle of unconscionability in the ownership of the home reflects

128. [1991] 1 AC 107 (HL).
129. [2011] UKSC 53.
130. *Gillies v Keogh*, [1989] NZCA 168.
131. *Pettkus v Becker*, [1980] 2 SCR 834; *Sorochan v Sorochan,* [1986] 2 SCR 38; *Peter v Beblow*, [1993] 1 SCR 980.
132. See Hedley, "Taxonomy", *supra* note 22 at 151.

a pure form of equity. The court considers the parties' relationship and considers whether it would be unconscionable to award or deny rights in the family home between them.[133] There is a principle at its heart which is concerned to observe Aristotle's requirement that formal land law rules should not permit injustice in individual cases. The idea of conscience that accompanies this doctrine has been developed in the cases. It has a generally moral project at its heart: to prevent the continuation of an unfairly gendered system in which one part of society tended to acquire property rights at the expense of the contributions of any other part of that society. There are no confusions in an equity of this sort. It is simply about preventing an identified form of unconscionable activity.

VII. Conclusion

As Hamlet noted, nothing is either good or bad but rather thinking makes it so. The same is true of equity. Thinking that law must *always* involve hard-and-fast rules makes equity seem bad; but hard-and-fast rules will get us only so far in a world in which aeroplanes are flown unexpectedly into tall buildings, in which the entire financial system is able to crash without anyone anticipating it, and in which large parts of the world continue to be at war on the basis of religious denomination. Rigid systems cannot serve all of our needs in the modern world because unanticipated events will require us to be able to react quickly and to create novel solutions for novel circumstances. That is precisely what Aristotle had in mind when equity appeared in his *Ethics*. Moreover, human beings continue to be held in thrall by "big picture" ideas like god, and intangible ideas like financial instruments, the internet and the metaphysical concept of hope. Equity and the idea of conscience fit exactly into this world of big picture ideas. In a world that is constantly changing there is obviously a need for stable moorings, and the law is an essential part of ensuring that there are some things which will last forever, but there is also a need for the courts to be flexible. The world was unpredictable in the sixteenth century and it still is today.

Pinocchio is exhorted always to let his conscience be his guide. In

133. *Baumgartner v Baumgartner* (1988), 62 ALJR 29 (HCA).

the Disney version of the story, he has to "give a little whistle" to call his conscience to him. In the Freudian version, the conscience is always there, monitoring the conscious mind. It comes unbidden. It is this internal monitor which the earliest Lords Chancellor had in mind when they created a system of principles based on Aristotle's *Ethics*. They identified a growing need for a moral core to the law which would be proof against the tricks and contrivances which had become the workaday tools of even medieval lawyers. Over time, with the efforts of Lords Nottingham, Eldon and Hardwicke, many of these moral maxims have hardened through precedent into predictable rules and principles. Parts of express trusts law resemble contract law as much as anything else. And yet, akin to the methodologies of family lawyers and regulation specialists, there is still a vast terrain in which equity operates to find just outcomes using only a weak discretion. What this open-textured idea of a conscience-based equity means is that the legal system is fulfilling a part of its function to set out a code of moral principles alongside its rules. Consequently, we have a little more to go on when our consciences trouble us than simply to "give a little whistle".

Panacea or Pandemic: Comparing "Equitable Waiver of Tort" to "Aggregate Liability" in Cases of Mass Torts with Indeterminate Causation

Craig Jones*

The equitable doctrine of "waiver of tort", in which a plaintiff surrenders the right to tort damages and seeks instead a disgorgement of the defendant's wrongful profits, has received a mixed reception in Canadian courts. In this article, the author explains the doctrine and its difficult history, and proposes that the problem against which waiver of tort is usually being applied — indeterminate causation in mass tort claims — is very real. However, the author concludes that the use of the doctrine of waiver is a partial solution at best, and advocates instead for a more fundamental rethinking of our approach to causation in class actions.

* Professor of Law, Thompson Rivers University. The author is grateful for the research assistance of Kendra Morris, J.D. Candidate, Thompson Rivers University.

I. Introduction

Is the equitable doctrine of waiver of tort, as they say nowadays, "a thing"? If so, should it be the "next big thing" in mass tort class actions?

My answers to these questions are a qualified "yes" and an emphatic "no". Waiver of tort should be recognized as a cause of action, I will argue, and may at times be useful, but it should not be the doctrine of first, or even second, resort in mass tort class actions, as it is generally inferior to the available alternative: the evolution of tort law to permit the aggregate determination of causation in large-scale claims.

What exactly is waiver of tort? After 20 tumultuous years during which it has been pleaded and occasionally argued before the courts in Canada, we still do not know much about the rather obscure doctrine. In its most robust formulation, waiver of tort allows a plaintiff who is able to prove all the constituent elements of a traditional wrong *except* that they have suffered a loss as a consequence of the defendant's breach to

"waive" tort compensation and claim only "disgorgement"[1] of the profits that the defendant earned as a result of the wrongful behaviour. As the Supreme Court of Canada noted in its decision in *Pro-Sys Consultants Ltd v Microsoft Corporation*[2] ("*Microsoft*"): "[a]n action in waiver of tort is considered by some to offer the plaintiff an advantage in that it may relieve them of the need to prove loss in tort, or in fact at all".[3]

The doctrine is based on the intuitively appealing notion, deeply rooted in equity, that a defendant ought not be able to profit from wrongdoing. It also bears on deterrence: equitable waiver allows a plaintiff who has not suffered from harm to perform a corrective role in depriving a wrongdoer of profits, disincentivizing antisocial behaviour. But where a class of persons have suffered harm as a result of a mass wrongdoing, one might ask, why would they give up what may well be the overwhelming bulk of their claim?

One answer is straightforward: a waiver of tort claim "may be the easiest cause of action to prove",[4] because disgorgement flows from the wrongdoing of the defendant, rather than the harm caused to the plaintiff. Under its theory, it is enough for a plaintiff to show that the defendant behaved in a way that *was wrong* — usually that it breached a duty somewhat "at large" or generic — but the plaintiff need not establish *a wrong* — that is, there is no requirement that all the elements of a complete tort be present. So, if a manufacturer produced a dangerously defective product, or failed to provide a necessary warning to its customers, or if an issuer of shares deliberately or negligently misrepresented facts in a prospectus, or if a factory breached pollution standards and exposed its neighbours to risk, the defendant could lose whether or not the plaintiff

1. The cases and literature on waiver of tort often use the language of a number of equitable remedies — "accounting", "disgorgement" or "constructive trust". But no matter how it is cast, the effect is the same: some amount equivalent to the defendant's ill-gotten profits will be calculated and surrendered to the plaintiff or class.
2. 2013 SCC 57 [*Microsoft* SCC].
3. *Ibid* at para 93.
4. Paul Perell, "Field Notes on Products Liability Claims in Class Actions" (2011) 38:2 Advocates' Quarterly 149 at 163 (describing possible advantages in the context of products liability claims).

could prove a direct connection between the wrongdoing and his or her loss. This is the attraction of the resort to equity: it is often a very difficult task to prove the necessary causal link, particularly where the loss is indirect, or issues of reliance or scientific uncertainty are in play in an individual case.

When it first appeared on the class action scene in the 2004 decision of *Serhan Estate v Johnson & Johnson*[5] ("*Serhan*"), waiver of tort was welcomed as a panacea for class claimants, allowing them to vault over many of the traditional rules of tort that barred recovery even where defendants had miserably failed in their duties of care or had been shown to have been fraudulent or malevolent. Waiver of tort claims, generally pleaded in the alternative, became a routine feature of class action pleadings,[6] and as they began to trickle before the courts, the class action bar held its collective breath.

It was a long wait. In the intervening two decades, judicial skepticism, or at least ambivalence, had calcified into what appears to be a trend of qualified rejection, and it appeared to some observers that the doctrine was a dead letter.[7] However, in the recent certification decision of *Microsoft*, a unanimous Supreme Court of Canada refused to strike a waiver of tort claim and permitted it to proceed to trial, holding that it was not "plain and obvious" that it could not succeed.

Judicial reluctance, however, remains in the wake of *Microsoft*, as exemplified by the subsequent decision of *O'Brien v Bard Canada Inc*[8] ("*O'Brien*"), where Justice Perell, while allowing that disgorgement through waiver of tort may be a suitable remedy in some cases, found it wildly *inappropriate* for a mass tort premised on personal injury with

5. (2004), 72 OR (3d) 296 (Sup Ct J) [*Serhan* Sup Ct], aff'd (2006), 85 OR (3d) 665 (Div Ct) [*Serhan* Div Ct].
6. See Perell, *supra* note 4 at 160-63.
7. Charles Murray, "An Old Snail in a New Bottle? Waiver of Tort as an Independent Cause of Action" (2010) 6:1 Canadian Class Action Review 5; H Michael Rosenberg, "Waiving Goodbye: The Rise and Imminent Fall of Waiver of Tort in Class Proceedings" (2010) 6:1 Canadian Class Action Review 37; and Kit Scotchmer, "Waiver of Tort: A Potential Sea Change in Class Action Law" (2011) 7:1 Canadian Class Action Review 159.
8. 2015 ONSC 2470 [*O'Brien*].

substantial potential damages.

In this article I will argue that the reluctance of Perell J to permit plaintiffs to rely on waiver of tort was well placed, even if the doctrine might have some residual utility. However, the problem that it sought to address still looms unbearably large in the legal landscape: what to do when a mass wrong has clearly been committed, where people have been harmed, but where it is difficult or impossible to identify the actual victims, or the degree of their individual loss, with any precision, if at all?

I suggest here that large-scale claims premised on widespread defendant wrongdoing can be usefully broken into two categories, only one of which is truly amenable to the application of waiver of tort. I conclude that, in cases where a defendant has committed a wrong of a type that is so serious that deterrence is called for even when it does not produce legally-cognizable harm (and particularly when the wrong is, by its nature, elusive of ordinary damages claims), waiver of tort can provide an important behaviour modification device where no other is available. This category might include criminal acts, some regulatory offences, and intentional common law or equitable wrongs such as fraud, deliberate misrepresentation or a wanton or otherwise egregious flouting of legal duties.

However, with respect to most true mass torts with individual causation issues, where it is known that the defendant was negligent and that this *had* caused harm but it is not knowable with certainty which victim's injuries could be attributed to the wrong as opposed to other causes, waiver of tort is not the preferable analysis when compared to other devices that are (or should be) available as a matter of tort law. In particular, I propose, the most appropriate solution in most cases is an aggregate or global treatment of causation issues: treating causation of harm as something that occurs in a *population* of persons, rather than a collection of *individuals*.

My argument here proceeds as follows. In Part II, I describe the doctrine and its origins, and trace its recent somewhat lurching progress through 20 years of Canadian class action jurisprudence. In Part III, I identify the built-in limitations of the doctrine and other problems that would be inherent in its widespread adoption. In Part IV, I describe the

principal alternative: the assessment of causation in mass tort cases on an aggregate, rather than individual basis, and suggest why, in such cases, it is preferable to waiver of tort. In my conclusion I outline and defend the classification of large-scale claims into the two categories I have earlier mentioned, only one of which should be dealt with through recourse to waiver of tort.

I argue here that resort to equitable waiver is not the preferable solution to the problem of indeterminate causation or a lack of nexus otherwise between wrongdoer and victim, for three reasons.

First, divorcing damage from wrongdoing *altogether* risks distorting the historical role of tort by removing an important (and occasionally maligned, including by me) limiting principle: restricting recovery through a nexus of harm-causation prevents the emergence of a purely-regulatory civil litigation regime where busybody plaintiffs and their lawyers are incentivised to ferret out even harmless wrongdoing, leading to a costly, inefficient and wasteful court "policing" of the economy that does little to advance the role of the civil courts as fora for the vindication of aggrieved victims. Having identified this problem, though, I suggest that it is not insurmountable and can be addressed through a principled application of the rules of standing.

A second and more serious concern regarding waiver of tort in class actions is not based on a fear that they could accomplish too much, but rather that they do too little, too easily. I will argue that where the wrongdoing has actually caused harm, disgorgement of profits is inferior to tort damages as a device of either compensation or deterrence, especially when measured against other possible innovations available to the courts.

This leads to my third and final argument for disfavouring the waiver approach in most cases: although it may be argued that "some compensation and deterrence is better than nothing", reliance on waiver of tort risks creating a schism between the interests of the class and the public on one hand (who would want the defendants to pay the full cost of harm, not just of profits) and class counsel on the other. This is because a waiver of tort claim would provide "low hanging fruit" for lawyers, who would be incentivised to pursue a high volume of "quick and dirty"

settlements or judgments rather than seeking to maximise recovery for their clients.

I argue here that the problem of under-deterrence and under-compensation in mass claims where the defendant's wrongdoing is plainly established is very real, and it is something that the substantive law of tort should, and can, accommodate through rules of causation specific to mass torts, and in particular class actions. This requires that courts treat mass torts as fundamentally distinct from individual claims, as harm in populations, rather than individuals. This is an idea that underlay (or at least should have underlain) the attempts to utilize waiver of tort, but to put it simply, tort law can do it better. In this respect, I will suggest that very recent decisions of the Quebec Superior Court of Justice and the Ontario Superior Court provide us an intriguing glimpse of the future.

II. Waiver of Tort in Canadian Law

A. The Difficult Doctrine

In introducing the concept of waiver of tort it is difficult to do better than quote the description by Justice Epstein of the Ontario Divisional Court in *Serhan*:

> [i]ts origin lies in the expression "waiver of tort and suit in assumpsit", the latter being the historical antecedent of many modern common law "quasi-contract" restitutionary claims. In invoking waiver of tort, the plaintiff gives up the right to sue in tort and elects to base the claim in restitution, thereby seeking to recoup the benefits the defendant has derived from his wrongful conduct. The practical purpose behind it is that in certain situations, where a wrong has been committed, it may be to the plaintiff's advantage to seek recovery of an unjust enrichment accruing to the defendant rather than normal tort damages.[9]

Justice Epstein went on to quote Peter D Maddaugh and John D McCamus in *The Law of Restitution*:

> [t]he doctrine known as "waiver of tort" is perhaps one of the lesser appreciated areas within the scope of the law of restitution. From the outset, it seems to have engendered an undue amount of confusion and needless complexity. The almost mystical quality that surrounds the doctrine is attested to by the following famous couplet penned by a pleader of old [[J.L. Adolphus, "the

9. *Serhan* Div Ct, *supra* note 5 at para 50.

> Circuiteers– An Eclogue" (1885) 1 L.Q Rev. 232, at p. 233]:]: "[t]houghts much too deep for tears subdue the Court when I assumpsit bring, and god-like waive a tort". One source of this confusion stems from the doctrine's very name. As one writer has pointed out, not entirely facetiously, it has "nothing whatever to do with waiver and really very little to with tort".[10]

Historically, waiver of tort was restricted to a narrow and discrete class of "predicate wrongs", cases involving conversion, detinue, trespass to chattels and deceit.[11] However, as *Serhan* (a products liability case) and its progeny suggest, and the Supreme Court of Canada's decision in *Microsoft* (a competition class action) confirms, its possible application may be almost limitlessly broad.

Waiver of tort is not compensatory, and in this sense it is distinct from an equitable claim of unjust enrichment, which requires not only an unlawful profit by the defendant but also a corresponding loss by the plaintiff,[12] a transactional relationship which has historically been required to be quite direct.[13] The waiver doctrine emerges from the basic equitable idea that holds (as more or less a standalone principle) that a wrongdoer should not be permitted to retain ill-gotten gains.

Much of the skirmishing around waiver of tort in class claims has so far centred on whether the doctrine's availability in a particular claim or generally should be decided at the preliminary certification stage, where

---

10. *Ibid* at para 51, citing Peter D Maddaugh & John D McCamus, *The Law of Restitution* (Aurora: Canada Law Book, 2005)(loose-leaf) at 24-1.
11. Paul Perell, "Serhan Estate v. Johnson & Johnson and Soulos v. Korkontzilas: Something Is Happening Here and We Don't Know What It Is", Case Comment, (2007) 33:1 Advocates' Quarterly 375 at 375-76.
12. The well-known elements required to establish an unjust enrichment are: (i) an enrichment of the defendant; (ii) a corresponding deprivation of the plaintiff; and (iii) an absence of juristic reason (such as a contract) for the enrichment: see *Alberta v Elder Advocates of Alberta Society*, 2011 SCC 24 at para 82 [*Alberta Elders*]; *Garland v Consumers' Gas Co*, 2004 SCC 25 at para 30; *Rathwell v Rathwell* [1978] 2 SCR 436 at 455; *Pettkus v Becker* [1980] 2 SCR 834 at 848.
13. But see *Microsoft* SCC, *supra* note 2 at para 87 (where the Supreme Court of Canada permitted an unjust enrichment pleading to survive the certification stage, notwithstanding that it was brought by indirect purchasers).

the test is the low threshold of "plain and obvious" and where novelty, the Supreme Court of Canada has held, cannot be a necessary bar to a legitimate claim lest it become a barrier to the progress of tort law through the discovery of new or evolved causes of action.[14] The Supreme Court of Canada seemed to settle this aspect of the controversy when it upheld the certification of the waiver claim advanced in *Microsoft*.[15] But the *Microsoft* decision, and those other "demurrer" cases that came before and since, put off the central questions to another day: if it exists as an independent cause of action, what would it look like? Clearly it must be premised on wrongdoing by the defendant. But could any kind of wrongdoing found recovery? And recovery by whom?

B. The Jurisprudence

It was the 2004 decision of Justice Cullity in *Serhan*, upheld by a majority of the Divisional Court two years later, that firmly fixed waiver of tort in Canadian legal consciousness. *Serhan* was a products liability class action brought by users of a device used to monitor blood glucose levels.

---

14. *Alberta Elders*, *supra* note 12 at para 20 (noting that the first criterion for certification of a class action is that the plaintiff's pleading must disclose a cause of action); and *Hollick v Toronto (City)*, 2001 SCC 68 at para 25 (where the Supreme Court of Canada confirmed that this requirement is assessed on the same basis as a motion to dismiss, as set out in *Hunt v Carey Canada Inc* [1990] 2 SCR 959 at 980: the question is, assuming all facts in the Statement of Claim are true, whether it is "plain and obvious" that the plaintiff's claim cannot succeed).

15. However, see *O'Brien*, *supra* note 8 (where Perell J suggested that Cromwell J's *dicta* should not apply in a products liability case, and he rejected waiver of tort as a legitimate cause of action in the facts before him: "*Pro-Sys Consultants Ltd.* was a competition law action. The case at bar is a products liability tort case. For decades, going at least as far back as *Donoghue v. Stevenson*, [1932] AC 562 (HL), and continuing to this day, courts have determined matters of policy in tort claims at the pleadings stage and if it were necessary to do so I would decide whether waiver of tort is a cause of action and, if it is a cause of action I would decide whether it is a viable cause of action for a products liability proposed class action so as to satisfy the cause of action criterion of certification" at para 158).

The device was admittedly defective, and, possibly as a consequence, harmful. But it appeared likely that none of its thousands of users in Canada could show any harm that could be attributed to the defect, and they had all received it for free. Class members seemed to have no real claim, in other words, under traditional tort or contract law, or under any statutory regime. Nevertheless, Cullity J found that the plaintiffs had pleaded material facts sufficient to support a claim under waiver of tort, notwithstanding that it had not been specifically invoked. On appeal, Epstein J of the Divisional Court allowed certification to proceed on the equitable claim, and the case settled in 2011 with the application of waiver of tort still deeply in doubt.[16]

Following *Serhan*, a number of Ontario decisions permitted waiver of tort claims to proceed through the certification, based primarily on the idea that the operation of the doctrine was not so settled so as to make such claims certain to fail on the "plain and obvious" standard.[17] A number of other decisions disposed of waiver of tort claims on the basis

16. *Serhan v Johnson & Johnson*, 2011 ONSC 128, (The settlement provides replacement devices for diabetics, and educational programs and other methods of *cy-près* distribution, totalling $4.5 million. In approving the settlement, the Court recognized that the case likely turned on the waiver of tort issue. In discussing the "likelihood of success" criterion for determining the reasonableness of the settlement, Horkins J wrote: "[m]ost importantly, does waiver of tort exist as an independent cause of action or is it only a remedy applicable to another tort? This difficult question is at the heart of this case. While Ontario Class Counsel were confident that a court would find that it was an independent cause of action, there was a considerable risk that it would not" at para 69).
17. *Griffin v Dell Canada Inc*, 2009 CanLII 3557 (Ont Sup Ct J); *Haddad v Kaitlin Group Ltd*, 2008 CanLII 66627 (Ont Sup Ct J); *Heward v Eli Lilly & Co* (2007), 39 CPC (6th) 153 (Ont Sup Ct J), aff'd (2008), 295 DLR (4th) 175 (Ont Div Ct); *Pollack v Advanced Medical Optics Inc*, 2011 ONSC 1966; *Peter v Medtronic Inc* (2009), 83 CPC (6th) 379 (Ont Sup Ct J), aff'd 2010 ONSC 3777; *Tiboni v Merck Frosst Canada Ltd* (2008), 295 DLR (4th) 32 (Ont Sup Ct J), leave to appeal refused (2008), 304 DLR (4th) 220 (Ont Sup Ct J).

that the operation of the doctrine, if applicable, was moot.[18] In each of these latter cases the claim was defeated either at a preliminary stage or after trial because no predicate wrong to support waiver of tort had been established (that is to say, not only that there was no completed tort, but that there was also no breach of duty or other illegality on which to found the waiver).

There was a somewhat different, but no more conclusive, outcome in *Koubi v Mazda Canada Inc*[19] ("*Koubi*"), where the British Columbia Court of Appeal decertified a waiver of tort class action on the basis that the predicate wrong alleged was breach of statutory standards for which the legislation itself provided comprehensive and exhaustive remedies — essentially finding that, at least with respect to those Acts, the statutory remedies had displaced all other private modes of redress, including waiver of tort.

Koubi was indicative of a trend whereby British Columbia's courts took, overall, a somewhat more restrictive view of waiver of tort than their Ontario counterparts, rejecting certification of claims in *Reid v Ford Motor Company*[20] ("*Reid*"), and *Strata Plan LMS 3851 v Homer Street Development Limited Partnership*[21] on the basis that "anti-harm" wrongs such as negligence could not provide the predicate breaches to found the restitutionary remedies available under the waiver of tort doctrine. In *Reid*, Justice Gerow cited *Networth Industries Ltd v Cape Flattery*,[22] for the proposition that unjust enrichment could not be founded on negligence, and therefore it was plain and obvious the claim for waiver of tort was similarly bound to fail. She held:

> [r]estitutionary claims are not made in negligence and nuisance because they are in the main "anti-harm wrongs" in relation to which it is impossible, even if they lead to an enrichment of the wrongdoer, to elevate the prevention of

18. *Aronowicz v Emtwo Properties Inc*, 2010 ONCA 96 (shareholder dispute); *Arora v Whirlpool Canada LP*, 2012 ONSC 4642 at para 300 (allegedly defective laundry machines); *Andersen v St Jude Medical Inc*, 2012 ONSC 3660 (allegedly defective heart valves).
19. 2012 BCCA 310.
20. 2006 BCSC 712 [*Reid*].
21. 2011 BCSC 569.
22. [1997] BCJ No 3174 (SC) [*Networth*].

> enrichment to the level of a primary purpose.[23]

The British Columbia Court of Appeal was somewhat more generous with anti-competition claims, permitting waiver of tort pleadings to survive certification in *Pro-Sys Consultants Ltd v Infineon Technologies AG*.[24] Justice Smith, writing for the Court, allowed that the plaintiffs might not need to show damage if the doctrine were applied, and that an aggregate monetary award could be certified as a common issue. Similarly, in *Steele v Toyota Canada Inc*,[25] Justice Hinkson (as he then was) permitted a waiver of tort claim premised on breach of the provincial competition legislation to proceed, also acknowledging that if waiver of tort were an independent cause of action, proof of caused damage may not be necessary in order for a global remedy to be available.

In *Microsoft*, the plaintiffs were indirect purchasers of Microsoft's operating systems who claimed that the software giant had conspired to fix prices. Two judges at first instance had certified a claim including waiver of tort;[26] the majority of the Court of Appeal overturned certification without expressly addressing that cause of action, simply holding that indirect purchasers of Microsoft's product had no competition claim (only the dissenting judge, Justice Donald dealt with it, finding that it did disclose a cause of action).[27] On appeal to the Supreme Court of Canada, the certification was reinstated, including waiver of tort as a possible cause of action.[28]

There is one decision since *Microsoft* that deserves more than mention. *O'Brien* was a products liability and failure to warn class action involving the manufacturers of pelvic mesh implants. The claim involved 19 different products and thousands of class members, who were alleged to have suffered one or more of a host of complications and injuries

---

23. *Reid*, *supra* note 21 at para 29, citing *Networth*, *ibid* at paras 24-26.
24. 2009 BCCA 503, leave to appeal refused, [2010] SCCA No 32.
25. 2011 BCCA 98.
26. *Pro-Sys Consultants Ltd v Microsoft Corporation*, 2010 BCSC 285; *Pro-Sys Consultants Ltd v Microsoft Corporation*, 2006 BCSC 1047; and *Pro-Sys Consultants Ltd v Microsoft Corporation*, 2006 BCSC 1738.
27. *Pro-Sys Consultants Ltd v Microsoft Corporation*, 2011 BCCA 186.
28. *Microsoft* SCC, *supra* note 2.

as a result of the implants. Justice Perell, while allowing that waiver of tort might be viable as a cause of action in some cases, said that where mass tort claimants were pursuing possibly billions of dollars in personal injury claims it was not a viable alternative pleading. I will have more to say about Perell J's concerns below.

So what is the state of the law in Canada today? Well, waiver of tort, like Professor Schrödinger's famous cat, presently appears to be both alive and dead (or perhaps it is better to say neither alive nor dead) pending an examination.[29] In order to make out a claim it would be necessary for the plaintiff to show a wrongdoing of the defendant (predicate wrong), and a profit that has accrued to the defendant from the activity that was unlawfully conducted. These conditions may in turn be subject to the overriding objectives of equity itself.

At virtually every step of the way, lawyers representing defendants in class actions have loudly — and as it turned out somewhat prematurely — declared that waiver of tort, as an independent cause of action, was dead or at least dying.[30] But it has proven stubbornly resilient, and I suggest that this is because the problem that it is being used to remedy is real: what happens when the plaintiffs can establish (i) that the defendant

29. Physicist Erwin Schrödinger proposed a thought experiment in 1935 in which decaying radioactive material would either kill a hidden cat or not. According to quantum theory of superposition, the cat would be both alive and dead until it was observed to be one or the other, a result Schrödinger regarded as absurd.

30. See *e.g.* Brandon Kain, "Waiver of Tort Gets a Reality Check at the B.C. Court of Appeal" (23 July 2012), *Canadian Appeals Monitor*, online: <www.canadianappeals.com/2012/07/23/waiver-of-tort-gets-a-reality-check-at-the-b-c-court-of-appeal/>; Christopher Naudie et al, "Class Actions Development: B.C. Court of Appeal Slams the Door on Waiver of Tort in Statutory Cases" (20 July 2012), *Osler, Hoskin & Harcourt LLP*, online: < https://www.osler.com/en/resources/critical-situations/2012/class-actions-development-b-c-court-of-appeal-sl>; Peter Kryworuk & Rebecca Case, "Waiver of Tort — So Long, It's Been Nice to Know You?" (2 September 2015), *Lerners* (blog), online: <http://lernersclassactionplaintiff.ca/blog/post/waiver-of-tort-so-long-it-s-been-nice-to-know-you?utm_source=Mondaq&utm_medium=syndication&utm_campaign=View-Original>.

committed a breach of some statutory or common law duty and; (ii) that the breach caused a certain amount of harm in persons who have been exposed through purchase or use of a product; but (iii) it is difficult or impossible to link, on a balance of probabilities, any *particular* victim with the wrong?

If waiver of tort exists, what are the elements in modern Canadian law? What constitutes a predicate wrong? Does negligence qualify? Some breaches of statute (competition laws, for instance) seem to be covered, where others that contain exhaustive recovery regimes (consumer protection legislation) may not be. Who can claim under the doctrine, if this is not to be determined by who has been harmed by the wrong? These questions might eventually be answered with reference to the equitable principles underlying the cause of action. So we might ask, first, whether the wrong is such "that in equity and good conscience [the] defendant should not be permitted to retain that by which it has been enriched";[31] and second, whether there is some connection or nexus between the plaintiff (or class) and defendant such that the former may equitably pursue and receive the benefit of the disgorgement. But at this point all of these matters remain unresolved.

I should conclude my discussion of the Canadian cases by disposing of one issue that has caused some angst among both academics and jurists: is waiver of tort a stand-alone cause of action, or is it simply a remedy?[32] If the latter, the doctrine would only permit disgorgement where a complete wrong (usually including proof of caused loss) is independently established.

As the Supreme Court of Canada noted in *Microsoft*: "[t]he U.S. and U.K. jurisprudence as well as the academic texts on the subject have largely rejected the requirement that the underlying tort must be

31. *Federal Sugar Refining Co v US Sugar Equalisation Board*, 268 F 575 (SDNY 1920 (US)) at 582 [*Federal Sugar*].

32. The controversy is well described in *Serhan* Div Ct, *supra* note 5 at paras 45-76, and I need not detail it here.

established in order for a claim in waiver of tort to succeed".[33] The Court appeared less persuaded by other cases restricting the doctrine to cases where the full tort, including proof of loss, was required[34] and permitted the waiver claim to proceed, apparently on the assumption that it might provide an independent basis for recovery, notwithstanding Microsoft's assertion that it had been pleaded only as remedy.

But in any event, the Supreme Court's analysis accords with a practical reality: if providing a remedy (disgorgement of profits) is *all* waiver of tort does, then it is a weak doctrine in most cases. Indeed it would be largely redundant with punitive damages, which courts recognize can, and perhaps often should, be measured by the profits of the defendant, with the equitable and deterrence objectives of depriving the defendant the fruits of his wrongdoing.[35]

It is difficult to see how equitable disgorgement, as a remedy, could improve on this, and indeed it appears far less efficacious than punitive damages which can be increased or decreased according to other relevant factors, including ensuring proper deterrence.

So if waiver of tort is to mean anything, it must be an independent cause of action, one that can succeed where ordinary tort or contract claims cannot, and the remainder of my analysis of its efficacy and desirability is premised on this, more robust, view of the doctrine.

---

33. *Microsoft* SCC, *supra* note 2 at para 96, citing *inter alia*, *National Trust Co v Gleason*, 77 NY 400 (App Ct 1879 (US)); *Federal Sugar*, *supra* note 31; *Mahesan v Malaysia Government Officers' Co-operative Housing Society Ltd*, [1979] AC 374 (PC (Malaysia)); *Universe Tankships Inc of Monrovia v International Transport Workers Federation*, [1983] 1 AC 366 (HL).
34. *Microsoft* SCC, *supra* note 2 at para 96, citing *United Australia Ltd v Barclays Bank Ltd*, [1941] AC 1 (HL) at 18; *Zidaric v Toshiba of Canada Ltd* (2000), 5 CCLT (3d) 61 (Ont Sup Ct J) at para 14; *Reid*, *supra* note 20.
35. *Whiten v Pilot Insurance Co*, 2002 SCC 18 (where the Supreme Court of Canada held "it is rational to use punitive damages to relieve a wrongdoer of its profit where compensatory damages would amount to nothing more than a licence fee to earn greater profits through outrageous disregard of the legal or equitable rights of others" at para 72).

III. What is Wrong With a Waiver of Tort Claim

A. What is the Problem, Exactly?

I propose here that there are two categories of cases where courts might apply the doctrine. They suggest two different problems, and the fact that waiver of tort is good at addressing one and poor at the other leads to my conclusion that it should not be a preferred doctrine in most mass-wrong cases.

The first category consists of cases where a defendant has profited from a wrong, but there has been no loss to the plaintiff giving rise to an ordinary cause of action. In such circumstances, waiver of tort can provide a measure of deterrence of activity that is potentially harmful and antisocial but did not cause harm "this time around". Whether any particular wrong should give rise to disgorgement is a question rife with policy considerations. Presumably, the intervention of the courts is warranted to deter wrongs of a type which society has an interest in absolutely prohibiting (such as a crime), instead of just regulating (such as, at least on one view, negligence). But even after this question is answered, the court might further ask: is the plaintiff or class the appropriate "prosecutor" of such an action, given that it is essentially regulatory in ambition?

The second category of cases to which waiver might apply is where we know the defendant has committed complete torts — that is, we know *some* people have been harmed by the wrong but we simply cannot tell which ones. Typically, this will arise in the context of mass claims that can be prosecuted through the device of the class proceeding. In such cases, I will argue, resort to waiver of tort is inappropriate and ineffective, and a much more straightforward solution — the determination of causation class-wide, in a population of persons exposed to the risk of harm — is available and preferable.

In order to justify my arguments in this regard, I must review the objections that are taken to waiver of tort as a cause of action. I hope to show that the problems with the doctrine are most acute in the second category of cases (*i.e.* where harm has in fact occurred but is indeterminate), and are insignificant, or at least manageable, in the first

category.

B. Wrongs "In the Air"

We know that crimes, fraud and other malfeasance can form the basis of a waiver of tort claim. But what of negligence-based cases, such as products liability or toxic tort claims? Should different principles apply? Some courts have expressed discomfort, at least in *obiter*, that waiver of tort claims could relate to "anti-harm" wrongs like negligence, which are overtly premised on compensatory rather than regulatory principles.[36] Causation — not just factual but *legal* causation — has proven to be an important limiting device in negligence law, and not just from the point of view of compensation objectives. Tort scholars might disagree on the *reason* why the right to pursue negligence claims has been limited to those whom the defendant has harmed through its fault, but it certainly has. And it may be that disconnecting the causation link altogether will unnecessarily depress socially or economically useful activity as "busybody plaintiffs" (or more likely lawyers) set themselves to ferreting out "wrongdoing" and launching a wave of litigation with no good purpose.

So there are good arguments why waiver of tort claims have never been, and should not be, prosecuted entirely "at large": that is, a person who is a complete stranger to the defendant and the wrong ought not be able to sue to obtain the defendant's profits. On this view, there should be *some* relationship between the wrong and the plaintiff sufficient to permit the plaintiff to obtain standing on the basis that the plaintiff is in an equitable position to pursue the claim.

I would suggest this nexus could be established on a couple of different bases. First, if the plaintiff (or the class) was "within the ambit of the risk" of harm created by the defendant's wrong,[37] they could

36. See for instance *Reid*, *supra* note 20 at paras 15, 29; *Serhan* Div Ct, *supra* note 5 at paras 66-67.

37. Individual plaintiffs who cannot prove causation but who fall within "the ambit of the risk" may be able to take advantage of exceptions to the traditional "but for" test in individual causation cases: *Resurfice Corp v Hanke*, 2007 SCC 7 at para 25 [*Resurfice*].

claim. So, for instance, consumers of a product, purchasers of shares, or victims of diseases epidemiologically linked with a pollution source or toxic substance might qualify as appropriate plaintiffs in a waiver of tort claim, where mere bystanders or busybodies might not. Thus in *Serhan*, for instance, it made sense to permit the purchasers of the medical device to bring the (eventually settled) waiver of tort claim. Another way of assessing the proposed plaintiff might be whether there will be some remedy (or settlement term) aside from simple disgorgement that will benefit the plaintiff. So, to again use *Serhan* as an example, each user of the diabetes testing device was, as part of the settlement, entitled to a replacement device.

In many class action cases, the "busybody" problem will not arise because waiver is almost always pleaded in the alternative to tort, fraud or unjust enrichment claims, and so the class is defined with regard to persons who have suffered harm as the result of the wrongs. If those pleadings survive certification, which would require that the tort claims have, at least "some basis in fact",[38] then this might provide a sufficiently restrictive class who should have standing to pursue waiver claims in the alternative.

Nevertheless, the separation of wrongdoer from victim that waiver of tort entails must be counted among the difficulties to a widespread use of the doctrine.

C. Under-Compensation and Under-Deterrence

Another obvious but more serious problem with resort to waiver of tort emerges in cases where we can know that the defendant's wrongdoing did in fact cause some harm. In some cases, application of the doctrine might represent a windfall for plaintiffs, who might have suffered little or no loss from a wrong associated with a highly profitable product or activity of the defendant. In many other cases, the harm will be of such a magnitude that any disgorgement of profits will be little more than

38. Brandon Kain, "The Supreme Court of Canada and the Still-Curious Requirement of 'Some Basis in Fact'" (2015) 68 Supreme Court Law Review (2d) 77.

token.

The main objectives of tort law are usually expressed in terms of compensation and deterrence. Though scholars will disagree about which has primacy, the question is usually moot: tort law operates by making the defendant pay the cost of the harm it has caused to the person it has injured. This achieves compensation for the victim and forces the defendant to internalize the cost of the harm, thus providing the economically-optimal level of deterrence, or regulation, of the risky activity.[39]

Disgorgement, as a sole remedy, upsets this balance. Certainly, in cases where the plaintiff has suffered little or no harm, the disgorgement of profits provides some disincentive to defendants to engage in risky or other antisocial behaviour. So we see cases like *Strand Electric and Engineering Co Ltd v Brisford Entertainments Ltd*,[40] where Lord Denning observed that in "[the] cases where the defendant has obtained a benefit from his wrongdoing he is often made liable to account for it, even though the plaintiff had lost nothing and suffered no damage".[41]

But reliance on disgorgement as a deterrent is inferior to a system based on damages, particularly where "habitual defendants" may be systemically incentivised to take risks in the pursuit of profit. If all a defendant risks are its profits, it might, overall, take *more* risks to obtain *more* profits, on the expectation that there is no real "downside" to doing so. If a wrongdoer gets caught, it gains nothing but loses nothing either. If it does not get caught, it retains the benefit of the wrong. If it internalizes the true costs of the harm it has caused, on the other hand, the equilibrium provided by tort law is restored.

So in cases where the defendant's wrong has *not* actually caused harm

39. See *e.g.* Guido Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (New Haven: Yale University Press, 1970)(the still-resilient "law & economics" analysis of tort law); see also Richard A Posner, "A Theory of Negligence" (1972) 1:1 Journal of Legal Studies 29; Richard A Posner, *Economic Analysis of Law*, 6d (New York: Aspen Publishers, 2003); Richard A Posner, "Utilitarianism, Economics, and Legal Theory" (1979) 8:1 Journal of Legal Studies 103.
40. [1952] 1 All ER 796 (CA).
41. *Ibid* at 800.

"this time around", waiver of tort and disgorgement provides a rational means of providing some deterrence for activity that is inherently antisocial or risky, but did not cause harm in the particular event (such as in *Serhan* itself).[42] The difficulty is that, in the majority of cases, *some* harm *has* actually been caused; the problem is that the connection between the defendant's wrong and each member of a plaintiff class cannot be confidently established.

D. The Divergence of Interests of Class Counsel

There is one further problem with reliance on waiver of tort that may lead to systemic under-deterrence and under-compensation. That is that recovery under the doctrine is so easy that it may actually prevent legitimate damages claims from being brought, or at least from being aggressively pursued.

The success of class actions depends on the interests of class counsel being aligned with the class members' own. The fact that this will not always be so has influenced many aspects of the class procedure, such as the requirements that class members be given notice of proposed settlements and an opportunity to object, and that settlements be approved by the court. The temptation is always that plaintiffs' counsel can, implicitly at least, collude with defendants to produce a quick settlement for a small amount that provides substantial payment for the lawyers but little

42. It is tempting to say, as a consequence of the distinction I propose here (between cases where there is no damage versus true "indeterminate causation" cases where damage is known but particular victims cannot be certainly identified), that waiver of tort is never required in class actions, because in the cases where I propose that it should be available could be as easily pursued as individual actions. However this is probably too simplistic a view for three reasons: (i) because waiver of tort may legitimately be pleaded in the alternative in cases where it is not certain into which category the claim should properly fall; (ii) because it may be more fair or just to distribute the disgorgement more broadly than to a single individual; and (iii) because a class action will dispose simultaneously of the claims of all persons who would have standing to pursue them individually, therefore avoiding an inefficient multiplicity of competing claims for the same remedy.

benefit for the members.[43]

In *O'Brien*, referred to earlier, Perell J found that even if it were a valid cause of action conceptually, waiver of tort should not be permitted to proceed to trial on the facts before him, not because it did too much but because it did too little:

> [i]n *Serhan v. Johnson & Johnson,* which is the case that started the debate about the nature of waiver of tort, there were zero monetary damages for the tort claim and waiver of tort was the route to access to justice and behaviour modification. In the case at bar, assuming Bard were negligent, a waiver of tort cause of action would not provide access to justice to class members or any meaningful behaviour modification. It would be reprehensible for Class Counsel to take a contingent fee based on an award calculated on the disgorgement of profits. A judgment or a settlement based on waiver of tort would create enormous conflicts between Class Members as to how the disgorged funds should be distributed. It would be a waste of the court's and the parties' litigation resources to expend discovery and trial time calculating what profits, if any, Bard made from its Pelvic Mesh Products, when assuming liability, everybody should be spending their litigation resources calculating compensatory damages. In my opinion, in these circumstances, regardless of whether waiver of tort is a reasonable cause of action, it would not be reasonable to prosecute it as a class action. Even if the pleading of waiver of tort satisfied the cause of action criteria, the class definition, and the common issues criteria, in my opinion, the waiver of tort claim in the circumstances of the case at bar would not satisfy the preferable procedure and the representative plaintiff criteria. In these circumstances, I conclude that the waiver of tort claim in the case at bar does not satisfy the cause of action criterion for a class action. I would not certify the waiver of tort claim.[44]

One might quibble with Perell J's mixing of the "cause of action" criterion with the other threshold requirements of class certification, particularly "preferability". And his concern over the problems of distribution of a class-wide award, which he saw as reflecting on the question of whether

43. See generally Bruce Hay & David Rosenberg, "'Sweetheart' and 'Blackmail' Settlements in Class Actions: Reality and Remedy" (2000) 75:4 Notre Dame Law Review 1377.

44. *O'Brien*, *supra* note 8 at paras 162-65.

the representative plaintiff was appropriate, may have been overstated.[45] But at least this one of his concerns over the use of the doctrine should cause serious reflection: he worried that the class in *O'Brien* would be "sold out" if their recovery were limited only to recovery of the profits when the harm alleged was of a much higher magnitude.

When we come to appreciate that plaintiffs' counsel in the class action bar approach their work from the point of view of investment and return, we can apprehend that there will come a point where a reliance on waiver of tort actually undermines the compensatory and behaviour-modification objectives of both tort law generally, and class actions in particular.

In the popular book *Freakonomics*, the authors identified a structural conflict of interest between realtors and their clients. The difficulty arises from the commission structures adopted by the industry: a realtor paid on commission has a comparatively small interest in maximising his client's selling price: that is, a realtor being paid 3% commission on the sale of a house worth $500,000 "loses" only $1500 if she quickly sells the house for a $50,000 discount. It will generally be easier to invest the realtor's efforts in selling two such houses quickly and cheaply (earning $27,000 commission) rather than taking the same time to sell one house for its full value (earning only $15,000 commission), and the authors suggested that a study showed that is exactly what realtors did.[46]

---

45. *Ibid* at paras 125-26 (the distribution problem Perell J identified would apply to any lump-sum award, including punitive and exemplary damages in tort. The deeper problem was the plaintiff's attempt to certify a single class for 19 distinct products, each with its own history and each with a different group of alleged victims).

46. Stephen J Dubner & Steven D Levitt, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (New York: Harper Collins, 2005) at 5-8, (The methodology was ingenious: the researchers compared realtors' behaviour when listing and selling their own homes versus their clients'. They found that realtors took longer to sell their own properties, and realized higher sale prices. It might be observed as an aside that commission structures that are non-linear, such as those providing 7% on the first $100,000 and 2% thereafter (usually split between the buyer's and seller's agents), are even more perverse in their incentives).

Justice Perell's caution in *O'Brien* seems well placed for the same systemic reasons. If the plaintiffs' counsel can get an easy but small recovery from a waiver of tort claim, why should they pursue lengthy, convoluted tort actions, at comparatively great investment risk? Judge Henry Friendly recognized the inclination in the context of an individual wrong in *Alleghany Corp v Kirby*,[47] when he wrote that a lawyer has "every incentive to accept a settlement that runs into high six figures or more regardless of how strong the claims for much larger amounts may be … [A] juicy bird in the hand is worth more than the vision of a much larger one in the bush".[48]

Yet if lawyers do not pursue the full measure of the harm, the tort "market" is distorted: defendants internalize only a fraction of the harm they have caused, and the class members receive only a fraction of their true losses.

So either of the present approaches — insistence on individual attribution of harm (with recovery denied in each case, even if the claims are economically viable one-by-one), or replacing that with limited recovery based on the profits of the defendant — are chronically unsatisfactory and will demonstrably lead to under-compensation and under-deterrence. Is there a better way?

IV. A Better Way

A. The Problem, Reiterated

As I hope is now apparent, the downsides of resorting to waiver of tort in large-scale claims are most acute in mass torts where it is possible to determine that the defendant's wrong has caused harm, but each plaintiff, or member of a class exposed to the risk of harm, cannot establish a causal link between the wrong and the damage they have suffered.

In a class action, it is possible to assess the harm on a collective basis, as harm caused within a *population* of persons, without the need to prove that any particular class member's harm was the result of a particular

47. 333 F (2d) 327 (2nd Cir 1964 (US)).
48. *Ibid* at 347.

defendant's misconduct. Damages can also be assessed in the aggregate, with only the problem of distribution remaining. As I will mention a bit later, this is often the basis of court-approved class action *settlements*.

The difficulty is that class actions are generally regarded as only a procedural device.[49] Class proceedings statutes generally permit the calculation of quantum of *damages* on an aggregate basis, but only once liability has been established.[50]

This permits defendants to argue that, even if there can be collective determination of "general causation" (*i.e.* that the defendant's wrong can cause the type of harm alleged, or even that, viewed in the aggregate, it did cause harm),[51] the claim cannot be legally made out until it is known which class member actually suffered the harm from the wrong. The Supreme Court of Canada has, from time to time, drank of this water, as when it said that the judge "must still be satisfied on a balance of probabilities that *each* element is present for *each* member".[52] In cases of indeterminate causation, it simply cannot be. As such, in mass tort cases where causation in individuals is indeterminate, tortious harm that can be plainly seen in the aggregate may go unaddressed by the tort system, which provides only, as some have called it, a "phantom remedy".[53]

---

49. *Dell Computer Corp v Union des consommateurs*, 2007 SCC 34 (a class action is "only a legal procedure" at para 106); *Bisaillon v Concordia University*, 2006 SCC 19 (a class action "neither modifies nor creates substantive rights" at para 17).
50. See Ontario's *Class Proceedings Act, 1992*, SO 1992, c 6, s 24(1) [Ont *Class Proceedings Act*]; or identically BC's *Class Proceedings Act*, RSBC 1996, c 50, s 29(1).
51. The same evidence that is used to establish *general* causation - that is that the wrong creates a risk of harm - is evidence for the proposition that it *did in fact* cause harm in a population, though the particular victims might not ever be identified.
52. *Bou Malhab v Diffusion Metromedia CMR Inc*, 2011 SCC 9 at para 53 [emphasis added].
53. William R Ginsberg & Lois Weiss, "Common Law Liability for Toxic Torts: A Phantom Remedy" (1981) 9:3 Hofstra Law Review 859 (examining how the economic realities of individual litigation, combined with causation and other challenges, effectively preclude individual claims for compensation in toxic tort cases).

A decade ago, Professor Jamie Cassels and I proposed that, in mass toxic claims, aggregate assessment of causation may prove not only necessary, but superior to assessments done case-by-case (in the sense that the former is both more accurate and more fair).[54] In 2005, we wrote:

> [c]ausation of harm in the aggregate becomes clearer even as the individual identity of the victims, and their individual connection with each wrongdoer, is lost… in our opinion, viewing inherently probabilistic causation in the aggregate — as a definite harm in a percentage of the population rather than a probabilistic harm in an individual — provides several advantages in the resolution of mass tort claims.[55]

In 2011, while the individual causation case of *Clements v Clements*,[56] was before the Supreme Court of Canada but before it was decided, I reiterated my concern that individual causation rules should not be crafted so as to frustrate mass tort claims:

> [i]n such instances, we know that the defendant has, in fact, caused a certain number of the injuries suffered in the population. We simply do not know which of the afflicted were harmed by the defendant, and which would have been injured in any event. Why should this be an insurmountable obstacle to

54. See for instance, Craig Jones, *Theory of Class Actions* (Toronto: Irwin Law Book, 2003); Jamie Cassels & Craig Jones, *The Law of Large Scale Claims: Products Liability, Toxic Torts and Complex Litigation in Canada* (Toronto: Irwin Law Book, 2005) [Cassels & Jones, *Large Scales*]; Craig Jones, "Reasoning Through Probabilistic Causation in Individual and Aggregate Claims: The Struggle Continues" (2011) 39:1 Advocates' Quarterly 18 [Jones, "Reasoning Through Probabilistic Causation"]; Jamie Cassels & Craig Jones, "Rethinking Ends and Means in Mass Tort: Probabilistic Causation and Risk-Based Mass Tort Claims after *Fairchild v. Glenhaven Funeral Services*" (2003) 82 Canadian Bar Review 597 [Cassels & Jones, "Rethinking Ends and Means in Mass Tort"]. My and Professor Cassels' work in this area was heavily informed by that of David Rosenberg, particularly two foundational articles: David Rosenberg, "Class Actions for Mass Torts: Doing Individual Justice by Collective Means" (1987) 62:3 Indiana Law Journal 561; and David Rosenberg, "Individual Justice and Collectivizing Risk-Based Claims in Mass-Exposure Cases" (1996) 71:1&2 New York University Law Review 210.

55. Cassels & Jones, *Large Scales*, *ibid* at 208.

56. 2012 SCC 32 [*Clements*].

> tort law?[57]

Professor Cassels and I had written these things as a series of cases, especially *Fairchild v Glenhaven Funeral Services Ltd*[58] and *Resurfice Corp v Hanke*,[59] suggested a relaxation of the rules of causation permitting persons "within the ambit of the risk" to bring personal injury claims notwithstanding that they could not demonstrate causation. We posited that these techniques, applied in the context of a class action, could considerably ease the problem of indeterminate causation in mass torts. In this sense, it was disappointing that the Court seemed determined to rein in the idea of "probabilistic causation" in the individual personal injury case of *Clements*.

But it appeared that the Supreme Court was alive to the problem, and might be prepared to relax the rules of causation in mass tort class actions. In *Clements*, the Chief Justice (writing for a unanimous Court on this point), upheld and reiterated the individualistic "but for" test of tort causation in single cases, but then said this:

> [t]his is not to say that new situations will not raise new considerations. I leave for another day, for example, the scenario that might arise in mass toxic tort litigation with multiple plaintiffs, where it is established statistically that the defendant's acts induced an injury on some members of the group, but it is impossible to know which ones.[60]

B. Assessment of Liability, Harm, and Damages on an Aggregate Basis

Let us take the example of a mass tort involving the exposure to a toxic substance or the use of a defective product, where evidence could establish a probability of harm over and above "background risk", and that this harm could be attributable to the defendant's wrong (sometimes called "general causation"). Let us further suppose that we could identify a class of persons who were exposed to the risk (consumers, or people living in

57. Jones, "Reasoning Through Probabilistic Causation", *supra* note 54 at 22 [emphasis in original].
58. [2002] UKHL 22.
59. *Resurfice*, *supra* note 37.
60. *Clements*, *supra* note 56 at para 44.

a geographic area) and whom had suffered harm of a type that would be expected from the wrong. But even if we know that some of that group was actually harmed by the defendant, we also know that others would have suffered the harm without the occurrence of the wrong. We just do not know who is who.

The ordinary tort system would provide no remedy, because attribution of any individual's harm to the defendant's wrong would be impossible.[61] However if we calculated the harm in the aggregate, we could impose liability on the defendant and recover damages in the amount of the harm they had caused. We would still, of course, have a problem of distributing the proceeds among the sufferers of harm. But the regulatory function of tort would be preserved through proper deterrence, and the longstanding tradition that negligence should be regulated by persons who had been harmed and only to the extent that it *has* caused harm, would be respected. Yes, it is only part fulfillment of tort's compensation/deterrence objectives, but it does provide the potential for *some* relief of victims, and more importantly, through deterrence, it helps avoid *all* harm to future victims who would be created if the tort system did nothing.

Tobacco litigation is paradigmatic of the problem, and has also been fertile ground for innovative solutions. The diseases caused by tobacco are, in the main, elusive of individual attribution: it is very difficult for an individual smoker who suffers from emphysema, or cancer, to prove with any certainty that he would not have contracted the disease "but for" smoking. All we can say for sure is that smoking increased the risk of the disease. But increased *risk* in an individual means increased *prevalence*

61. This is a problem for plaintiffs only, of course, if the background risk was higher than the probability of the harm resulting from the wrong. I suggest that this is the case in most toxic torts and many products liability claims, but I allow that in some cases the "balance of probabilities" in the individualistic system could result in *every* claimant succeeding even where we know that some of the harm was not defendant-caused. This does not, in my view, weaken the case for aggregate assessment of causation that I make in this article, and in fact the opposite: it is fairer to both defendants and plaintiffs because it neither over- or under-deters, and exacerbates neither the "sweetheart" or "blackmail" settlement problem.

in a population. We can know with scientific certainty that *some* smokers with cancer would not have got the disease "but for" smoking, and we may even be able to know, with some certainty, how many. To move to a further level of abstraction, if the smoker's claim is based on a failure to warn, the question of whether an adequate warning would have prevented the smoking is elusive in an individual case, while we can at the same time know that warnings do reduce the prevalence of smoking in populations. Thus, it should be possible to determine how much tobacco-related disease could have been prevented by an adequate warning, even if each individual's claim must, under the principles of tort law, fail. We can then place a dollar figure on the global loss.

In British Columbia a statute, the *Tobacco Damages and Health Care Costs Recovery Act*[62] ("BC *Tobacco Act*"), swept aside the particularistic rules of tort in favour of an "aggregate action" by the government to recover damages caused by tobacco-related wrongs, regardless of whether any particular smoker could prove a complete tort. Rules were introduced permitting liability, harm and damages to be assessed collectively. The government filed its suit, and the defendant manufacturers challenged the BC *Tobacco Act*, *inter alia*, on the basis that such rules were unfair. The government argued that aggregate assessment of damages in cases of causal indeterminacy was superior to individual adjudication, because harm in populations can be more, rather than less, accurately measured as a whole rather than as a sum of parts. In *British Columbia v Imperial Tobacco*,[63] Justice Major, writing for the unanimous Supreme Court of Canada, wrote:

> [t]he rules in the Act with which the appellants take issue are not as unfair or illogical as the appellants submit. They appear to reflect legitimate policy concerns of the British Columbia legislature regarding the systemic advantages tobacco manufacturers enjoy when claims for tobacco-related harm are litigated through individualistic common law tort actions.[64]

This echoes the view of the trial judge, who had found in 2000, when reviewing an earlier iteration of the BC *Tobacco Act* that:

62. SBC 2000, c 30.
63. 2005 SCC 49.
64. *Ibid* at para 49.

> [t]he basic tenet that causation within a population may be more accurately identified statistically than by means of attribution of individual causation in a multiplicity of conventional tort-based actions appears sound. The use of statistical and epidemiological evidence is an essential aspect of an aggregate action. The question in issue becomes causation in the group rather than of any individual member.[65]

Justice Holmes reiterated this endorsement when a revised version of the BC *Tobacco Act* came before him again in 2003:

> [t]here is nothing inherently unfair about an aggregate action. In fact, it may often balance unfairness of proceeding either by individual actions or by other forms of collective proceeding. Neither is the use of statistical or epidemiological evidence itself evidence of an unfair trial. They are aids to the resolution of issues in a unique but appropriate form of action.[66]

The breathtaking possibilities of the true aggregate approach have been further demonstrated in the recent decision of the Quebec Superior Court in a pair of tobacco-related class actions, *Blais v JTI-Macdonald Corp* ("*Blais*") and *Létourneau v JTI-Macdonald Corp*[67] ("*Létourneau*"). In those cases the plaintiffs had the benefit of some statutory provisions that eased their problems of proof and permitted the trial judge to assess liability, harm and damages in the aggregate.[68]

The claim in *Blais* was by a class of smokers who had contracted

65. *JTI-Macdonald Corp v British Columbia (Attorney General)*, 2000 BCSC 312 at paras 74-75.
66. *R v Imperial Tobacco Canada Ltd*, 2003 BCSC 877 at para 156.
67. 2015 QCCS 2382 (*sub nom Blais v JTI-Macdonald Corp* 2012 QCCS 469) [*Létourneau*].
68. Quebec's *Tobacco-Related Damages and Health Care Costs Recovery Act*, RSQ c R-2.2.0.0.1 ("[i]n an action brought on a collective basis, proof of causation between alleged facts, in particular between the defendant's wrong or failure and the health care costs whose recovery is being sought, or between exposure to a tobacco product and the disease suffered by, or the general deterioration of health of, the recipients of that health care, may be established on the sole basis of statistical information or information derived from epidemiological, sociological or any other relevant studies, including information derived from a sampling". This provision is made applicable to class actions by the last paragraph of section 25, which states that the rules in, *inter alia*, section 15 "also apply to any class action based on the recovery of damages for the (tobacco-related) injury": ss 15, 25).

cancer and emphysema, allegedly from smoking. In *Létourneau*, the class members were smokers who were claiming for the harm of addiction. In both classes, individual causation was, at least, uncertain: given the nature of disease processes and other causal issues such as causation-reliance, no one could conclusively attribute any particular disease to smoking, nor disease or addiction to the wrong of the defendant.

The fundamental finding of the court was that the three Canadian manufacturers of cigarettes were at fault because they did not warn of the dangers inherent in their product, adequately or at all. But with the breach demonstrated, what then? Did the plaintiffs still have to show that every individual for whom recovery was sought had suffered a harm causally linked to the defendant's wrong? Justice Riordon did not think so. He documented a 50-year campaign of deception and misinformation, and then went on to impose staggering damages ($15 billion, after interest) despite the fact that there had been no evidence that any particular class member had suffered harm.

The specific provisions of the Quebec legislation provided only that statistical and epidemiological evidence, including sampling, could be used to establish liability as well as damages. The defendants argued that this did nothing more than permit questions of individual harm to be decided with resort to such evidence. It did not, they suggested, extend to permitting proof of causation in populations and assessment of causation on an aggregate basis.

Justice Riordon rejected this interpretation. He wrote:

> [t]he objective of the TRDA is to make the task of a class action plaintiff easier, inter alia, when it comes to proving causation among the class members. When the legislator chose to favour the use of statistics and epidemiology, he was not acting in a vacuum but, rather, in full knowledge of the previous jurisprudence to the effect that each member of the class must suffer the same or similar prejudice. It thus appears that the specific objective of the act is to move tobacco litigation outside of that rule.[69]

This, the trial judge observed, effectively overrode the "previous jurisprudence calling for proof that each member suffered a similar

69. *Létourneau*, *supra* note 67 at para 692.

prejudice".[70] In Riordon J's analysis, this led to the conclusion that individual proof of causation was unnecessary altogether.

The *Blais* and *Létourneau* decision represents the first clear Canadian manifestation of what is, if class actions are to fulfill their promise as an effective compensatory and regulatory device,[71] inevitable: the adaptation of *substantive* law of causation in tort to accommodate the scale and difficulties associated with truly massive wrongs.

Justice Riordon in *Blais* and *Létourneau* seemed sufficiently pleased with the aggregate approach that he mused openly about whether the techniques should be available in all class actions. He said:

> [i]t will be interesting to see if the National Assembly eventually chooses to broaden the scope of this approach to have it apply in all class actions. Although such a move would inevitably be challenged constitutionally, its implementation would go a long way towards removing the tethers currently binding class actions in personal injury matters.[72]

Justice Riordon's decision, while singular in the Canadian jurisprudence, was not entirely without precedent. Judges faced with massive claims spanning large periods of time have before been driven to techniques of "wholesale justice" in order to "fit the forum to the fuss". I have described and discussed the resulting innovations, such as "market share"[73] or

---

70. *Ibid* at para 693.
71. There are three commonly-accepted objectives of class actions in Canada: compensation; behaviour modification (that is to say, the deterrence of wrongdoing); and "access to justice". The third objective may be seen as valuable principally to the extent that it facilitates the first two, although it can be argued also to be an independent social good.
72. *Létourneau*, *supra* note 67 at para 693, n 319.
73. Market share liability operates in individual cases as well as class actions, and holds defendants liable on the basis of their risk contribution, where a causal nexus between victim and wrongdoer can't be established: *Sindell v Abbott Laboratories*, 26 Cal (3d) 588 (Sup Ct 1980 (US)). In Canada, the theory has been permitted to proceed through certification in *Gariepy v Shell Oil Co* [2000] 52 OR (3d) 181 (Sup Ct J) at para 11, and referred to as a potential claim in *Ragoonanan Estate v Imperial Tobacco Canada Ltd* [2000] 51 OR (3d0 603 (Sup Ct J) at para 27.

"sampled" liability[74] and other rules facilitating proof of causation in populations, extensively elsewhere.[75]

Recent Ontario decisions seem also to hint at the willingness of courts to entertain questions of liability on an aggregate basis. In *Ramdath v George Brown College*[76] ("*Ramdath*"), the defendant college had negligently misrepresented that completion of its courses would lead to three professional designations, something that was not true. At the trial of the common issues, Justice Belobaba found that the defendants had breached their duty to the plaintiffs both under the *Consumer Protection Act, 2002*[77] and under negligence law, but noted that this alone would not entitle them to recovery for the tort without proof that each had relied on the misrepresentation. He wrote:

> *[f]urther evidence may still be needed to establish legal liability for negligent misrepresentation, namely, evidence of individual reliance. This question will no doubt be addressed in the next phase of this litigation*. However, legal liability has been established under the CPA because, as already noted, under this statute, evidence of actual reliance is not required … The common issues trial has now been concluded. *The next step in this class proceeding is to schedule a case conference to discuss the "damages" phase of this lawsuit*. Counsel should contact my office to arrange a convenient date for the case conference.[78]

---

74. In *Hilao v Estate of Marcos*, 103 F (3d) 767 (9th Cir 1996 (US)) the Court directed that the class of claimants against the late Philippine dictator for human rights abuses could be sampled to determine the number of valid claims, to permit a global assessment of damages. Similar approaches adopted by federal trial judges in the US, however, were subsequently disfavoured by appeal courts: see *e.g. Cimino v Raymark Industries, Inc*, 151 F (3d) 297 (5th Cir 1998 (US)) and *McLaughlin v Philip Morris USA Inc*, 522 F (3d) 215 (DC Cir 2009 (US))(these decisions, which found that abandoning individual proof of causation was a violation of the 5th and 7th Amendments to the US Constitution or the "predominance" requirements of US Federal Rule 23 (establishing the class action procedure), have little utility in the Canadian analysis as these provisions have no equivalent here).
75. See *e.g.* Cassels & Jones, "Rethinking Ends and Means in Mass Tort", *supra* note 54.
76. 2012 ONSC 6173 [*Ramdath* 2012].
77. SO 2002, c 30, Schedule A.
78. *Ramdath* 2012, *supra* note 76 at paras 94-95 [emphasis added].

On appeal, the College challenged the finding of the trial judge that the college owed a duty of care to its students, because the tort required "reasonable reliance". The British Columbia Court of Appeal appeared to treat the question, central to liability, as an aggregate, generic issue, rather than an individualistic one, holding:

> *[t]he appellant's concerns are primarily directed at whether each of the class members reasonably relied on the representations and can prove damages*. The issue of damages was not certified as a common issue and will be determined, with evidence, at the individual issues phase of the trial.[79]

So, without confronting the matter directly, the Court of Appeal appeared to believe that the question of individual reliance, an element of liability, could be dealt with simultaneously with the question of financial loss, thereby blending the concept of liability and damages.

This "blending" appeared complete when the question was returned to the trial judge for the damages phase of the hearing. By that time, counsel had agreed to proceed solely on the *Consumer Protection Act* remedy for "damages". The main controversy was whether the legislative remedy still required, like negligence, a causal link between the wrong and loss. The trial judge had concluded in the prior hearing that it did not, but the matter was reargued before him again:

> GBC, however, argues that even if reliance is not required to establish an unfair trade practice under the Act, or to rescind the consumer agreement and get a refund of monies paid, some measure of causation must still be shown if the consumer is claiming "damages". The entitlement to claim damages, says GBC, does not vitiate the need to prove causation. There has to be at the very least some evidence of a causal connection or nexus between the unfair practice and the damages being claimed. And this nexus can only be determined, argues GBC, on an individual, *i.e.* not aggregate, basis. This issue - whether or not the s. 18(2) damages remedy requires proof of a causal connection - dominated both the written and oral submissions. The issue has not been addressed in the case law and is not self-evident. Fortunately, I do not have to decide the matter. I am satisfied on the uncontroverted findings that have already been made in this litigation that a sufficient causal connection (for the purposes of the s. 18(2) damages remedy) has been established. I refer in particular to the following findings in the Common Issues and Appeal Decisions: (i) It was the opportunity to complete the three industry designations that attracted the plaintiffs to GBC in the first place, not the GBC certificate. None of

79. *Ramdath v George Brown College of Applied Arts and Technology*, 2013 ONCA 468 at para 8 [emphasis added].

> them wanted or needed another college graduate certificate. (ii) The plaintiffs claim they would not have enrolled in the Program but for the representation about the industry designations. For each of them, and for the students they represent, the value of the Program was the promised opportunity to complete the requirements for the CITP, the CCS and the CIFF "in addition to" the GBC graduate certificate. (iii) The promise of these industry designations made the program very attractive to prospective students. I also rely on the common sense observation that students applying for an eight-month college program (especially those that are coming from foreign lands) will most likely review the Program description before applying and paying a substantial tuition. In short, *I have no difficulty concluding that if the s. 18(2) damages remedy requires some nexus or causal connection with the unfair practice, this has been sufficiently established*.[80]

So let us be clear on what is happening here. The trial judge found that the requirement of individual reliance in all members of the class (if indeed it was required), had been satisfied, not through individual evidence, but rather on (i) the representative plaintiffs' own pleadings and evidence; and (ii) judge's class-wide inference, based on "common sense" and the circumstances of the misrepresentation, that *all* the class members had probably relied. He then moved quickly to an assessment of aggregate damages.

Ramdath was followed by Justice McEwen in *Trillium Motor World Ltd v General Motors of Canada Limited*[81] ("*Trillium*"), a case turning on "loss of chance". There, the court decided it could award class wide damages to car dealers who had lost an opportunity to negotiate due to the default of the defendant. But *would* each in fact have negotiated? Even though the evidence had been specific to the representative plaintiff alone, the court felt comfortable extrapolating this causation question to the entire class, and disposing of it simultaneously with the question of class-wide, aggregate damages:

> [t]he third precondition, s. 24(1)(c), is the most critical. In this regard, I must determine the aggregate or a part of the defendant's liability to the Class Members and give judgment accordingly where the aggregate or part of the defendant's liability to some or all of the Class Members can reasonably be determined without proof by individual Class Members. Justice Belobaba

80. *Ramdath v George Brown College*, 2014 ONSC 3066 at paras 17-19 [*Ramdath* 2014] [emphasis added, italics in original].

81. 2015 ONSC 3824.

> recently analyzed this precondition in *Ramdath v George Brown College*, 2014 ONSC 3066 (CanLII), 375 D.L.R (4th) 488 at para 47, identifying three requirements: (a) the reliability of the non-individualized evidence that is being presented; (b) whether the use of this evidence will result in any unfairness or injustice to the defendant (for example, by overstating the defendant's liability); and, (c) whether the denial of an aggregate approach will result in "a wrong eluding an effective remedy" and thus a denial of access to justice. In my view, Trillium has satisfied all three requirements. The basis for Trillium's claim in aggregate damages is loss of chance. This chance relates to the affected dealers as a group, and the likelihood that negotiations of the terms of the WDA would have taken place between the group as a whole and GMCL. The non-individualized evidence is reliable, the use of the evidence does not result in any unfairness to Cassels, and to deny the Class Members the aggregate approach would amount to the denial of a remedy. Acting collectively in negotiations with GMCL is a critical component of the Class Members' claim against Cassels. An individualized approach to damages would not only be unfair to the individuals who would have banded together, it would be misguided given the nature of their action. Determining how much more money would have been available from GMCL for the Class Members had they had an opportunity to negotiate for it does not cause any injustice to the defendant Cassels by overstating its liability; rather, it simply quantifies that liability.[82]

In other words, where it can be inferred that the wrong has had generic consequences across the class, and where there is no injustice to either plaintiff or defendant, the requirements of section 24(1) of the *Class Proceedings Act* of Ontario have been met, and group wide *liability* in the sense of causation can be assessed simultaneously with "monetary liability".[83] Section 24(1) of the Ontario Act, therefore, appears to be operating not as simply a procedural device, but at least as a framework

82. *Ibid* at paras 540-41 [emphasis in original].

83. Ont *Class Proceedings Act*, *supra* note 50 provides that aggregate damages may be awarded where: "(a) monetary relief is claimed on behalf of some or all class members; (b) no questions of fact or law other than those relating to the assessment of monetary relief remain to be determined in order to establish the amount of the defendant's monetary liability; and (c) the aggregate or a part of the defendant's liability to some or all class members can reasonably be determined without proof by individual class members": s 24(1).

within which courts may modify or extend the substantive law of tort.[84]

It is a very short step from these cases finding, without individual evidence, causation in the entire class, to a court finding, on a similar basis, that causation was made out in a *portion* of the population. Once the decision is made to consider the question on an aggregate basis, then argument can be introduced on the extent of the harm throughout the class. So in *Trillium*, for instance, if the defendant had produced evidence that *some* class members would not have negotiated and therefore suffered no loss, that need not send the matter for individual adjudication of each class members claim. The judge could still fairly assess aggregate liability if the harm could be equitably assessed proportionately through expert evidence, sampling or other devices (assuming also distribution concerns could be assessed in the aggregate). This would still satisfy the three-part test in *Ramdath*, and further the access to justice goals identified in that case. Justice Belobaba in *Ramdath* had introduced his damages judgment with the following paragraph:

> [a]ggregate damages are essential to the continuing viability of the class action. If all or part of the defendant's monetary liability to class members can be fairly and reasonably determined without proof by individual class members, then class action judges should do so routinely and without hesitation. Aggregate damage awards should be more the norm, than the exception. Otherwise, the potential of the class action for enhancing access to justice will not be realized.[85]

The same points can, and should, be made with respect to liability. If class-wide causation can be fairly and reasonably determined without individual proof, then class action judges should do so routinely and without hesitation. And we appear to be cautiously starting down that road. This is how it should be: recognizing that class proceedings Acts did not, in themselves, modify the substantive law does not mean that they froze the substantive law at the time of their enactment. The law of causation can adjust to the procedural context of a class action, and should do so when it so obviously improves the efficacy of the tort system,

84. That these cases were decided recently is significant, as they appear to be, at least to some extent, at odds with Rothstein J's comments in *Microsoft SCC*, *supra* note 3 at para 131 to the effect that class-wide liability must be established *before* the aggregate damages provisions could be applied.

85. *Ramdath* 2014, *supra* note 80 at para 1.

improves access to justice, secures compensation and effects appropriate deterrence.

Finally, I cannot resist pointing out that aggregate assessment of liability *already* underpins the judicial resolution of mass wrongs in Canada. Most certified cases settle before trial, and the settlement terms are subject to judicial oversight and approval. These settlements routinely rely on estimates of the global liability of defendants to "groups of persons" without proof of individual loss.[86] If one is "substantively, procedurally, institutionally, or circumstantially fair", as class action settlements must be,[87] then how can the other (the use of the same devices to determine a fair outcome at trial) not be?

Aggregate assessment of causation is a solution to the same problem that has led to reliance on waiver of tort. But because the aggregate award is actually premised on the true harm caused by the defendant's wrong (rather than just by the extent to which it has profited from it), the assessment of liability and quantification of the damages on an aggregate basis is far preferable whether viewing the question from a perspective of adequate compensation or optimal deterrence. This was the central insight of Riordon J in *Blais* and *Létourneau*, and the logic underpinning *Ramdath* and *Trillium*.

V. Conclusion

I have endeavoured to demonstrate, throughout this article, that in most class actions with problems of indeterminate causation, tort law is a better avenue of address than equitable waiver of tort.

Courts are beginning to recognize the efficiency of adapting the

86. See for instance the decision approving the settlement in *Pro-Sys Consultants Ltd v Infineon Technologies AG*, 2014 BCSC 1936 (where Masuhara J approved a settlement which incorporated principles of aggregate causation and market share liability). I would argue that many, if not most, class actions that settle before trial are "rough and ready" estimates of class-wide liability that are premised on disconnecting the causal link between wrongdoer and the individual victim.

87. See John C Kleefeld, "Facets of Fairness: *Kidd v Canada Life Assurance Company* and the Approval of Class Action Settlements" (2015) 10:1&2 Canadian Class Action Review 33 at 64.

substantive rules of tort to further the objectives of compensation and deterrence, so that class action procedures can fulfill their promise as one of the principal avenues for access to justice in large-scale claims. The courts' willingness to entertain waiver of tort in collective litigation is a step on the way, and indicates that judges are alive to the problems of indeterminate causation in class claims.

But the recognition of waiver of tort would represent only a partial solution, one that might yield poor results and risk backfiring on the victims of tort, undercompensating them and, through under-deterrence, ensuring that more victims will be created in the future.

Conversely, a narrow, particularistic application of tort causation rules is little better. It is possible for courts to use presently recognized devices, such as "robust inferences of causation" or reversed onuses, to overcome individual attribution issues even in class claims. On one reading, this is what the Ontario courts were doing in *Ramdath* and *Trillium*, and it also could be viewed as operating in *Blais* and *Létourneau.* But these devices, which are artificial in individual claims, are even more plainly so in a class claim, where defendants are *inferred* or *presumed* to have caused a magnitude of harm that we know they did not cause. This could be heavy-handed, even absurd, in many toxic torts and many products liability cases: where a defendant's wrong has caused a measurable, but still incremental, increase in the incidence of a particular disease or injury, the "robust inference" or reversed onus might mean it would be on the hook for *all* such injuries.

The "better way" is to continue down the road blazed by the Quebec and Ontario courts, and embraced by legislatures in provincial tobacco legislation: viewing classable claims as cases of harm in populations, rather than in individuals. True aggregate estimation of the harm, with the plaintiff retaining the burden of proof, can be both more accurate and fair, and, where it is, it should be employed.

Undermining Justice: The Two Rises of Freedom of Contract and the Fall of Equity*

Hila Keren**

This article explores a crucial moment in American legal history, known as the Lochner era, in which the rise of freedom of contract was sharp enough to defeat equity concerns, and then argues that a second rise of the freedom of contract has recently been developed by the Supreme Court in the domain of arbitration agreements. It contends that this second rise is not only a revival of Lochnerism but also, and more so, what the article names "neoliberal-Lochnerism": a process of legal dissemination of neoliberal common sense outside of the world of contracts. Via close reading of leading recent cases, the article demonstrates that the genus of arbitration agreements now allowed by the US Supreme Court represents an assault on fairness, morality, and justice that is larger than the eye can see at first glance. The result, it is argued, is "law without equity", a form of neoliberal jurisprudence that allows, and even incentivizes, humans who have accumulated enough power to act opportunistically. Without equity's restraining power, the article concludes, those possessing a combination of economic means, political influence, and intellectual sophistication can and will exploit the legal rules to undermine justice.

---

* The title of this article responds to Patrick Atiyah, *The Rise and Fall of Freedom of Contract* (Oxford & New York: Clarendon Press & Oxford University Press, 1979). For an earlier response to the same book see FH Buckley, ed, *The Fall and Rise of Freedom of Contract* (Durham: Duke University Press, 1999).

** Professor of Law, Southwestern Law School. Part of this article was selected for presentation at a working conference organized by The Vulnerability and the Human Condition Initiative at Emory Law and I thank Martha Fineman for the opportunity. I deeply thank Danielle Hart who took the time to read this piece and share valuable insights. For their excellent research assistance, I thank Alice Zakaryan and Aykanush Tabadzhyan. Finally, I am grateful for the dedicated work of the editors of the *Canadian Journal of Comparative and Contemporary Law.*

I. Introduction

That a true freedom of contract necessitates legal enforcement of agreements by courts is an accepted premise from time immemorial. Throughout the years, and until today,[1] numerous American cases have quoted a famous English case from 1875 that stated the idea as follows:

> if there is one thing which more than another public policy requires it is that men … shall have the utmost liberty of contracting, and that their contracts when entered into freely and voluntarily shall be held sacred and shall be enforced by Courts of justice.[2]

1. See *e.g.* the 1900 case of *Baltimore & OSR Co v Voigt*, 176 US 498 (1900) at 505-506; and the 2015 case of *Royston, Rayzor, Vickery, & Williams, LLP v Lopez*, 467 SW (3d) 494 (Tex Sup Ct 2015).

2. *Printing & Numerical Registering Co v Sampson* (1875), 19 LR Eq 462 at 465 (CA (Eng)).

Although few would argue against the general presumption that the freedom of contract requires an enforceability of contracts, there has also been a much larger ongoing debate regarding the *limits* of this idea. Do courts always have an obligation, or a duty, to enforce contracts, or are they allowed to refuse enforcement under some circumstances? With regard to this question numerous American cases have quoted another old and famous English case, authored in 1751 in the "courts of conscience" — England's courts of equity.[3] In this case, Lord Chancellor Hardwicke explained that the courts of conscience would not enforce agreements that "no man in his senses and not under delusion would make on the one hand, and … no honest and fair man would accept on the other".[4]

Moreover, he also described those undeserving agreements by directly referring to equity and conscience, naming them "*unequitable* and *unconscientious* bargains".[5] And, although it was not the first time that courts had refused enforcement of unfair contracts,[6] it was certainly one of the first times the refusal was theorized in conscience-oriented terms and reflected the logic of equity. Lord Chancellor Hardwicke's words established the unconscionability principle as an equity-based limit on the freedom of contract and his words have proven appealing to generations of judges and legal commentators on both sides of the

3. See *e.g.* Dennis R Klinck, "The Nebulous Equitable Duty of Conscience" (2005) 31:1 Queen's Law Journal 206 at 208 [Klinck, "Nebulous"] citing *Ewing v Orr* (1883), 9 App Cas 34 (HL)(in which the court said "[t]he courts of equity in England are, and always have been, courts of conscience" at 40); see also Klinck, "Nebulous" (stating that "no doubt historically conscience and equity were intimately allied, even synonymous" at 211).
4. *Earl of Chesterfield v Janssen* (1751), 28 ER 82 at 100 (Ch) [*Earl of Chesterfield*].
5. *Ibid* [emphasis added].
6. For the "ancient roots" of unconscionability, see *e.g.* Stephen E Friedman, "Giving Unconscionability More Muscle: Attorney's Fees as a Remedy for Contractual Overreaching" (2010) 44:2 Georgia Law Review 317 at 334-43 (citing sources which connect the idea to ancient Jewish and Roman law).

pond.[7] Notably, Hardwicke LC's words have also proven influential in all courts, regardless of the traditional separation between courts of equity and courts of law.[8]

The conflict between the freedom of contract and the unconscionability principle — with the former construed as demanding courts to enforce contracts and the latter understood as ordering courts to refuse enforcement — has yielded an ongoing and intense jurisprudential debate, depicted by a number of scholars.[9] However, for the most part American courts have managed to strike some sort of balance by routinely enforcing contracts while occasionally utilizing unconscionability and other equity-based principles to deny enforcement. And, although the pendulum has shifted from time to time,[10] by and large there has been no definite loser or winner.

An exception emerged, however, during a defined period in American jurisprudence known today as the *Lochner* era.[11] During this

---

7. The first American case to refer to the words of Hardwicke LC in *Earl of Chesterfield*, *supra* note 4 is *Powell v Spaulding*, 3 Greene 443 (Iowa Sup Ct 1852) [*Powell*]. The Supreme Court has adopted the full definition in *Hume v United States*, 132 US 406 (1889) [*Hume*]. To date, the latest case citing the definition in full (including the archaic term "unconscientious bargains" as opposed to only referring to *Earl of Chesterfield*, *supra* note 4 is *Brown v Genesis Healthcare Corp*, 228 W Va 646 (Sup Ct App 2011 (US)) at 67-80 [*Brown*], vacated, *Marmet Health Care Center, Inc v Brown*, 132 S Ct 1201 (2012 (US)) [*Marmet*].
8. See the decision of the US Supreme court in *Hume*, *supra* note 7 (citing Hardwicke LC's definition, and other cases that had followed it, and affirming the lower court's finding that "[t]hese citations are sufficient to show that in suits upon unconscionable agreements the courts of law will take the matter in their own control, and will, without the intervention of courts of equity, protect the parties against their enforcement" at 406, 411).
9. See *e.g.* Carolyn Edwards, "Freedom of Contract and Fundamental Fairness for Individual Parties: The Tug of War Continues" (2009) 77:3 University of Missouri Kansas City Law Review 647.
10. Charles L Knapp, "Rescuing Reliance: The Perils of Promissory Estoppel" (1998) 49:5&6 Hastings Law Journal 1191 at 1202.
11. The era got its name from the infamous case of *Lochner v New York*, 198 US 45 (1905) [*Lochner*].

era, the Supreme Court created a new jurisprudence, giving the freedom of contract a *constitutional* status strong enough to command the enforcement of contracts even in the face of state legislation specifically designed to invalidate them. This article explores this crucial moment in history, in which the rise of freedom of contract was sharp enough to defeat equity concerns, and then argues that a second rise of the freedom of contract has recently been developed by the Supreme Court in the domain of arbitration agreements.

Much like in the *Lochner* era, this rise involves an emergence of a new jurisprudence, and in this case, one that entails an original form of interpretation of the *Federal Arbitration Act*[12] ("*FAA*"). First in *AT&T Mobility LLC v Concepcion*[13] ("*AT&T*"), and then in *American Express v Italian Colors Restaurant*[14] ("*American Express*"), the Court decided to reverse decisions of the lower courts that refused to enforce class arbitration waivers. In both situations, the lower courts had refused to enforce the waivers because enforcing them would deny the waiving parties — consumers in *AT&T* and a small restaurant in *American Express* — any access to justice. However, as the Supreme Court explained, each of the lower courts — California's Supreme Court in *AT&T* and the US Court of Appeals for the Second Circuit in *American Express* — erred because, according to the new reading of the *FAA*, the freedom of contract is not limited by fairness or morality concerns in the arbitration context. Rather, "courts *must* rigorously enforce arbitration agreements",[15] even if the drafting parties deliberately have used their superior bargaining power not to design a private forum of litigation, but as a way *to avoid* litigation, thereby circumventing their legal liability and leaving future claimers with no legal recourse.

What makes this approach comparable to the one developed in the *Lochner* era is that the Court, once again, has taken the freedom of contract to a new level, making it a concept that has the power to de-authorize and delegitimize state actors as they make efforts to prevent

12. 9 USC §2 (1947) [*FAA*].
13. 131 S Ct 1740 (2011 (US)) [*AT&T*].
14. 133 S Ct 2304 (2013 (US)) [*American Express*].
15. *Ibid* at 2309 [emphasis added].

injustice. Surely, *AT&T* and *American Express* were not the first in which the Supreme Court reversed decisions that invalidated arbitration agreements. However, until recently such reversals were explained by the Court's belief that the terms were not as inequitable as the lower court had seen them. And here lies the "revolution":[16] under the new jurisprudence, arbitration agreements should be enforced even if they are inequitable. Put another way, never before were courts categorically forbidden from using equity principles to overcome extreme injustice caused by an opportunistic and manipulative use of the law. And, the change is especially remarkable given the fact that the ability to utilize equity tools in the context of arbitration is explicitly guaranteed under the *FAA* itself; in its relevant section the *FAA* clarifies that an arbitration agreement "shall be valid, irrevocable, and enforceable, save upon such grounds as exist at law or *in equity* for the revocation of any contract".[17]

While the parallels between the *Lochner* era and the new arbitration jurisprudence are prominent,[18] the current "neo-*Lochnerism*"[19] is more than a revival of an old belief for a few reasons. First, because applying *Lochnerian* ideas to our times exponentially magnifies their effect mainly due to the incalculable amount and variety of mass contracts that impose arbitration across the entire market. Secondly, despite the similarities, there is something thoroughly different about the new rise of the freedom of contract. This time around, I argue, we face a "neoliberal-*Lochnerism*", and this neoliberal version means, to quote political scientist Wendy

16. J Maria Glover, "Disappearing Claims and the Erosion of Substantive Law" (2015) 124:8 Yale Law Journal 3052 [Glover, "Disappearing Claims"] (describing the Supreme Court's new arbitration as a revolution).
17. *FAA*, *supra* note 12 [emphasis added].
18. I am not the only one to point to this similarity. See *e.g.* Burt Neuborne, "Ending Lochner Lite" (2015) 50:1 Harvard Civil Rights-Civil Liberties Law Review 183. However, my focus is on the difference between the original *Lochnerism* and its contemporary version.
19. For the use of this term in a different context (relating to the Supreme Court's new First Amendment jurisprudence that generally speaking assigned corporations the rights formerly reserved to humans), see Jedediah Purdy, "Neoliberal Constitutionalism: Lochnerism for a New Economy" (2014) 77:4 Law & Contemporary Problems 195.

Brown, that "more is at stake … than support for capital in the name of freedom".[20] This article suggests that what we are witnessing is no less than a very troubling phase in the neoliberal project — one that uses the power of law to thwart equity principles, denies the logic of equity, and will eventually eradicate the notion of justice.

Beyond the actual outcome of particular cases, neoliberal-*Lochnerism* is produced by the emergence of an original legal discourse that operates to transform the meaning of everything. Offering a close reading of the rhetoric used by Justice Scalia in both *AT&T* and *American Express*, this article demonstrates how the new arbitration jurisprudence works to disseminate neoliberal rationality, in at least three major ways. First, Scalia J's analysis translates every idea to "*Economish*" — the language of economy — and reframes the issue to fit "the logic of profit-making".[21] For example, unlike public litigation, private arbitrations are presented as offering "greater efficiency",[22] by "reducing the cost and increasing the speed of dispute resolution".[23] Second, the logic of speed and efficiency is used to conduct an assault on a collective agency of legal subjects. For instance, Scalia J asserts that allowing class arbitrations is "likely to generate procedural morass",[24] and therefore only individual arbitration is rational. Such a "divide and conquer" tactic not only isolates weaker subjects, it also assigns to them the sole responsibility for their poor fate. And third, this neoliberal reasoning is capable of gaining broad popularity outside the elite because it is presented as universal, while the fact that it is biased and reflects only the viewpoints of the most powerful market actors is carefully disguised. All in all, the new rationality takes over not only the economy but also the justice system, which it leaves devoid of

20. Wendy Brown, *Undoing the Demos: Neoliberalism's Stealth Revolution* (New York & Cambridge, Mass: Zone Books & MIT Press, 2015) at 153.
21. Eric J Weiner, *Private Learning, Public Needs: The Neoliberal Assault on Democratic Education* (New York: Peter Lang, 2005) at 20 citing Henry A Giroux, "Neoliberalism, Corporate Culture, and the Promise of Higher Education: The University as a Democratic Public Sphere" (2002) 72:4 Harvard Educational Review 425.
22. *AT&T*, *supra* note 13 at 1751.
23. *Ibid* at 1749.
24. *Ibid* at 1751.

equity logic.

Before proceeding, an explanation is warranted for my use of the term equity and my recurrent references to equity principles, equity discourse, and the logic of equity. For the purposes of this article, "equity" is deliberately used in a non-technical and non-formalistic fashion. To be sure, as an old concept that survived centuries of use by different humans in a variety of countries and cultures, equity cannot possibly have a simple meaning. And yet, I believe that a *modern working definition* is essential in order to capture the problem at hand. I also believe that the old idea of equity would not have endured unless the specific doctrines developed under this broad title shared a core logic and together brought to modern law a unique focus that goes beyond the particulars.

To conceptualize such a general message, I start from adopting Dennis Klinck's view that "[o]ne cannot delve very far into judicial equity without encountering the notion of 'conscience'";[25] combined with Irit Samet's argument that conscience offers equity "a workable legal standard".[26] I continue with an observation made by the former Chief Justice of Australia, Sir Anthony Mason, who wrote that "equity came to reflect a strong sense of morality", and "equitable principles were shaped with a view to inhibiting unconscientious conduct and providing for relief against it".[27] Additionally, I draw on two recent works done at the intersection of equity and private law by an American legal scholar, Henry Smith,[28] and a Canadian philosophy scholar, Dennis Klimchuk.[29] Both have attributed to equity the role of inhibiting opportunists from

25. Dennis R Klinck, *Conscience, Equity, and the Court of Chancery in Early Modern England* (Burlington, Vt: Ashgate, 2010) at vii [Klinck, *Conscience*].
26. Irit Samet, "What Conscience Can do for Equity" (2012) 3:1 Jurisprudence 13 at 20.
27. Sir Anthony Mason, "Equity's Role in the Twentieth Century" (1997-98) 8 Kings College Law Journal 1 at 1.
28. Henry E Smith, "Property, Equity, and the Rule of Law" in Lisa M Austin & Dennis Klimchuk, eds, *Private Law and the Rule of Law* (Oxford: Oxford University Press, 2014) at 224 [Austin & Klimchuk, *Private Law*].
29. Dennis Klimchuk, "Equity and the Rule of Law" in Austin & Klimchuk, *Private Law*, *supra* note 28 at 247 [Klimchuk, "Equity"].

exploiting the generality, formality, and strictness of the law. Thus, all in all, I use the idea of equity as an insistence that judicial discretion should be applied with conscience in mind, and that the legal outcome must deter exploitation of the law while promoting fairness, moral behavior, and social justice.[30]

This article unfolds in three parts. Part II tells the story of the first rise of the freedom of contract during the *Lochner* era. Part III argues and demonstrates that a second rise of the freedom of contract is taking place in recent years. It also contends that this second rise is not only a revival of *Lochnerism* but further reflects a process of legal dissemination of neoliberal rationality. Part IV explains why and how such neoliberal rationality works to defeat principles of equity, and cautions that "law without equity" severely undermines the quest for justice. The article concludes in a somewhat more hopeful tone, emphasizing the power of equity to counter neoliberal rationality and to offer a better narrative of justice.

II. The First Rise of Freedom of Contract

The jurisprudence developed by American courts during the "*Lochner* era" had made the limitation of state powers the essence of the freedom of contract. Within a few decades, between the end of the 19th century and the first quarter of the 20th century,[31] the freedoms held by market actors were not only articulated as a chief liberty, but more importantly, were instilled with new meaning. In 1909, only four years after the *Lochner* decision, Roscoe Pound stated that "liberty of contract" was a new term in an article he authored, titled "*The Liberty of Contract*".[32] In it, Pound

30. For a longer discussion of this role of the judiciary as applied to market behaviors see Hila Keren, "Guilt-Free Markets? Unconscionability, Conscience, and Emotions" (Forthcoming, 2016) Brigham Young University Law Review [Keren, "Guilt-Free Markets"].

31. See *e.g.* Morgan Cloud, "The Fourth Amendment During the Lochner Era: Privacy, Property, and Liberty in Constitutional Theory" (1996) 48:3 Stanford Law Review 555 (presenting different periods covered by the term).

32. Roscoe Pound, "The Liberty of Contract" (1909) 18:7 Yale Law Journal 454.

embarked on a description of the rise of freedom of contract and the meaning attached to the term. The result is an authentic portrayal of the inception of an idea that controls our life until today, without the problems attached to hindsight wisdom.

As Pound explained, the term "liberty of contract" itself was new then and was not used by courts as a legal constraint on governments' powers prior to 1886. The novelty of the American courts of the period was less in recognizing the freedom of choice available to individuals dealing with one another within the market, and more in using this freedom, or liberty, as a *constitutional principle*. Case after case, the newly constitutionalized "liberty of contract" was used to invalidate state efforts to protect weaker market players, mainly workers, via regulation of the market. With a critical tone, Pound described how, by elevating the freedom of contract to the level of a constitutional principle, the courts have deemed unconstitutional regulations that limited labor hours of women[33] and workers of bakeries,[34] required employers to pay wages in money rather than with credit to the employer's store,[35] determined how coal should be weighed for purposes of compensating miners,[36] and so on. As Pound generalized: "[i]n this way [freedom of contract] became a chief article in the creed of those who sought to minimize the functions of the state".[37]

But what about certain individuals who may be pressured into agreeing to harmful contracts? Shouldn't the state protect them? The answer emerging in the *Lochner* era was bluntly negative. Attempts to argue that some individuals or groups of individuals — such as industrial workers or women — are more vulnerable and need special protection by the state failed. The reasoning, which Pound himself criticized at length as "academic"[38] and hence false, is worth our attention. Freedom

33. *Ibid* at 475, citing *Ritchie v People*, 155 Ill 98 (Sup Ct 1895).
34. *Ibid* at 479, citing *Lochner*, *supra* note 11.
35. *Ibid* at 472, citing *State v Goodwill*, 33 W Va 179 (Sup Ct App 1889); and citing *Frorer v People*, 141 Ill 171 (Sup Ct 1892) at 473 [*Frorer*].
36. *Ibid* at 471, citing *Jones v People*, 110 Ill 590 (Sup Ct 1884).
37. *Ibid* at 456.
38. *Ibid* at 487.

of contract, courts of the period insisted, is a natural liberty that, as such, applies equally to all. The employer and the employee, for example, were regarded as having equal rights in designing their contract, which included a determination of the amount of working hours. And, since "the employer and the employee have equality of right", as one court famously stated, "any legislation that disturbs that equality is an arbitrary interference with the liberty of contract, which no government can legally justify in a free land".[39]

Courts further insisted that to decide otherwise — that is, to affirm protective legislation that invalidates certain contractual provisions — would mean not to protect, but rather, to mistreat and harm the seemingly protected parties. This is because according to this *Lochnerian* approach, only inferior people lack the full capacity — or liberty, or freedom — to contract. Thus, for example, preventing an employee from selling his ability to work for long hours, would mean to degrade and insult him, or to treat him like a fool.[40] Similarly, invalidating a statute requiring wages to be paid solely in money the Supreme Court of Kansas reasoned, "places the laborer under guardianship, classifying him in respect of freedom of contract with the idiot, the lunatic, or the felon in the penitentiary".[41] What even Pound could see in the midst of the *Lochner* era and without the perspective of time was how hypocritical and unrealistic the hypothesis of equality between stronger and weaker market players was. The problem, of course, was that the equality logic and the presumption that weaker parties enjoy the same freedom of contract as their counterparts was used *against* these vulnerable parties to deprive them of state assistance or state protection.

This harmful effect is not surprising given the fact that insistence by the courts on the idea that miners, bakers, and women equally enjoy the same freedom of contract as their powerful employers never originated from a concern for the dignity of the less powerful individuals. Rather, the true motivation was an antiregulatory approach, aimed at delegitimizing any form of state intervention in the market, with special hostility to acts

39. *Ibid* at 454, citing *Adair v United States*, 208 US 161 (1908) at 175.
40. *Ibid* at 463.
41. *Ibid* at 477, citing *State v Haun*, 61 Kan 146 (Sup Ct 1899) [*Haun*].

of legislation. The concept of a universal and natural liberty of contract and the persistence that such a liberty exists for all were merely tools used to that end. Or, as Pound remarked upon the new jurisprudence of his time: "the idea of liberty of contract has been invoked to defeat legislation".[42]

However, the very use of the contractual idea in order to defeat the state has transformed the meaning of the idea itself. What once mainly stood for a celebration of individualism and private ordering has now begun to chiefly symbolize a paralyzed state lacking the power to take care of its own citizens or to promote public interests. As Pound himself warned back in 1909, the rhetoric of equal freedoms is not only "artificial",[43] but also quite dangerous: it has the potential to "defeat the very end of liberty".[44] When courts insist that the state cannot and should not interfere if vulnerable parties are exploited by contracts the result for those parties is *less* freedom.

Looking at things from a contemporary perspective, it is important to realize that the liberty of contract jurisprudence that had developed in Pound's days significantly deviated from the law of contracts of the period. To be sure, *the common law* often called for the enforcement of promises regardless of gaps in bargaining power between the parties and without much concern for fairness. However, principles developed in *equity* were part of the law as well, offering legitimate legal ways to protect vulnerable parties from unfair contracts. As Pound reminds us: "[f]rom the time that promises not under seal have been enforced at all, equity has interfered with contracts in the interests of weak, necessitous, or unfortunate promisors".[45] For example, equity courts have for a long time released sailors from promises to dispose of their wages "where they appeared unfair, one-sided, or inequitable"; and similarly ignored waivers of necessitous borrowers of their right for redemption.[46] And, despite the ongoing use of such equity-based protections, and prior to the emergence

42. *Ibid* at 470.
43. *Ibid* at 487.
44. *Ibid* at 484.
45. *Ibid* at 482.
46. *Ibid* at 482-83.

of the *Lochnerian* logic, there was no assertion that their use degraded the protected parties or classified them with "the idiots, lunatics, or felons".[47] Rather, these protections symbolized a concern for justice, arising from a realistic understanding that not all humans enjoy the same kind of freedom of contract.

Apparently, judges committed to the newly developed jurisprudence of liberty of contract were fully aware of this state of the law, which via principles of equity took into account gaps in the availability of the freedom of contract for individuals in differing socio-economic classes. In *Frorer v People*[48] ("*Frorer*") for example, the court reviewed legislation requiring that employees be paid monetary wages rather than with credit to their employer's store. In this context the judge explicitly discussed gaps of power between borrowers and lenders, explaining that "the borrower's necessities deprive him of freedom in contracting and place him at the mercy of the lender".[49] The *Frorer* court further explained that such inequality is the reason that "all civilized nations of the world, both ancient and modern" have some version of usury laws.[50] However, as Pound noted, it did not occur to the judge that the same logic was relevant to the miners and workers protected by the legislation under his review, that they too suffered from unequal freedom of contract, and that their inequality called for affirmation rather than invalidation of the reviewed protective legislation. Instead, and in the spirit of the new liberty of contract jurisprudence, the *Frorer* court decided that since all are equal in their freedom to contract no legislation can legitimately limit the ability of workers to "consent" to non-monetary wages.

The insistence of courts during the *Lochner* era that the freedom of contract is equally available to all members of society soon attracted criticism. Pound himself argued that such equality exists only in theory and blamed the courts of his time for their "academic" view of

47. *Ibid* at 477 (referring to the reasoning of *Haun*, *supra* note 41).
48. *Ibid* at 473, citing *Frorer*, *supra* note 35.
49. *Ibid.*
50. *Ibid.*

individualism that, he argued, has no factual basis.[51] Part of the damage of such an unrealistic approach, cautioned Pound, was a growing feeling of bias. Quoting "an acute and well-informed observer", Pound reported "a growing distrust of the integrity of the courts", coming from a "belief that the present judge has been a corporation attorney, that his sympathies and experience and his whole view of life is on the corporation side".[52] The contemporary version of this salient bias argument will be further developed later.[53] For now, however, it is worth to note that the contention that letting freedom of contract take over equity principles risks the integrity of the law was expressed as early as in the midst of the *Lochner* era.

Much as Pound predicted in a hopeful tone, the *Lochner* era eventually came to an end. Specifically, the Supreme Court reversed its *Lochner* decision in 1937,[54] and more generally, protective legislation was no longer systematically deemed unconstitutional. However, as professor Robin West recently argued, the *concept* of freedom of contract as it was formed in the *Lochner* era — as "a natural right to determine whether, with whom, and on what terms we will take on other-regarding obligations, that in turn determines a sphere of freedom into which the state may not intrude" — remained alive.[55] As we shall now see our contemporary understanding of the idea of freedom of contract, to quote West, "is still strikingly *Lochnerian* in its content — we just do not think the Constitution protects it any longer".[56] However, in the coming section I go beyond West's argument to caution that our current version of *Lochner* has further transformed the original concept of freedom of contract, elevating its status and deepening its reach in a manner that

51. For a similar contemporary argument see David Strauss, "Why Was Lochner Wrong?" (2003) 70:1 University of Chicago Law Review 373 at 383-86.
52. Pound, *supra* note 32 at 487.
53. See Part IV, below.
54. Robin L West, "The Right to Contract as a Civil Right" (2014) 26:4 Saint Thomas Law Review 551 at 558 [West, "The Right to Contract"], citing *West Coast Hotel Co v Parrish*, 300 US 379 (1937) at 391.
55. West, "The Right to Contract", *ibid* at 558.
56. *Ibid.*

presents a new level of threat to equity principles.

III. The Second Rise of Freedom of Contract

A. The New Arbitration Jurisprudence

For several decades it seemed as if the freedom of contract was on the fall. For a while, and with a dip in the period known in American history as the New Deal, legislators and judges had shown an increased willingness to limit the freedom of contract in order to promote public goals and social justice. For example, new antidiscrimination laws limited the freedom in selecting contractual partners, emphasizing that certain categories, such as gender or race, could no longer be a legitimate basis for rejecting a potential partner. Notably, courts had used their equity powers, and especially the unconscionability principle, to invalidate unfair contracts and to release parties with inferior bargaining power from predatory obligations.[57]

During this post-*Lochner* era it seemed as if the old rivalry between the state and the market had been settled; the state was no longer required to leave the market alone ("*laissez-faire*"); and some political supervision of the economy was appropriate. While the freedom of contract still functioned as a symbol of autonomy, agency, and choice, and as imposing on courts an almost absolute duty to enforce contracts, it seemed to have lost its *Lochnerian* face that had previously debilitated the state and its legal powers and as imposing on courts an almost-absolute duty to enforce contracts. The Supreme Court's special liberty of contract jurisprudence — as described by Pound — was put to rest, and instead the Court's constitutional attention, especially under the leadership of Chief Justice Earl Warren, turned away from the protection of the market and towards matters more associated with modern democracy and its operation.

However, at some point during the 1970s, things began to change,

57. See *e.g. Williams v Walker-Thomas Furniture Co*, 350 F (2d) 445 (DC Cir 1965).

creating what many refer to today as our "neoliberal" age.[58] In addition to deregulation, privatization, and a general revival of the conflict between market and state, the neoliberal decades have brought back the *Lochnerian* understanding of freedom of contract. In this second rise of the freedom of contract, the idea has re-gained its meaning as a legal and political restraint on the state and its legal powers. And again, similar to the *Lochner* era, the renewed high status of the freedom of contract has resulted from the emergence of a new jurisprudence that has been developed by the Supreme Court: in this case, a new arbitration jurisprudence.

The Supreme Court's new arbitration jurisprudence, recently referred to as the arbitration "revolution",[59] is an original form of interpretation of the *FAA* established by the Supreme Court in recent years. Unlike the *Lochner* era, the Court's new jurisprudence is aimed not at legislators but at the work of judges in lower courts who have used equity principles, and mainly the doctrine of unconscionability, to avoid enforcement of what they have perceived as unfair arbitration agreements. What is exceptional about this approach is that — parallel to the *Lochner* era — the Supreme Court has taken the freedom of contract to a new level, making it a concept that has the power to de-authorize and delegitimize state actors as they make efforts to prevent injustice.

At the beginning, long before this revolution, the *FAA* was an act of legitimization. Its enactment in 1925 was interpreted as establishing the once-doubted freedom of contractual parties to agree on private dispute resolution outside of the public legal system. At that period, enforcement of arbitration agreements was only a way to publicly support the freedom of contract of parties with similar bargaining power(s) who chose to negotiate and conclude an agreement to arbitrate their future disputes.

---

58. See *e.g.* David Singh Grewal & Jedediah Purdy, "Introduction: Law and Neoliberalism" (2014) 77:4 Law & Contemporary Problems 1. Defining "neoliberalism" is a hard task that is beyond the scope of this work. However, the discussion that follows explains some of the main features of neoliberalism. See generally, David Harvey, *A Brief History of Neoliberalism* (Oxford: Oxford University Press, 2005) at 206.
59. Glover, "Disappearing Claims", *supra* note 16.

In those times, most corporations and big businesses did not use this freedom in their relationships with their customers, workers, franchisees, or other weaker parties.[60] Indeed, in 1953, the Supreme Court indicated that enforceability is limited to transactions made at arm's length as opposed to situations in which the plaintiff "had no choice but to accept the arbitration stipulation, thereby making the stipulation an unconscionable and unenforceable provision in a business transaction".[61]

Next came privatization. In 1983, time that importantly accords with the inception of the neoliberal age, the Supreme Court announced that the *FAA* reflects "a liberal federal policy favoring arbitration agreements" and therefore "any doubts concerning the scope of arbitrable issues should be resolved in favor of arbitration".[62] In the years that followed, the Court consistently applied its pro-arbitration policy, expanding the *FAA*'s reach far beyond enforcing agreements between businesses with similar powers. Most importantly, the Court affirmed, and by that encouraged, an increasing use of arbitration clauses in contracts of adhesion drafted by legal teams working at the service of big corporations. Those form contracts compelled millions of weaker parties — consumers, workers, clients, patients, franchisees, and other would-be-claimants — to forgo public litigation and instead commit to resolving disputes in arbitration.[63] To enhance the transference of disputes from the public to the private system, many restrictions on the types of claims considered arbitrable were removed, giving drafters of form contracts the leeway to subject to arbitration a variety of statutory rights which never before were discussed in arbitration, including those protected under antidiscrimination laws.

As in any other act of privatization, the shift from public to private legal services was followed by a transformation of the service itself.

60. J Maria Glover, "Beyond Unconscionability: Class Action Waivers and Mandatory Arbitration Agreements" (2006) 59:5 Vanderbilt Law Review 1735 at 1740 [Glover, "Beyond Unconscionability"].
61. *Wilko v Swan*, 346 US 427 (1953) at 440.
62. *Moses H Cone Memorial Hospital v Mercury Construction Corp*, 460 US 1 (1983) at 24-25.
63. David Horton, "The Shadow Terms: Contract Procedure and Unilateral Amendments" (2010) 57:3 UCLA Law Review 605 at 621-22.

Using their excessive power and their control of the drafting process, corporations not only insisted on transferring litigation to a private forum, but also invested resources in changing the rules governing the private process and its results. Form contracts were used, for example, to limit discovery rights, shorten limitation periods, and eliminate specific remedies.[64]

The judicially supported privatization of dispute resolution services combined with the ability to modify the rules of private litigation in a way that would better serve the interests of the drafters of the arbitration agreements certainly sacrificed the interests of their weaker counterparts. And yet, before the most recent change of jurisprudence, this sacrifice was still somewhat cabined mainly by the idea of unconscionability. The Court's pro-arbitration approach prior to the revolution was explained by the virtues that the Court attributed to the mechanism of arbitration. Time and again, courts emphasized that their willingness to enforce the arbitration agreement is based on their belief, or assumption, that arbitration agreements benefit *both parties* by offering them an effective path of dispute resolution. Following this logic courts refused enforcement when the agreement dictated an arbitration that was clearly designed to harm one of the parties. Accordingly, courts used the doctrine of unconscionability to invalidate, for example, an arbitration agreement that was designed to make the arbitration process prohibitively expensive for purchasers of computers.[65]

The most significant change brought by this court-approved privatization was the imposition of class action waivers. Those waivers — which quickly became part of every standard arbitration agreement — work to ensure that weaker parties remain isolated from each other and forego the ability to join others similarly harmed by a corporation.[66] As we shall soon see, it is this change that eventually gave rise to the Supreme

64. Glover, "Beyond Unconscionability", *supra* note 60 at 1742.
65. See *e.g. Brower v Gateway 2000, Inc*, 676 NYS (2d) 569 (App Div 1998) (invalidating a term requiring arbitration due to a minimum up-front fee of $4,000 and explaining that such fee would "deter consumers from invoking arbitration" at 573).
66. Horton, *supra* note 63 at 631-32.

Court's new arbitration jurisprudence. Instead of merely turning a public process into a private one, those waivers have a special potential to harm weaker parties by effectively preventing them from the pursuit of their rights, even in a limited private forum. This effect is especially relevant within the context of claims that have a relatively small monetary value due to the cost of litigation far exceeding the potential reward. In those cases, enforcement of the arbitration agreement does not lead to private litigation but rather to the prevention of *any* litigation, vanquishing any legal right(s) the weaker party may have against the stronger party. This potential of class arbitration waivers to minimize or even eliminate litigation did not escape the awareness of drafters of form contracts, and the waivers became a widespread strategy used by big businesses in their contracts with consumers, employees, and other weaker parties. In this age of privatization, many courts enforced class arbitration waivers as part of the protected freedom of contract and despite their infringement on weaker parties' rights.

And yet, in 2005, California's Supreme Court famously responded to the problem by declaring class arbitration waivers unconscionable provided that these waivers were not negotiated and practically amounted to a deprivation of the right to litigate not only in courts but also in arbitration. Importantly, this admittedly narrow limitation was linked by the court to the idea of an exploitation of the legal right to contract by those with excessive bargaining power. Creating the *Discover Bank v Superior Court*[67] ("*Discover Bank*") rule, which would later be abrogated by the new arbitration jurisprudence, the California Supreme Court emphasized the need to limit in certain situations — and via the doctrine of unconscionability — the freedom of contract of the stronger party. The Court explained:

> when the [class arbitration] waiver is found in a consumer contract of adhesion in a setting in which disputes between the contracting parties predictably involve small amounts of damages, and ... the party with the superior bargaining power has [allegedly] carried out a scheme to deliberately cheat large numbers of consumers out of individually small sums of money, then ... the waiver ... [is] unconscionable under California law and should not be enforced.[68]

67. 113 P (3d) 1100 (Cal Sup Ct 2005) [*Discover Bank*].
68. *Ibid* at 1110.

For several years, courts followed the *Discover Bank* rule and applied it both in the context of consumer contracts as well as in some employment contracts.[69] This is precisely what the California Court of Appeals did in the litigation between *AT&T* and its consumers who contested the fact that the company charged them taxes on mobile phones it advertised as being given for free.[70] Since the overcharge was relatively small and could not possibly allow for an individual litigation, the consumers initiated a class action thereby creating a need for the courts to decide the validity of the class arbitration waivers that were included in all of *AT&T*'s contracts. Following the rule of *Discover Bank*, the court of appeals decided to invalidate the waivers as unconscionable. The Ninth Circuit Court of Appeal affirmed.[71] And then came the "revolution".

After granting a *certiorari*, the Supreme Court decided in 2011 to abrogate the *Discover Bank* rule, forbidding courts around the country to use the doctrine of unconscionability in order to invalidate class arbitration waivers. This, of course, was not the first time the Supreme Court reversed a lower court's decision ordering the enforcement of class arbitration waivers in consumer contracts.[72] However, before *AT&T,* such a reversal was explained by a belief that the terms were not as inequitable as the lower court had seen them. And here lies the revolution: under the new jurisprudence the arbitration agreement is enforceable even when it is inequitable. In *AT&T* the Supreme Court emphasized for the first time that, when it comes to arbitration, freedom of contract trumps even when it is strategically used by the stronger party, not to channel litigation away from courts and into the private sphere, but rather *to avoid litigation*

69. The *Discover Bank* rule relates to consumers' contracts of adhesion. It was later extended to employment contracts: see *Gentry v Superior Court*, 165 P (3d) 556 (Ca Sup Ct 2007) [*Gentry*]. But compare with Peter Danysh, "Employing the Right Test: The Importance of Restricting AT&T v Concepcion to Consumer Adhesion Contracts" (2013) 50:5 Houston Law Review 1433 at 1462 (arguing that the court in *Gentry* ruled only on a claim for overtime pay pursuant to the Labor Code and its decision should not be read in a broader way).

70. *Laster v T-Mobile USA Inc*, 407 F Supp (2d) 1181 (SD Cal 2005).

71. *Laster v AT&T Mobility LLC*, 584 F (3d) 849 (9th Cir 2009).

72. See *e.g. Green Tree Financial Corp-Alabama v Randolph*, 531 US 79 (2000).

altogether. The Court confirmed corporates' opportunistic drafting of contracts in a way that would block both the path to public litigation *and* the path to collective private litigation, even in cases where the third path — of individual arbitration — is undoubtedly futile. Notably, the new jurisprudence requires lower courts to refrain from using equity principles, and especially the unconscionability doctrine, even when one party — namely a big corporation — has abused its freedom of contract, creating a set of terms appearing like an agreement to settle disputes in arbitration but realistically having the opposite goal and effect: to make any review of disputes as difficult and improbable as possible.

Later efforts of lower courts to narrow the new *AT&T* rule of unlimited freedom and disempowered courts failed. In fact, two years after *AT&T* the Supreme Court extended its new jurisprudence even further, clarifying that the *FAA* preempted not only protections awarded under contract state law — mainly via the doctrine of unconscionability — but also rights secured under Federal law. In a litigation between American Express and a small restaurant named Italian Colors, the Court affirmed a class arbitration waiver that effectively *prevented* arbitration with regard to rights secured under Federal antitrust laws.[73]

What is unique about this arbitration jurisprudence and what marks it as "new", or even "revolutionary", is the willingness of the Court to approve and enforce any product of the freedom of contract as exercised by stronger parties, regardless of the consequences to the other party and to society as a whole. While this is not the first time we have seen a strong pro-arbitration jurisprudence that admittedly shrinks the ability of weaker parties to enforce their rights, in the past there were some safeguards and some limitations that remained in place. Specifically, until recently, lower courts always kept the tools of equity at their disposal. And, what is more, the legitimacy of utilizing those equity tools is explicitly guaranteed under the *FAA* itself. In its relevant section, the *FAA* clarifies that arbitration agreement "shall be valid, irrevocable, and enforceable, save upon such grounds as exist at law or *in equity* for the

73. *American Express*, *supra* note 14.

revocation of any contract".[74]

In other words, never before *AT&T* and *American Express* had it reached a point in which the Court recognized the harm to the weaker party but responded by saying that it is, to quote Justice Kagan's dissent, "[t]oo darn bad".[75] And never before had the *FAA* been interpreted in a way that would officially offer corporations, to quote Kagan J again, "de facto immunity",[76] and "a foolproof way of killing off valid claims".[77] Finally, never before were courts categorically forbidden from using equity tools to overcome extreme injustice caused by opportunistic use of the law.

B. The Meaning of New Arbitration Jurisprudence

What is the meaning of such new jurisprudence? At its most basic level, it reflects a heyday of the freedom of contract, comparable to the rise of the idea during the *Lochner* era. Again, we are witnessing a jurisprudence that refuses to limit the freedom of contract regardless of the consequences of such an unfettered version of freedom. And, similar to its predecessor, this second rise of the freedom of contract includes de-legitimization and a rejection of any state response that attempts to protect the rights of weaker parties and/or maintain a minimal level of social justice. To the extent that equity stands for the general quest for fairness and justice and for inhibiting the exploitation of the formality of law, the new jurisprudence does not leave room for it.

Moreover, and still in parallel to the *Lochner* era, the new jurisprudence aligns itself with the interests of stronger parties, such as those in *AT&T* and *American Express*, without admitting to it. The effect

74. *FAA*, *supra* note 12 [emphasis added].
75. *American Express*, *supra* note 14 at 2313.
76. *Ibid* at 2315.
77. *Ibid.* For this aspect of the new jurisprudence see Glover, "Disappearing Claims", *supra* note 16 (arguing that the "new approach erodes substantive law itself by empowering private parties, through contract, to frustrate or altogether eliminate claiming in any forum, and thereby to rewrite the scope of their obligations under substantive law" at 3066).

is achieved by celebrating, and relentlessly enforcing, contracts drafted by strong parties and imposed on weaker parties, while portraying the contractual process as equal and reciprocal. To see that suffice is to pay attention to Scalia J's rhetoric. In *AT&T,* Scalia J presents the logic of the freedom of contract in the context of arbitration, stating, "[t]he point of affording parties discretion in designing arbitration processes is to allow for efficient, streamlined procedures tailored to the type of dispute".[78] This is the chosen rhetoric despite Scalia J's awareness of the fact that the consumers in this case, and in any other case for that matter, lacked any freedom to design the arbitration process. Such awareness is expressed when Scalia J concedes, albeit in a different part of his decisions, that "the times in which consumer contracts were anything other than adhesive are long past".[79] And yet, awareness of this reality does not stop Scalia J from presenting the contractual process as free and further referring to its result — six times along the decision — as establishing a "bilateral" arbitration,[80] implying that both parties exercised their freedoms and chose arbitration as the best way to mutually resolve their future disputes.

So far the new arbitration jurisprudence may seem as though it is mainly a revival of the *Lochner* era. However, much more is at stake. First of all, applying *Lochnerian* ideas to our times exponentially magnifies their effect. While old *Lochnerism* glamorized the freedom of contract mainly in the domain of industrial labor, the new jurisprudence covers incalculably larger amount of contracts that impact consumers, patients, clients, borrowers, franchisees, small investors, and virtually all other individuals and businesses who do not belong within a powerful "corporate America". Similarly, while *Lochnerism* limited state legislators from interfering in the market, the new jurisprudence additionally restricts, as seen in *AT&T* itself, the work of courts around the country. And finally, while *Lochner* deprived weaker parties of rights awarded to them under law, the new jurisprudence, as I will further discuss later, *additionally* deprives them of rights arising from the principles of equity. In other words, part of what is new here is the magnitude of the phenomenon of

78. *AT&T*, *supra* note 13 at 1749.
79. *Ibid* at 1750.
80. *Ibid* at 1745, 1749-51.

celebrating the freedom of contract at the expense of other values.

However, this difference in scale, devastating as it is due to its role in intensifying inequalities and injustice, is only the tip of the iceberg. "More is at stake", writes political scientist Wendy Brown, "than support for capital in the name of freedom".[81] In the coming part I will argue and demonstrate that the new-*Lochnerism* amounts to a "neoliberal-*Lochnerism*" and that what we are witnessing is no less than a salient and very troubling phase in the dissemination of neoliberalism — one that thwarts equity principles both in theory and in practice.

C. The New Jurisprudence as a Neoliberal-*Lochnerism*

To be able to evaluate the magnitude of the risk, it is imperative to consider neoliberalism beyond all the ways in which it promotes economic policies aligning with the idea of a free market. Conceived by Pierre Bourdieu as "a political project", neoliberalism works *outside* the economic field, not only within it. Indeed, careful observation with a critical eye reveals that neoliberalism stealthily operates to create a new "order of normative reason",[82] to redefine rationality, and to enforce itself as the common sense of all subjects, while denying the possibility of other logics. In this way, neoliberalism has the ability to change the way we understand the world, process our experiences, and respond to challenges. As Margaret Thatcher once said: "[e]conomics are the method, but the object is to change the soul".[83] But how is the objective of changing souls achieved? As Wendy Brown argues, *law* is an important medium through which neoliberalism disseminates its logic beyond the economy; and, as I will argue next, the new arbitration jurisprudence provides a powerful demonstration of such an operation.

The legal reasoning of the Supreme Court in the cases creating the new arbitration jurisprudence works to disseminate neoliberal rationality and to revise our common sense in at least three major ways. First, it frames the legal question of arbitration contracts' enforceability solely

81. Brown, *supra* note 20 at 153.
82. *Ibid* at 30.
83. *Ibid* at 153.

in neoliberal terms, creating a strong economized discourse in an arena formerly belonging to the justice system: that of dispute resolution. This strong discourse silences and vanquishes a host of other discourses, with the main victim being — as later further discussed — an equity discourse aimed at the prevention of unconscionable conduct. Second, and relying on this economized discourse, the same legal reasoning operates as an assault on the collective agency of legal subjects, applying a "divide and conquer" tactic that not only isolates and wears off weaker subjects but also, and alchemically, makes *them* — and them alone — responsible for their poor fate. And third, this neoliberal reasoning is gaining broad popularity outside the elite because it is presented as universal, while the fact that it is biased and reflects only the viewpoints of the most powerful market actors is carefully disguised. All in all, the new rationality takes over not only the economy and the market, but also the justice system. Consequently, and for the second time in history, the law conflicts with equity and all it stands for. This time, however, King James I is not there to side with equity.[84] In what follows I will demonstrate each of the three neoliberal moves that, together, create this effect.

1. Economized Discourse

Under the new rule of *AT&T,* the *FAA* is interpreted as preempting the state-level contractual doctrine of unconscionability. Similarly, under *American Express*, the *FAA* preempts the doctrine of effective vindication of Federal rights. Together, these two decisions may leave parties with inferior bargaining power with no legal recourse against the big corporations that had them sign a class arbitration waiver. But why? What is the shared logic that justifies this dramatic result? The key, according to Scalia J and the four other justices that sided with him, is *efficiency.* And not just an abstract notion of efficiency: as Scalia J explains and then reiterates in *AT&T*, it is all about economic efficiency, concretely measured by "costs", "savings", and "speed".[85] Instead of the "costliness

84. King James I famously decided to favor equity after the *Earl of Oxford's Case in Chancery* (1615), Rep Ch 1 (Eng) at 7 [*Earl of Oxford*].

85. *AT&T*, *supra* note 13 at 1751-52.

and delays of [public] litigation", states Scalia J as if he were the proud CEO of a successful firm, private arbitration offers "greater efficiency", and allows for "streamlined proceedings and expeditious results", while "reducing the cost and increasing the speed of dispute resolution".[86] Similar rhetoric is used by Scalia J again in *American Express* both by him citing the words of *AT&T* and by branding arbitration as a "speedy resolution".[87]

Admittedly, the mechanism of arbitration is not perfect even according to Scalia J, as without judicial review there is an increased risk that "errors will go uncorrected".[88] However, to further economize the discourse, those possible errors, too, are framed not as justice-related issues, but rather are described in market terms. Translated to economic lingo those potential errors are analyzed as "costs". Moreover, the proliferation of arbitration contracts despite the problem of errors — now re-termed as "costs" — is itself explained based on "the logic of profit-making".[89] As Scalia J points out, drafters of arbitration contracts still prefer to opt out of the public justice system because it *makes sense* to do so under the only rationality presented in the case — that of a cost and benefit analysis. He explains: "[d]efendants are willing to accept the costs of these errors in arbitration, since their impact is limited to the size of individual disputes, and presumably *outweighed by savings* from avoiding the courts".[90]

The constant use of the idea and jargon of "incentives" further transforms the issue of arbitration into a purely economic one. For example, as an important part of his legal reasoning, Scalia J stresses in *AT&T* that invalidating class arbitration waivers "will have a substantial deterrent effect on incentives to arbitrate".[91] Similarly, in *American Express*, Scalia J translates the argument of the restaurant that was not able to fund an individual litigation into "Economish" (the language of

86. *Ibid* at 1749, 1751.
87. *American Express*, *supra* note 14 at 2312.
88. *AT&T*, *supra* note 13 at 1752.
89. Weiner, *supra* note 21 at 20.
90. *AT&T*, *supra* note 13 at 1752 [emphasis added].
91. *Ibid.*

economics), presenting the problem not as a lack of access to justice,[92] but as having "no economic incentive to pursue [the restaurant's] antitrust claims individually in arbitration".[93] And, surrendering to Scalia J's rhetoric, the dissent in *American Express* replies that with the enforcement of class arbitration waivers, "companies have every incentive to draft their agreements to extract backdoor waivers of statutory rights, making arbitration unavailable or pointless".[94]

The economized message of the new jurisprudence is additionally amplified by the type of evidence offered by the court. To "prove" the efficiency argument, Scalia J presents empirical data often used in the domain of economics: statistics. The information, produced by the American Arbitration Association as *amicus curiae*, is introduced due to its ability, according to Scalia J, to illustrate *in numbers* both the efficiency of individual arbitration and the inefficiency of class arbitration.[95] This use of data is yet another signature move of neoliberalism where all things, including justice systems, could and should be measured in numbers and in terms of productivity. In order to decide whether to enforce arbitration contracts, the theory goes, we only need to measure how many disputes were resolved per time unit. Thus, "speed" is a repeating theme in *AT&T* and the arbitration proceedings are evaluated solely by their ability to bring a dispute to a quick resolution.

Notably, despite their appearance, numbers and statistics are far from being objective measuring tools and can be manipulated. For the purpose of calculating the speed and quantifying the efficiency of legal proceedings, Scalia J assigns the concept of "resolution" with an oddly narrow meaning. For him, the term resolution only refers to the production of "judgment on the merits", while other possible resolutions, such as "settlement[s], withdrawal[s], or dismissal[s]" do not

92. See *e.g.* Charles L Knapp, "Taking Contracts Private: The Quiet Revolution in Contract Law" (2002) 71:3 Fordham Law Review 761 (describing arbitration as the "denial of access … to the law itself" at 782).
93. *American Express*, *supra* note 14 at 2310.
94. *Ibid* at 2315.
95. *AT&T*, *supra* note 13 at 1751.

count.[96] The choice to define a resolution so narrowly may be misleading, especially given the fact — recognized by Scalia J in another part of the decision — that class arbitrations often do not end in judgments on the merits precisely because defendant companies *prefer* to settle in order to avoid negative precedents. In any case, while the logic of defining a resolution in such an unusual manner is not explained in the case, it surely helps in presenting individual arbitrations as more "efficient" than class arbitrations. This special way of counting simply makes individual arbitrations produce more "resolutions".

To conclude: by heavily using an economized efficiency rhetoric, applying a cost and benefit analysis, committing to incentive thinking, and supporting arguments with only quantitative measures, the legal reasoning in *AT&T* and *American Express* amounts to the building of what Bourdieu called a "strong discourse".[97] What Scalia J achieves through creating such discourse is an "economization" of the domain of dispute resolution. As any other economization, this too is done by "extending a specific formulation of economic values, practices, and metrics"[98] to noneconomic dimensions of life, in this case the functioning of our justice system as well as the civic rights and values that the system stands for. And, because economization targets noneconomic domains, it works to disseminate neoliberalism not just as an economic approach, but more so as a rationality, defining a new common sense for all.

Another salient component of the neoliberal economization of everything is the strong negation of alternative rationalities. To use economic terms: economization is also an effort to create a neoliberal monopoly within the "market of ideas", a monopoly that would have the unleashed power to crush already existing competing logics and prevent the emergence of new ones. A few examples from *AT&T* and *American*

---

96. *Ibid*.
97. Pierre Bourdieu, "Utopia of Endless Exploitation: the Essence of Neo-liberalism", *Le Monde Diplomatique* (December 1998), online: Le Monde diplomatique <http://mondediplo.com/1998/12/08bourdieu> ("neoliberal discourse is not just one discourse among many"; rather, it is a "strong discourse") [Bourdieu].
98. Brown, *supra* note 20 at 30.

Express can illustrate this feature. First, in *AT&T*, Scalia J determines that "contrary to the dissent's view", his interpretation of the *FAA* as a norm designed to promote arbitration is "beyond dispute".[99] The use of such language expresses not only disagreement with the dissent, but also a denial of the very effort to challenge the economized analysis and its presumption that arbitration is always positive (read: "efficient"). Put differently, once established that arbitration makes economic sense, the discussion ends and the matter becomes "beyond dispute" — so, it must be that the *FAA* means prioritizing arbitration under all circumstances and regardless of the consequences.

Second, initially in *AT&T* and later in *American Express*, Scalia J dismisses the issue of injustice, calling it "unrelated". Remarkably, situations in which the enforcement of arbitration agreements consequently leaves plaintiffs without neither public nor private path to redress is somehow classified as irrelevant to the discussion. To clarify: my point here is not that Scalia J had an obligation to agree with the dissent. Quite to the contrary, I argue for the importance of pluralism of approaches and against the monopolization of economized logic. Thus, it is not the disagreement between the justices that is my concern, but the attempt of Scalia J to deny *the relevance* of the dissent's reasoning. How and why concerns about the unavailability of remedies are "unrelated" to the very matter of private dispute resolution are questions that remain unanswered by Scalia J, and thus must be answered by the underlying claim of neoliberalism to exclusivity. Indeed, this last proposition is confirmed in *American Express*, when Scalia J bluntly dismisses the possibility of a competing rationality, declaring an approach that prefers the goals of antitrust law to the goals of arbitration as nothing but "simply irrational".[100] The overall result is a claim that the economized understanding of the issue of private dispute resolutions is the only logical understanding and other perspectives are not less desirable, but rather, no longer imaginable or worth raising.

Interestingly, in both *AT&T* and *American Express*, the dissent

99. *AT&T*, *supra* note 13 at 1749.

100. *American Express*, *supra* note 14 at 2309.

attempted to resist the monopoly of economized thinking. In *AT&T* Justice Breyer contests the need to make decisions based on efficiency as the sole criteria, stating: "[t]he intent of Congress requires us to apply the terms of the Act without regard to whether the result would be possibly inefficient".[101] Likewise, in *American Express* Kagan J attacks the economized logic by pointing out its absurd result which offers corporations "a foolproof way of killing off valid claims", instead of securing a "method of resolving disputes".[102] Furthermore, Kagan J goes beyond rejecting the logic of efficiency and insists on introducing significant fairness concerns. Enforcing class arbitration waivers when the plaintiff has no real way to use the path of individual arbitration, argues Kagan J, deprives the plaintiff "of his day in court",[103] and "confers immunity on a wrongdoer".[104]And yet, as compelling as this resistance is, what is missing, perhaps due to long years of neoliberal dominance, is a fuller model of thinking; an alternative discourse which could be strong enough to counter the one produced by neoliberalism. As I later further discuss, it is at this point that equity-based discourse could have proved invaluable.

In any case, Scalia J's response to Kagan J's dissent in *American Express* further demonstrates that his economized logic tolerates no alternative approaches. "Truth to tell", he writes quite impatiently, "our decision in *AT&T* Mobility all but resolves this case".[105] However, since what was at stake in *AT&T* was state law and what was at stake in *American Express* was federal law, it is unlikely that whatever was already decided in *AT&T* could directly settle the dispute in *American Express*. Perhaps what Scalia J means to say in this part of his decision is that in *AT&T* neoliberalism had already won the race between competing rationalities and therefore Kagan J's dissent is repetitive in trying to challenge the victory; there is simply no room to reopen the question of *how*, that is according to which

101. *Dean Witter Reynolds Inc v Byrd*, 470 US 213 at 219 (1985) [*Dean Witter*] as cited in *AT&T*, *supra* note 13 at 1758.
102. *American Express*, *supra* note 14 at 2315.
103. *Ibid* at 2314.
104. *Ibid* at 2318.
105. *Ibid* at 2312.

rationality, to analyze the issue of class arbitration waivers.

2. An Assault on Collective Agency

Attention to rhetoric will also reveal that in both *AT&T* and *American Express* the Supreme Court did more than economize the issue of class arbitration waivers. Portraying class arbitrations as "slow" and "inefficient" is a necessary step but not a sufficient one for achieving the greater goals of the neoliberal project. Thus, an additional effort was made to represent class arbitrations as dangerous and malicious. Reading *AT&T*'s legal reasoning one first learns that class arbitrations are illegitimate because they are, to cite Scalia J, "manufactured"[106] by the pre-revolution *Discover Bank* rule. Second, readers are told about the specter of chaos when, in a move typical to neoliberalism, the very idea of collective dispute resolution becomes a metaphor for public disorder.[107] Allowing class arbitrations, writes Scalia J in *AT&T*, "is likely to generate procedural *morass*".[108] And, proving that he chose his words intentionally and that these words are essential to his legal reasoning, Scalia J uses the same rhetorical reference to chaos, word for word, in *American Express*. Additionally, and in the same intimidating vein, the Court cautions that if courts are obligated to ensure that an individual resolution is viable as a condition to enforcing a waiver of class arbitration, the result would be a "litigating *hurdle*",[109] which "would undoubtedly *destroy* the prospect of speedy resolution".[110]

And third, in addition to describing class arbitrations as a manufactured, chaotic, and destructive method of dispute resolution, the procedure is also represented as unfair towards those being sued. According to this approach, class arbitrations are not a way for powerless claimants to make their small (*AT&T*) or prohibitively expensive

106. *AT&T*, *supra* note 13 at 1751.
107. Henry A Giroux, "Neoliberalism, Corporate Culture, and the Promise of Higher Education: The University as a Democratic Public Sphere" (2002) 72:4 Harvard Educational Review 425 at 428.
108. *AT&T*, *supra* note 13 at 1751 [emphasis added].
109. *American Express*, *supra* note 14 at 2312 [emphasis added].
110. *Ibid* [emphasis added].

(*American Express*) claims. Instead, according to Scalia J, class arbitrations are a vicious tool utilized to bully corporations "into settling questionable claims".[111] And, tantamount to class actions carried in courts, private class arbitrations too carry "the risk of 'in terrorem' settlements".[112] The artificial victimization of powerful corporations will be further discussed in the next section, but for the time being the main point is this: the assault on collective legal actions includes judicial effort to demonize class arbitrations.

Awareness of the Court's evident hostility to class arbitration as expressed in *AT&T* and *American Express* can aid in understanding why the new arbitration jurisprudence belongs with the neoliberal project. If neoliberalism is, as theorized by Bourdieu, "a programme of the methodological destruction of collectives",[113] then the ongoing judicial attack on class arbitrations fits the premise. It is aimed at preventing consumers, workers, small businesses, and the like, from creating legal collectives. Furthermore, such an approach deprives individuals with less market power of the one tool that may help them protect themselves: the ability to "band together to fight corporate abuses".[114]

But why does neoliberalism attack collectives? The rationale may be clear if we recall that the suppression of both counter ideas and attempts of resistance rests at the core of the neoliberal aspiration for hegemony. For that reason, when collectives try to use joint power as a method of countering the power of big businesses, neoliberals identify a target for attack. From a neoliberal perspective, and in a paradoxical treatment of autonomy and freedom, "while individuals are supposedly free to choose, they are not supposed to choose to construct strong collective institutions".[115] The general neoliberal assault on collectives, and the particular attack on institutionalized legal collaborations such as class procedures, is fundamentally a way to divide and conquer.

And yet, even if the assault on collectives makes sense from a neoliberal

111. *AT&T*, *supra* note 13 at 1752.
112. *Ibid.*
113. Bourdieu, *supra* note 97.
114. Brown, *supra* note 20 at 153.
115. Harvey, *supra* note 58 at 69.

viewpoint: why would the Court get involved and spread the anti-solidarity message? Since the judiciary is an arm of the state, the answer has to do with the general role of the state under the neoliberal scheme. The neoliberal state is supposed to have no public goals and is instead expected to serve the interests of powerful market actors. As a result, explains David Harvey, "the neoliberal state is necessarily hostile to all forms of social solidarity that put restraints on capital accumulation".[116] And, if need be, the state will use its legal arm, resorting "to coercive legislation and policing tactics … to disperse or repress collective forms of opposition to corporate power".[117] Notably, this understanding of the role played by the Court accords with Kagan J's metaphorical statement in *American Express*. Blaming the majority for promoting an ongoing and deliberate agenda against collective legal procedures, class actions and class arbitrations alike, Kagan J writes: "[t]o a hammer, everything looks like a nail. And to a Court bent on diminishing the usefulness of [class procedures] everything looks like a class action, ready to be dismantled".[118]

But the assault on collective actions not only forces people to cope with power in isolation. Rather, its effect is enhanced by the neoliberal signature practice of *responsibilization*. As demonstrated by the history of dispute resolution, the neoliberal order first leaves subjects to the mercy of the market via the process of privatization, and then assigns to them liability for whatever harm they suffer while operating in the market. If under *Lochnerian* classical liberalism the contractual "choice" was framed as somehow serving the self-interest of the person who agreed to work for very long hours, this is no longer the case today. Neoliberalism has long abandoned the idealization of individuals' self-interests, freely defined, as well as the belief that an invisible hand ensures the outcome is beneficial for all. Instead, subjects are reconfigured as self-investors and self-providers and — regardless of their interests — are expected to align themselves with what would best serve the economy.[119]

116. *Ibid* at 75.
117. *Ibid* at 77.
118. *American Express*, *supra* note 14 at 2320.
119. Brown, *supra* note 20 at 84.

Accordingly, because individual arbitrations — as we have seen — are presumed "efficient", *i.e.* making economic sense, it is the responsibility of individuals alone, and not of the state, to use arbitrations to solve their disputes with each other. And, when individuals neglect to self-provide for justice by failing to arrange for a mechanism that actually secures resolution, they have no one else but themselves to blame. In other words, people — and not the state — are now responsible for both *supplying* a private mechanism of dispute resolution and for the *quality* of the resolution the mechanism yields.

It is important to note that neoliberal "responsibilization" is an imposed process that significantly differs from a liberal exercise of human autonomy and agency. As explained by Wendy Brown, people are being responsibilized because neoliberalism, for its own political goals, "solicits the individual as the only relevant and wholly accountable actor".[120] Such responsibilization is evident in the legal reasoning of Scalia J in both *AT&T* and *American Express*. The Court goes a long way to place gaps of power aside and to emphasize the binding contract to which the weaker party in each case had committed itself to. As opposed to the times of *Lochner*, the obligation to arbitrate of the consumers in *AT&T* and the small restaurant in *American Express* is no longer portrayed as serving their interests and thus as deserving enforcement. Rather, the justification for enforcement stems directly from formalistically framing the entire issue a contractual matter and then responsibilizing the parties who signed the dotted line.

In *AT&T*, Scalia J opens his decision in describing the contract as if it were an agreement between equals, declaring: "Vincent and Liza Concepcion entered into an agreement for the sale and servicing of cellular telephones with AT&T Mobility LCC".[121] Next, readers learn that the contract included a duty to arbitrate and, when AT&T was sued, it sought nothing more than to enforce its rights "under the terms of its contract with the Concepcions".[122] To remove any doubt regarding the appropriate nature of the discussion Scalia J reminds readers that

120. *Ibid* at 133.
121. *AT&T*, *supra* note 13 at 1744.
122. *Ibid* at 1744-45.

"arbitration is a matter of contract",[123] strongly implying what it is not: a privatized form of dispute resolution carried as an alternative to a service once supplied exclusively by the state via the court system. And, defined as a contract, arbitration is, according to this logic, merely a matter of business exchange in which one simply gets what one bargained for. Or, as Scalia J describes the bargain: "parties forgo the procedural rigor and appellate review of the courts in order to realize the benefits of private dispute resolution".[124]

But what if the bargain conflicts with the interests of one of the parties? That, we learn, has no room in the analysis. The only line of reasoning is repeated, as if it were a mantra: "[a]rbitration is a matter of contract, and the *FAA* requires courts to honor parties' expectations".[125] And in this way, the Concepcions, like all people not lucky enough to receive services from an in-house legal department that drafts smart contracts, are configured as self-investors who may have made a bad investment decision. Did they have any alternatives? Could they have negotiated better with AT&T or purchased their mobile phones from another corporation that allows class arbitration in the event that things go wrong? Those questions lie outside of the responsibilizing neoliberal analysis under which responsibility is imposed notwithstanding context.

Similar logic is applied in *American Express*. Once again, Scalia J opens with representing the contract as if it suffered from no gaps in the bargaining power of the parties. Italian Colors, described by the dissent as "a small restaurant",[126] is represented by the majority as "merchants who accept American Express cards".[127] Next, the first and primary fact readers learn about those "merchants" is that "[t]heir agreement with … American Express … contains a clause that requires all disputes between the parties to be resolved by arbitration".[128] Echoing *AT&T*, the Court then declares that "arbitration is a matter of contract", adding that "courts

123. *Ibid* at 1752.
124. *Ibid* at 1751.
125. *Ibid* at 1752.
126. *American Express*, *supra* note 14 at 2313.
127. *Ibid* at 2308.
128. *Ibid*.

must 'rigorously enforce' arbitration agreements".[129] But what happens to the concern for "the policies of the antitrust laws"[130] that might be compromised if small restaurants cannot have any legal recourse against American Express? At this moment responsibilization seems to be the Courts' only way of answering: small restaurants have "to litigate their claims individually — *as they contracted to do*".[131]

It is important to recognize that responsibilization is a key stage in the "divide and conquer" strategy of neoliberalism. After institutionalized solidarities — such as class legal procedures — are extinguished, what remains is an isolated entity. At this point responsibilization finishes off what anti-solidarity has started: the subject, and the subject alone, is held responsible for the consequences of his or her behavior. Moreover, the assault on legal solidarities and the responsibilization that follow it are not only *facilitated by* the economization of everything, but are also *facilitators of* enhanced economization. At the most basic level economization invites anti-solidarity and responsibilization because, under economic measures, the inefficiency of class arbitrations justifies their abolition and the frame of a bargain (or contract) calls for "rigorous enforcement" of class arbitration waivers. However, perhaps less noticeably, anti-solidarity and responsibilization also operate to establish the supremacy of economized rationality. It is precisely because subjects are responsibilized that social questions regarding fairness and justice can be dismissed as "unrelated", and not worth serious consideration. Due to the totality of responsibilization, which assigns a hundred percent of the responsibility to individual subjects, other candidates are completely released from responsibility. Accordingly, the Court does not have to discuss questions of fairness and justice because these issues were already fully resolved by responsibilizing the parties who signed arbitration agreements that effectively deprived them of access to justice. Or, as Kagan J more straightforwardly phrases it, when parties to contracts find themselves alone and with no legal recourse, it is "[t]oo darn bad";[132] no

129. *Ibid* at 2309.
130. *Ibid.*
131. *Ibid* [emphasis added].
132. *Ibid* at 2313.

other alternative can be imagined.

3. A Biased Approach in a Universal Disquise

As we have seen, the neoliberal approach to arbitration agreements aligns itself with the interests of the economic elite: arbitration is to be judged according to economic criteria, those criteria dictate resistance to collective resolution of disputes, and it is thus sensible to enforce arbitration agreements while leaving the powerless party without any recourse against their powerful drafters. However, an additional step is required in order to spread these ideas successfully and to turn them into the new common sense: the biased nature of the approach must be concealed and the supporting reasoning has to be represented as a universal truth. Indeed, the neoliberal turn was accomplished, and its results have since been maintained and enhanced, via an intentional effort to gain broad popularity and to appeal to everyone. A key strategy has been aligning neoliberalism with the American dream and establishing it "as the exclusive guarantor of freedom".[133] This general strategy is particularly powerful whenever the law is used as a medium by which neoliberal ideas are disseminated. The foundation, as we have seen, has been in place since the *Lochner* era's celebration of the idea of the freedom of contract as a universal freedom shared by employers and employees alike. In the new arbitration jurisprudence, however, freedom in the *Lochnerian* sense, defined as "the liberty of the individual in adopting and pursuing such calling as he may choose",[134] is less emphasized, which should not be surprising given neoliberalism's disregard for matters of personal fulfillment. Instead, the Court is making statements that, coming from the judiciary, seem universal and neutral but in fact represent only the perspective of the economic elite.

In reading *AT&T* we learn, for example, that the reason for allowing unlimited freedom of contract regarding the design of arbitration processes is "efficiency". This freedom can be used in many ways; one such way may be to ensure "that the decisionmaker be a specialist in the

133. Harvey, *supra* note 58 at 40.

134. *Lochner*, *supra* note 11 at 63.

relevant field, or that proceedings be kept confidential to protect trade secrets".[135] At first glance, this statement appears as a neutral description of the advantages of arbitration, one that applies equally to both parties regardless of each party's respective economic positioning. In reality, however, this is far from being the case.

First, at the most general level, counter to what is implied by Scalia J's rhetoric, it is not the "parties", in the plural, who are afforded the "discretion in designing arbitration",[136] but only the drafting party. The less powerful party has to follow the "design" imposed by the drafting party or otherwise forgo the transaction altogether — a step that would be pointless considering that all similar big businesses happen to insist on arbitrations with a similar "design". Second, another large crack in the neutral façade is the utilization of the freedom of design to protect trade secrets. Evidently, workers, consumers, patients, and other powerless parties simply do not have "trade secrets" to protect. And third, regarding the freedom to have experts decide disputes, data gathered and published by journalists and consumer advocates demonstrates that substituting public judges with specialists imposes an expensive burden often carried by plaintiffs who are required to pay filing fees.[137] For all these reasons combined, what is presented as a freedom enjoyed by both parties equally, in fact is only a freedom from the perspective of the powerful party, and in contrast, a major obstacle for the weaker party.

Furthermore, when the *AT&T* Court describes the disadvantages of class arbitrations, it does so while entirely adopting the perspective of the powerful drafters and ignoring that of their weaker counterparts. However, the biased analysis is still presented as a universal truth. For example, the Court counts as a general problem the fact that, in class arbitrations, "[c]onfidentiality becomes more difficult".[138] However, this is a detriment only from the perspective of the stronger party who is interested in fending claims off. The other, weaker party, in contrast, may

135. *AT&T*, *supra* note 13 at 1749.

136. *Ibid.*

137. Jamie Court, *Corporateering: How Corporate Power Steals Your Personal Freedom* (New York: Jeremy P Tarcher & Putnam, 2003) at 102.

138. *AT&T*, *supra* note 13 at 1750.

be quite interested in sharing information with other similarly-situated claimants. Comparably, Scalia J states that "[a]rbitration is poorly suited to the higher stakes of class litigation",[139] but readers should ask themselves: higher stakes for whom? Again, this general statement is only true for stronger parties while weaker parties, such as consumers or workers, each typically have much smaller individual claims. The same pattern appears when the Court comments that allowing class arbitrations "will have a substantial deterrent effect on incentives to arbitrate";[140] once more, this statement only applies to big businesses while the opposite is true for their weaker counterparts.

In response to the bias argument one may answer that the new arbitration jurisprudence merely expresses a belief in free markets without necessarily reflecting a biased outlook in favor of economic elites. And yet, the decision in *American Express* demonstrates that, when faced with a need to choose between a loyalty to the market and a loyalty to big businesses, the Court chose to protect the interests of the economic elite, disguising its choice with a seemingly neutral formalistic reasoning regarding the hierarchy between two Federal norms. As mentioned earlier, in *American Express* a small restaurant claimed that a giant financial company used its monopolist power in a manner that threatens the level of competition in the market. The Court, however, blocked the claim from being discussed by limiting the restaurant to an individual arbitration that it could not possibly afford. In other words, the Court was willing to enforce a class arbitration waiver even when such enforcement risks the level of competition in the market by decreasing the effectiveness of antitrust laws.[141] Such willingness demonstrates that the Court adopted a legal reasoning that reflects a specific perspective — which belongs to the drafting powerful parties — even when it meant sacrificing central principles of economic efficiency.

---

139. *Ibid* at 1752.
140. *Ibid* at 1752, n 8.
141. See Glover, "Disappearing Claims", *supra* note 16.

IV. The Fall of Equity

With its contribution to the dissemination of neoliberal rationality, the new arbitration jurisprudence does not leave much room for equitable principles and directly attacks the logic of equity. This theoretical and practical threat operates at various levels. First, and most evident, is the negative effect that recent decisions, and mainly the one in *AT&T,* have on judges' ability to use the unconscionability principle. However, my argument is not limited to this unfortunate result and seeks to go much deeper. The second problem is that in a contractual world without an effective unconscionability principle, the freedom of contract becomes a biased concept and thus a prison for most individuals. Third, and more broadly, when the law is being used to disseminate neoliberalism as a controlling rationality, the logic of equity, and with it, its theoretical justification are on the fall. Finally, as a result of this three-layer process, law itself loses touch with justice and its legitimacy is severely undermined.

A. The Fall of Unconscionability

Recall that the *FAA* itself allows courts to refuse to enforce arbitration agreements on "such grounds as exist at law or in equity for the revocation of any contract".[142] According to this clear language, there should not have been a question regarding the legitimacy of using the unconscionability doctrine to invalidate the terms of arbitration agreements because the doctrine allows revocation of contracts under both equity and the law.[143] Admittedly, even before *AT&T* many judges and theorists argued against excessive use of unconscionability for the purposes of invalidating predatory contracts, within and outside of the context of arbitration.[144] And yet, there is something fundamentally different in *AT&T*'s approach. Unlike other unconscionability opponents Scalia J is not engaging in a debate regarding the desirability of frequent use of the doctrine. Instead,

142. *FAA*, *supra* note 12.
143. In California, the state of the *AT&T* litigation, the unconscionability doctrine is part of a legislated code.
144. For a description of this approach see Keren, "Guilt-Free Markets", *supra* note 30.

his approach entirely *disallows* judges to use the doctrine in the context of arbitration. It does so despite the clear language of the *FAA* permitting the use, and more importantly, precisely at the core of the domain of the unconscionability principle — when it is obvious that the drafters of the contract have intentionally misused the contractual tool to prevent dispute resolution rather than to secure it.

Not surprisingly, some state courts have found the new ban on an old judicial tool frustrating and have demonstrated resistance.[145] In at least two states, namely West Virginia and Oklahoma, the state Supreme Court openly disregarded the decision, invoking a reprimanding response from the Supreme Court that powerfully enforced its new arbitration jurisprudence.[146] Other state courts have persistently tried to find ways to limit the scope of the new jurisprudence and also tried to define at least some areas in which unconscionability — or other tools — can continue to be utilized in order to release parties from unfair contractual waivers of their rights. Important examples of such an approach, within the context of employment, emerged post-*AT&T* in California. In *Sonic Calabasas A, Inc v Moreno*,[147] Justice Goodwin Liu wrote on behalf of the majority that, even after *AT&T* and *American Express*, "the exercise of [unconscionability] as applied to arbitration agreements remains intact, as the *FAA* expressly provides".[148] However, this exercise "remains intact" primarily in theory, while in practice the new arbitration jurisprudence makes it hard to carve out areas not preempted by the *FAA*.

In a creative effort to define such areas, despite the decision in *AT&T*, Liu J drew a line between waivers of class arbitration, which can no longer be invalidated, and waivers of special protections offered to employees

145. Lyra Haas, "The Endless Battleground: California's Continued Opposition to the Supreme Court's Federal Arbitration Act Jurisprudence" (2014) 94:4 Boston University Law Review 1419 at 1421 (describing California's continued resistance to the Supreme Court's expansive interpretation of the *FAA*); James Dawson, "Contract After Concepcion: Some Lessons from the State Courts" (2014) 124:1 Yale Law Journal 233.

146. See Salvatore U Bonaccorso, "State Court Resistance to Federal Arbitration Law" (2015) 67:5 Stanford Law Review 1145 at 1158.

147. 311 P (3d) 184 (Cal Sup Ct 2013).

148. *Ibid* at 201.

under a "Berman hearing" — an informal administrative proceeding designed to help employees bring wage claims. Justice Liu acknowledged that post-*AT&T*, courts can no longer use unconscionability in a manner that will categorically delay or prolong the arbitration process. And yet, he insisted, courts are still allowed to exercise an unconscionability inquiry regarding the total fairness of the scheme of arbitration because, in doing so, they are taking into account the substantive rather than the procedural protections that were surrendered by the employee. In this way, utilizing unconscionability does not necessarily have the effect of slowing down what is supposed to be a speedy process. Such an inquiry, explained Liu J, should focus "on whether the arbitral scheme imposes costs and risks on a wage claimant that make the resolution of the wage dispute inaccessible and unaffordable, and thereby 'effectively blocks every forum for the redress of disputes, including arbitration itself'".[149] As one court recently summarized, Liu J's approach is "that post-*Concepcion*, California courts may continue to enforce unconscionable rules that do not 'interfere with fundamental attributes of arbitration'".[150]

Despite those efforts to save a sliver of the traditional uses of unconscionability much of the original power of this equity-based doctrine seems to be lost. For hundreds of years unconscionability has been used by courts as a broad standard and as an overarching principle. Its power — as well as its vulnerability to criticism — stem from its generality and resilience to manipulation. At least in theory, the principle applies regardless of whatever *form* of contractual harshness was used by the stronger party and courts are always authorized to undo the unfair results achieved by abusing the contractual mechanism and the freedoms it affords. Thus, chopping unconscionability to pieces while disallowing courts the use of most of the pieces — that is when there might be an interference with attributes of arbitration — is insidious to the core of the principle.

149. *Ibid* at 293, citing *Gutierrez v Autowest Inc*, 114 Cal (4th) 77 at 90 (App Ct 2003).

150. *Carlson v Home Team Pest Defense Inc*, 2015 Cal App Ct Lexis 702.

To be sure, the "rise and fall of unconscionability"[151] is a general trend in American jurisprudence not limited to arbitration agreements and in and of itself relates to our neoliberal age.[152] However, it is imperative to realize that paralyzing the unconscionability principle in the context of arbitration looms larger than any other limitation of the principle and hence risks its *general* survival. This is the magnitude of the problem because, by drafting sophisticated arbitration schemes, stronger parties use their freedom of contract to avoid all their other obligations — statutory and contractual alike. Even in a neoliberal age, non-drafting parties such as workers and consumers may have some rights under applicable regulations and/or under their respective contracts. For example, employees may be entitled to a certain combination of wages and benefits. And yet, these rights are meaningless when they are subjected to an arbitration scheme that makes them virtually unenforceable — an effect that is achieved by arbitration terms that block access not only to the public justice system but also to its privatized alternative. Put differently, when courts are subject to the new arbitration jurisprudence and can no longer use unconscionability against unjust arbitration schemes, they are left powerless and helpless. Placed in this weakened position, courts cannot cope with abuses of the freedom of contracts. This, in turn, yields a passive affirmation of predatory behavior and encourages wrongdoers to further exploit their superior status and their unlimited — and ever growing — power.

B. Freedom as Prison

Initially, the Court of the *Lochner* era established a general freedom of contract strong enough to constitutionally defeat the legislative power of the states. About a century later, the Court has engaged in a second wave of *Lochnerism*, framing private dispute resolution as a contract and demanding rigorous enforcement of its terms even if the drafters have used their freedom to design a contract that makes them immune

151. Anne Fleming, "The Rise and Fall of Unconscionability as the 'Law of the Poor'" (2014) 102:5 Georgetown Law Journal 1383.

152. For a discussion see Keren, "Guilt-Free Markets", *supra* note 30.

to legal challenges. This version of freedom is strong enough to defeat the judicial power of the state, thereby rendering salient contractual and statutory rights unenforceable. Taken together, the two rises of the freedom of contract accumulate to a version of freedom that ends up imprisoning masses of people. Freedom turns into imprisonment because of the biased way it has been defined. As it stands today, the freedom of contract encompasses only aspects of freedom that are fully available and meaningful for the most powerful market players: big corporations and their most wealthy individual owners. At the same time, the concept does not include the kind of freedoms that matter most to all other members of society who have to work for a living and struggle with handling their finances. To see this point more clearly, it is important to distinguish between the different components of the freedom of contract and the way these components operate to both enhance legitimate choices and inhibit state interference.

Considering the freedom of contract as a freedom of private choice reveals that the freedom refers to several sets of choices. Organized chronologically, according to the typical progress of the conventional contracting process, there are three main freedoms of choice. First, people have a choice in assuming contractual liability. At the very core of private ordering lies the notion of self-imposed obligations and thus, at least in theory, a party should be subject to a contractual liability only if he or she voluntarily assumed such a liability. Unless and until one consents, one has what theorists call a freedom *from* contract — the right to avoid contractual liability despite being involved in a bargaining process.[153]

Second, once a desire to contract exists, a party should be free to choose with whom to form it, and thus can ignore or reject those he or she does not wish to engage with. I shall call this aspect of the freedom of contract the freedom *to select* a contractual partner. Third, and most commonly associated with the general idea of the freedom of contract, contracting parties have the freedom to choose the content of their contractual obligations. Following others, I shall call this aspect of the

153. Omri Ben-Shahar, "Foreword to Freedom from Contract Symposium" (2004) 2 Wisconsin Law Review 261.

freedom of contract the freedom *to design* the contract.[154]

Combined, the three freedoms — the freedom *from* a contract, the freedom *to select* one's partner, and the freedom *to design* the terms of the deal — create the conventional positive meaning of the "freedom of contract" — a party's freedom "to decide whether, with whom, and under what terms to enter contracts that then become the source of his other-regarding obligations".[155] Most of the time, however, theorists do not engage in an analysis of the idea, let alone in critically reviewing its meaning. This is partially a result of treating the freedom of contract as an axiom, or, under the influence of the *Lochner* era, as a *natural* right or liberty of humans that consequently needs no justification. Most importantly, as rooted in the human will and in the autonomy of individuals, the general idea of the freedom of contract has been presumed universal, held by *all* humans living in Western societies.

The freedoms comprising the freedom of contract have been understood as requiring the imposition of severe limitations on the exercise of state power. Put simply, for the freedom of contract to mean anything the state has to refrain from interfering in the private arena. Respecting individuals' autonomy essentially means enforcing the contracts they have made while using their freedom *to design* their contracts. As we have seen, in the name of the freedom of contract, state acts have been harshly criticized and emphatically invalidated. For more than a century, the idea of the freedom of contract has become synonymous with *laissez-faire* capitalism, anti-regulatory approach, and an outright rejection of centralized efforts to promote welfare. And, as the literal translation of "*laissez faire*" from French suggests, the main idea has been that the state must "leave alone" its citizens to determine their economic affairs.

My argument here is that the kind of freedom of contract that we ended up having post-*Lochner* and post-*Concepcion* celebrates the freedom of choice of the strongest market players at the expense of the many weaker parties who contract with them. Instead of a universal freedom guaranteed to all, as the myth goes, we face an unlimited

154. *Ibid* at 263.

155. West, "Right to Contract", *supra* note 54 at 555.

freedom for some while all others are weakened and enslaved by the idea. A leading feature of the biased nature of the freedom of contract is that the three recognized sub-freedoms discussed above — the freedom *from* contract, the freedom *to select* one's contractual partner and the freedom *to design* the deal — are mostly enjoyed by subjects who belong within the economic elite.

First, only the most powerful market players enjoy a true freedom *from* contract and can choose whether or not — and when — to subject themselves to contract law. Those with less market power seldom have such privileges. This is partially because in an age of growing privatization — after the state had removed itself from so many areas of life — people have to get what they need, including vital necessities such as education and health services, via private contracts. More importantly, oftentimes individuals with limited economic power are not as free to opt out of a contractual regime as their stronger counterparts may be. For instance, typically, a weaker party must get a job that requires signing a contract, followed by a cell phone and a service plan — via a contract — in order to keep in touch with his or her children while on the job. Moreover, since this individual may not own property he or she ought to rent the apartment to which he or she will return at night after a day on the job. It becomes clear that, in this way of life, there is little to no place for the freedom *from* contract.

Second, and related, with limited means comes a way of life that is remarkably different from that of the upper class. This difference often entails an inability to select one's contractual partner. For example, those living in smaller towns or rural areas have to work for the main employer in the area in order to support their families. Others, living in urban neighborhoods, must buy their groceries from the local food store, especially if they are limited to using public transportation due to a lack of a car. Many people around the country have no choice but to send their children to certain schools or colleges and to see only certain doctors primarily because other education or medical opportunities are not affordable. As these examples demonstrate, the freedom to select a contractual partner is not a universal privilege, but a luxury only enjoyed by the wealthiest within society.

Third, and most relevant to the arbitration context, is the freedom *to design* the contract. It goes without saying that with limited bargaining power and without the resources to hire the best lawyers or even receive legal advice, many subjects — like the bakery's employees in *Lochner* or the consumers in *AT&T* — have virtually no such freedom to design nor any other way to influence the content of the contract. As Scalia J pointed out in *AT&T*, in our days most contracts are contracts of adhesion, which are drafted solely by the stronger party, or, more precisely, by the legal department held by the stronger party. The results, described in detail by Margret Radin in her notable book *Boilerplate*,[156] are lengthy and complex contracts that no common person can read or understand — a fact that gives the drafters ultimate control over the terms of the contract. It is now acknowledged and well documented, for example, that in the notorious case of subprime lending much of the damage to vulnerable borrowers has been caused by complex terms that were *deliberately* drafted to be incomprehensible due to the overuse of fine print and confusing legalese.[157] Moreover, the freedom to design the contract includes not only controlling the terms of the deal but also its fundamental structuring. A recent example is the design of the agreements between Uber and its drivers, in which the drafting company structured the relationship as existing outside of the employment realm, thereby depriving the drivers of rights that are only available to employees.[158] In other words, aside from rare occasions in which weaker parties interact with one another, they enjoy no freedom to design their contracts.

What is worse, more often than not the one-sidedly designed contracts are offered on a take-it-or-leave-it basis and thus weaker parties are also lacking the more basic freedom of *negotiating* the contract that was originally drafted by the other party.[159] Many times, the pre-

156. Margaret Jane Radin, *Boilerplate: The Fine Print, Vanishing Rights, and the Rule of Law* (Princeton: Princeton University Press, 2013).
157. Elizabeth Warren, *A Fighting Chance* (New York: Metropolitan Books, 2014) at 139.
158. See *e.g. O'Connor v Uber Techs*, 2015 US Dist Ct Lexis 116482 (ND Cal).
159. Brian H Bix, "Boilerplate, Freedom of Contract, and Democratic Degradation" (2013) 49:2 Tulane Law Review 501.

designed contract is not negotiable because weaker parties so badly and so urgently need the underlying transaction — a job, a loan, a place to stay — that they have no leverage to negotiate anything. At other times, even when alternative transactions are available — several phone carriers, various lenders, or a host of landlords — the parties offering the necessities all knowingly insist on the same terms and rely on the fact that their weaker counterparts have no other choice regarding the terms of the deal. Arbitration agreements are a pertinent example of this type of negotiability problem because most, if not all, of the drafting parties demand similar harsh terms such as the class arbitration waivers discussed earlier. For these reasons, whenever a gap in the bargaining power exists — which is the rule rather than the exception — only stronger parties enjoy the freedom to design their contracts and control their content.

Another considerable aspect of the biased conceptualization of the freedom of contract is that, at the same time that the recognized freedoms are those that matter most to stronger parties, other freedoms — crucially needed by subjects with less economic resources — are either completely unacknowledged or significantly undervalued. Despite the inherent difficulty in describing things that have no name, I would like to suggest that, in general, the missing freedoms reflect areas in which stronger parties typically do not struggle and therefore never had to phrase, frame, and defend as freedoms.

One missing freedom is the freedom *to have* a contract. Notwithstanding the myth of equality, not everybody enjoys this basic freedom.[160] When an owner of a property refuses to lease apartments to people who are, for example, disabled, elderly, or single mothers, the rejected tenants have no "freedom of contract" whatsoever. In fact, they cannot even enter the contractual arena. While the freedom to participate in the market via contracts is axiomatically guaranteed to powerful subjects, it is an uncertain and fragile freedom from the perspective of those with limited means. A current and prominent example is the rapidly growing fringe banking industry. This industry is based on vulnerable

160. Hila Keren, "'We Insist! Freedom Now': Does Contract Doctrine Have Anything Constitutional to Say?" (2005) 11:1 Michigan Journal of Race & Law 133 [Keren, "We Insist"].

borrowers who have to take high-cost predatory loans, usually because big banks *refuse* to contract with them. Notably, such refusal relies on the recognized freedom — discussed above — of the banks to select their contractual partners — a freedom which allows them to reject those with nonexistent or unsatisfactory credit history. Deprived of "prime" lending options, those borrowers are then forced to pay interest rates ranging from 300% per annum to over 1,000% per annum — rates that are sometimes 100 times higher than what is offered in the mainstream credit market.[161] Notably, the freedom to have a contract has never been recognized as part of the canonic freedom of contract precisely because it has always been available for stronger parties and thus was taken for granted and did not require definition or protection.

Another absent freedom is the freedom *from* contract, not in the recognized sense of freedom from having a contract or being subject to contract law, but rather a freedom from exploitation by contract. Borrowers in desperate need of a loan, for example, and many others with limited bargaining powers, need a different kind of freedom *from* contract than their stronger counterparts. They necessitate neither the ability to be left alone (*laissez faire*) nor the power to reject the supervision of the state. This is because these weaker parties are constantly manipulated by stronger market players who are targeting and exploiting them while using their superiority and own unbounded freedom of contract,[162] and thus need support and protection in the form of relief from inappropriate contracts. They have very little freedom of contract when the contractual tool is used *against* them and further impoverishes their limited resources. Importantly, like the freedom to have a contract, this freedom from contractual exploitation has never been recognized as part of the canonic freedom of contract because it is not within the interests of the economic elite. Since powerful market players are unlikely to be

161. Nathalie Martin, "Public Opinion and the Limits of State Law: The Case for a Federal Usury Cap" (2014) 34:2 Northern Illinois University Law Review 259 at 269.

162. Daniel Immergluck, *Foreclosed: High-Risk Lending, Deregulation, and the Undermining of America's Mortgage Market* (Ithaca: Cornell University Press, 2011).

exploited, but rather are more likely to take advantage of less-powerful parties, a meaningful solution to the risk of exploitation has never been developed as a part of the freedom of contract.

Yet another aspect of the biased freedom of contract has to do with the involvement of the state. Despite the myth of a neutral and non-interfering state that stays away from the free market, it must be recognized that the freedom from state intervention is selectively applied. Under the neoliberal rationality, the state is paradoxically "compelled to serve and facilitate an economy it is not supposed to touch, let alone … challenge".[163] One leading way by which the state actively supports the strongest market players is by offering them unyielding enforcement services and presenting these services as necessarily flowing from the idea of the freedom of contract. The Supreme Court's decision in *American Express*, for example, actively "permits and creates an incentive for entities to self-deregulate through private contract".[164] This form of active state support is crucial to the existence of the package of freedoms known as the freedom of contract since, without enforcement, the freedom to enter into a contract and freely design its terms would have meant very little. Furthermore, the biggest businesses and their organizations heavily rely on the state's enforcement services and lobby for the preservation and expansion of these services. For example, many leading corporations and coalitions affiliated with them were involved in the legal battles over the enforceability of class arbitration waivers and invested considerable resources in convincing the Supreme Court to adopt the new arbitration

163. Brown, *supra* note 20 at 40.

164. Glover, "Disappearing Claims", *supra* note 16 at 3091.

jurisprudence[165] and, after their success, in resisting proposals to restore the pre-revolution order.[166] It is therefore false to say that the state takes a neutral and minimal approach, nor is it true that such an approach is what the biggest corporations demand from the state. Rather, in reality what takes place and what is lobbied for is an asymmetric and biased format of state action: high levels of support and involvement on behalf of the few, combined with little to no activity on behalf of all others.

To further explain the last point, it may be useful to take a second look at some of the biases discussed above and to realize that amending each one of them would require a type of state intervention that is

165. See *e.g. AT&T*, *supra* note 13 (*amicus* brief submitted to the Supreme Court by the Pacific Legal Foundation ("PLF"), 2010 US S Ct Briefs Lexis 1043 stating that "PLF was founded more than 35 years ago and is widely recognized as the largest and most experienced nonprofit legal foundation of its kind. PLF litigates matters affecting the public interest at all levels of state and federal courts and represents the views of thousands of supporters nationwide. Among other things, PLF's Free Enterprise Project defends the freedom of contract, including the right of parties to agree by contract to the process for resolving disputes that might arise between them. To that end, PLF has participated as *amicus curiae* in many important cases involving the *Federal Arbitration Act (FAA)* and contractual arbitration in general, including this case at the petition stage" at 1); see also *American Express*, *supra* note 14 (*amicus* brief submitted to the Supreme Court by the Voice of the Defense Bar (DRI) 2012 US S Ct Briefs Lexis 3573 stating that DRI "… is an international organization of more than … [23,000] attorneys involved in the defense of civil litigation". And adding that "DRI has long been a voice in the ongoing effort to make the civil justice system more fair … [and] efficient … " at 1). In addition to special legal coalitions such as PLF and DRI, many big corporations, such as DIRECT TV, Comcast, and Dell, have directly submitted *amicus* briefs supporting rigorous enforcement of individual arbitrations: see *e.g.* 2010 US S Ct Briefs Lexis 1044 (*amicus* brief by the above corporations).

166. See *e.g.* Brian T Fitzpatrick, "The End of Class Actions?" (2015) 57:1 Arizona Law Review 161 at 194-96 (describing the academic support of restoring the pre-Concepcion status quo and the pending bill that followed it, but estimating that "[g]iven the business community's power in Washington, however, no one thinks this bill has much of a chance in the foreseeable future" at 197).

fundamentally different from the current "rigorous enforcement"[167] of contracts. As far as the freedom of design of the contract or at least the opportunity to negotiate the contract's content is concerned, the state would need to regulate the length, language, and negotiability of standard contracts drafted by powerful businesses. Similarly, with regard to the missing freedom to have a contract, to ensure the freedom of those left behind would require the state to compel stronger market players to give up some of their freedom to select their partners and to impose on them a duty to contract with others without discrimination.[168] In the same vein, to establish freedom *from* exploitation by contracts would require the state to limit its enforcement services and refrain from offering enforcement services in cases in which the freedom to design the contract's terms was abused to create a predatory result.

The two rises of the freedom of contract — in the *Lochner* era and under the new arbitration jurisprudence — demonstrate how the neoliberal state moves in the opposite direction, using freedom not to enhance true individual sovereignty and wellbeing for all, but rather to lock people into their inferior status. The combination of two freedom-branded ideas — the "freedom of contract", which is in fact only the freedom of a small but strong group of subjects and the "freedom from the state", which in reality includes extended enforcement services to the same group — turns the resulting "freedom" into a prison for all non-elite subjects. Surely, such an imprisoning notion of freedom severely harms those in the middle and the bottom of the socioeconomic spectrum while concealing the true nature of the process. The magnitude of the resulting harm equals the endless power reserved for contracts in a capitalist society. However, as the coming section suggests, another major harm is underway: as these changes occur in the market and within contract law, their effects spill over and corrode key principles of equity and with them the morality of the law.

---

167. *Dean Witter*, *supra* note 101 at 221.

168. State action via antidiscrimination laws is not awarding people a contractual freedom but only a tort claim. On the advantage of recognizing a contractual freedom in those situations, see Keren, "We Insist", *supra* note 160.

C. Undermining Equity's Rationality

As my close reading of the new arbitration jurisprudence has sought to demonstrate, the dissemination of the neoliberal rationality via the medium of law entails an unwavering rejection of competing rationalities. Non-economic theories based on morality, fairness, justice, or equality — all once at the heart of any jurisprudence — are silenced, marginalized, and delegitimized. In the Anglo-American context, such a rejection of non-economized ways of thinking directly undermines equity's rationality.

To be sure, the legal tradition captured by the term equity is rich, diverse, and much contested among scholars, making it inherently imprecise to talk about equity's rationality. And yet, since much of the power of the neoliberal project stems from effectively marketing simplified ideas (*e.g.* arbitration is efficient), a willingness to generalize and draw the picture with broad strokes is needed in order to expose the full impact of neoliberalism. With this goal in mind, I will put aside specific procedural and doctrinal aspects of equity, despite their importance, and focus on two core ideas that have been associated with equity and are particularly relevant in times of neoliberal economization.

One core idea is the role equity plays in keeping the law in line with conscience, which includes the prevention of immoral abuses of the law's generality and formality. The other is the importance of a broad availability of legal remedies for any system of justice. Each of these interrelated core ideas, I argue, is an essential component of equity's rationality — a rationality that strongly supports limiting the power of the freedom of contract. Together, they create a distinct legal "common sense" that generally disagrees with the one offered by the neoliberal project and particularly conflicts with the celebration of an unleashed contractual freedom. For that reason, the increasing legal support of neoliberalism severely challenges equity's relevancy and legitimacy.

First comes the deep, and almost obvious, connection between equity and conscience. "[E]quity has long been associated with conscience",[169]

169. Dennis R Klinck, "The Unexamined 'Conscience' of Contemporary Canadian Equity" (2001) 46:3 McGill Law Journal 517 at 574.

and there is little doubt that "historically conscience and equity were intimately allied, even synonymous".[170] As declared in the seminal *Earl of Oxford's Case*, equity's defined purpose has been "to correct men's consciences".[171] Alternative explanations and criticisms regarding the ambiguity of the concept notwithstanding,[172] what matters most is *the strong discourse* created by equating equity and conscience, making equity "the official discourse of conscience in the legal sphere".[173] Under such discourse, the law must include moral considerations and moral reasoning, features that have the effect of "taming" the law,[174] while enabling "judges to justify equitable intervention on a moral basis".[175]

How does this conscience-informed discourse relate to judicial decisions regarding the enforcement of contracts? The answer was most famously articulated in 1751 when Hardwicke LC explained that courts of conscience would not enforce agreements that "no man in his senses and not under delusion [could] make on the one hand, and … no honest and fair man would accept on the other".[176] Moreover, he also described those undeserving agreements by directly referring to conscience, naming them "unequitable and *unconscientious* bargains".[177] And, although it was not the first time that courts had refused the enforcement of unfair contracts,[178] it was certainly one of the first times the refusal was rationalized in conscience-oriented terms. Lord Chancellor Hardwicke's words, which emphasize the potential conflict between contractual rights and conscience, have proven to be appealing to generations of judges

170. Klinck, "Nebulous", *supra* note 3 at 211.
171. *Earl of Oxford*, *supra* note 84 at 7.
172. See *e.g.* Samet, *supra* note 26 at 14-17 (reviewing criticisms against anchoring equity in conscience).
173. Klinck, *Conscience*, *supra* note 25 at 11.
174. *Ibid.*
175. Mason, *supra* note 27 at 1.
176. *Earl of Chesterfield*, *supra* note 4 at 155.
177. *Ibid.*
178. For the "ancient roots" of unconscionability, see *e.g.* Friedman, *supra* note 6 at 334-43.

and legal commentators on both sides of the pond.[179] For centuries those words have been cited numerous times,[180] thus spreading a logic born in the "courts of conscience" to all modern courts, regardless of their operation under equity or law.[181] As the reason for unenforceability came to be known as "unconscionability", alluding to Hardwicke LC's "unconscientious bargains", it retained conscience and morality as the logic that explains the refusal to enforce the contract. It can be contended, therefore, without much risk, that equitable limitations on the freedom of contract reflect a moral-based rationality.

More specifically, equity's rationality defines as unconscientious the exploitation of the structure of the law, or the letter of the law, by opportunists. Accordingly, the role of equity is understood as discouraging opportunism, forbidding the abuse of legal rights,[182] and preventing a "behavior that is technically legal but is done with a view of securing unintended benefits from the system".[183] Put another way, and following Aristotle's account of equity, it is the role of equity to constrain the way people *exercise* their legal rights, and for that purpose motives matter. Thus, a person can be a stickler and formalistically insist on fully exhausting his or her legal rights, but he or she is not permitted to be a stickler "in a bad way". Or, in other words, a person is not permitted to exercise his or her legal rights outside of the purposes for which they were created and in a manner that harms others. In that sense, the logic of

179. The first American case to refer to the words of Hardwicke LC in *Earl of Chesterfield*, *supra* note 4 is *Powell*, *supra* note 7. The Supreme Court has adopted the full definition in *Hume*, *supra* note 7.

180. To date the latest case citing the definition in full (including the archaic term "unconscientious bargains" as opposed to only referring to *Earl of Chesterfield*, *supra* note 4) is *Brown*, *supra* note 7 at 678-80, vacated *Marmet*, *supra* note 7.

181. See the decision of the US Supreme court in *Hume*, *supra* note 7, citing Hardwicke LC's definition, and other cases that had followed it, and affirming the lower court's finding ("[t]hese citations are sufficient to show that in suits upon unconscionable agreements the courts of law will take the matter in their own control, and will, without the intervention of courts of equity, protect the parties against their enforcement" at 411).

182. Klimchuk, "Equity", *supra* note 29 at 257-58.

183. Smith, *supra* note 28 at 233.

equity is moral as it fits Kant's view that it is immoral to treat people as mere means rather than as ends in and of themselves.[184]

Applying equity rationality instead of neoliberal rationality to the lengthy and complex contracts drafted by corporations will therefore yield the opposite result than the one achieved under the new arbitration jurisprudence. To be sure, the right to the legal enforcement of contracts according to their terms is a salient aspect of the freedom of contract and, under the *FAA*, this right includes the enforcement of arbitration agreements. However, to seek the enforcement of arbitration agreements that were intentionally designed to deprive the other party of access to any measure of legal help necessarily makes a party a stickler in a bad way and acting opportunistically in exploitation of the general law of enforceability. It is an abuse of the legal tool of contract and the legal rights that are aimed at securing this tool. To be clear, the main problem with the new arbitration jurisprudence is not with affirming the effort of stronger market players to cast the scope of their freedom of contract too widely. Rather, the wrong is in denying that those contractual rights must be exercised with a conscience, in the service of *allowing* rather than *preventing* dispute resolution.[185] While equity's logic would condemn such opportunistic behavior, the neoliberal rationality confirms and incentivizes it.

D. Law without Equity

In the context of the Anglo-American legal world, with its centuries-long tradition of equity, legal adherence to neoliberal rationality effectively means divorcing law from equity principles. The result, which I term "law without equity", presents severe risks to the legitimacy of law and to the future of the social contract holding subjects together. Law without equity allows — and even incentivizes — humans who have

184. Robert Johnson, "Kant's Moral Philosophy" in Edward N Zalta, ed, *The Stanford Encyclopedia of Philosophy* (Stanford: Stanford University, 2004), online: Stanford Encyclopedia of Philosophy Archive <http://plato.stanford.edu/archives/sum2014/entries/kant-moral/>.

185. Compare to the analysis of *Hollywood Silver Fox Farm v Emmett*, [1936] 2 KB 468 (CA (Eng)) in Klimchuk, "Equity", *supra* note 29 at 259.

accumulated enough power to act opportunistically. Without equity's restraining power, those with a combination of economic means, political influence, and intellectual sophistication can and will exploit the legal rules of private ordering to gain even more power and influence.

As the case of class arbitration waivers demonstrates, private entities increasingly use legal rules — such as rules pertaining to the enforceability of contracts — to "delete legal remedies" for others in order "to enhance their economic position".[186] While the formal letter of the *FAA* may tolerate an interpretation that requires an enforcement of *any* arbitration agreement,[187] including one that is effectively designed to prevent arbitration, allowing an opportunistic exploitation of this possibility, and of the formality and generality of the law, is harmful not only to the weaker parties it directly leaves without redress. At a deeper level, a second-order harm occurs: the entire neoliberal legal regime which enforces contracts without equitable limitations increasingly conflicts with key principles of the rule of law.

Notably, in the context of private law in general and contract law in particular, most of the threat to the rule of law comes from powerful market actors and from the failure of the law to limit them. This, however, should not stop us from recognizing the big-picture problem. Although the rule of law has traditionally been a public law doctrine, there is a growing recognition of the importance of linking it to the way private law treats the exercise of private powers, especially in times of

186. Margaret Jane Radin, "Boilerplate: A Threat to the Rule of Law?" in Austin & Klimchuk, *Private Law*, *supra* note 28 at 297 [Radin, "A Threat"].

187. But note that many have argued that this was never the intention of the legislator. See *e.g.* Hiro N Aragaki, "The Federal Arbitration Act as Procedural Reform" (2014) 89:6 New York University Law Review 1939 (contesting the contractual model of arbitration and arguing that "the FAA was understood by merchants, judges, and lawyers as a vehicle for improving the procedure by which commercial disputes were adjudicated fair and square" at 1943).

increasing privatization.[188] In other words, the rule of law pertains not only to the relationship between individuals and the state, but also to the interactions between individuals. On this view, then, when legal actors decide questions of private law they heavily influence the rule of law.[189] Or, as Radin has compellingly argued in the context of contracts:

> [f]irms that use contract to destroy the underlying basis of contract, that deploy contract against itself, are using contract to destroy the ideal of contractual ordering, which the rule of law is formulated to protect. In this way, they not only undermine the idea of a regime of private ordering, *they effectively undermine the rule of law*.[190]

1. Arbitrariness

A core element of the rule of law is that "a right to exercise power arbitrarily cannot be conferred or upheld by law".[191] And yet, the new arbitration jurisprudence does exactly that. By offering unconstrained enforceability services, this jurisprudence awards the strongest market players an ability to exercise their own power over those who need to contract with them. Accordingly, the new arbitration jurisprudence facilitates the subjection of large portions of the public — workers, consumers, clients and the like — "to the arbitrary will of others".[192]

The will of powerful market players is arbitrary not because it is random or capricious; it is in fact quite systematic and deliberate in its uncompromising demand for waivers of rights and remedies. Rather, it is arbitrary in the sense that the power conferred by law — to design and enforce arbitration contracts — is being abused. The drafters of contracts use their legal power not as a way to shape a private mechanism of dispute resolution as authorized by the *FAA*. Quite to the contrary, these drafters deploy their power(s) arbitrarily to grant themselves immunity from legal

188. Lisa M Austin and Dennis Klimchuk, "Introduction" in Austin & Klimchuk, *Private Law, supra* note 28 at 6-14 [Austin & Klimchuck, "Introduction"].
189. William Lucy, "The Rule of Law and Private Law, in Private Law and the Rule of Law" in Austin & Klimchuk, *Private Law, supra* note 28 at 41.
190. Radin, "A Threat", *supra* note 186 at 301 [emphasis added].
191. Austin & Klimchuk, "Introduction", *supra* note 188 at 1.
192. Klimchuk, "Equity", *supra* note 29 at 249.

claims. Such an exercise of power "undermines the legal infrastructure of private ordering, and becomes a scheme of arbitrary power".[193]

2. Inequality

Another core principle of the rule of law is equality before the law. Under this principle, Dicey famously explained, "every man, whatever be his rank or condition, is subject to the ordinary law of the realm and amenable to the jurisdiction of the ordinary tribunals".[194] However, the neoliberal celebration of the freedom of contract — in a biased form — allows one group of powerful market actors to use their freedom in a manner that positions their group *above* the law, exempt from the "jurisdiction of the ordinary tribunals".[195] At the same time, by enforcing the contracts drafted by this preferred group, the new arbitration jurisprudence deprives most other members of the public from having access to a working mechanism of dispute resolution, thereby marking their inferiority before the law. Put differently, when — with approval from the law — some can put themselves above the law while others remain with no "reasonable access to remedies",[196] the principle of equality before the law is severely corroded.

3. Risking the Social Contract

Where does law without equity — which conflicts with core elements of the rule of law — leave us as a society? Not in a very hopeful place. When the freedom of contract has a biased definition and is subject to almost no limitation even in circumstances when it is clearly abused, the legitimacy of the law is at risk and the trustworthiness of courts is jeopardized. Recall that Pound had warned us about this risk back at the midst of the *Lochner* era, reporting that the then-new "liberty of contract" approach had generated "a growing distrust of the integrity of

193. Radin, "A Threat", *supra* note 186 at 300.
194. Austin & Klimchuk, "Introduction", *supra* note 188 at 4-5 citing AV Dicey, *Introduction to the Study of the Law of the Constitution,* 10d (London: Macmillan, 1959) at 193.
195. *Ibid.*
196. Radin, "A Threat", *supra* note 186 at 290.

the courts".[197] About a century later, a Montana judge responded to the pro-arbitration jurisprudence of his time with similar suspicions, albeit less carefully expressed, writing: "[i]t seems to me that judges who have let their concern for their own crowded docket overcome their concern for the rights they are entrusted with should step aside and let someone else assume their burdens".[198]

And by risking the legitimacy of the law, the very fabric of our social contract is at risk. The belief that we all should sacrifice *some* of our individual freedoms for the sake of maintaining a safe and just civil society is severely undermined when the law so clearly works on behalf of a certain group affiliated with the economic elite. Instead of subjecting themselves to the rule of law — and in return, to being protected by the law — mass parts of the public are controlled by the unbridled interests of the elite as those elite interests are enforced by law. It does not help, of course, that in addition to the emergence of the new arbitration jurisprudence, other recent decisions of the American Supreme Court conferred upon corporations key political and constitutional rights traditionally reserved to individuals.[199] The problem is that the more the law aligns itself with neoliberal rationality and the interests of corporations, the more it becomes part of the general "shift from the social contract to savage forms of corporate sovereignty".[200]

Indeed, our current situation may be even worse since the said "shift" sometimes seems to amount, as suggested by Wendy Brown, to a complete "*inversion*" of the social contract.[201] Instead of being equally limited and protected by the state, subjects are divided into winners and losers, and losers are treated as means rather than ends. As proved by the new arbitration jurisprudence, the end is no longer the wellbeing of individuals, but rather economic efficiency, which is defined from the

197. Pound, *supra* note 32 at 487.
198. *Casarotto v Lombardi*, 268 Mont 369 at 386 (Sup Ct 1994).
199. *Citizens United v FEC*, 558 US 310 (2010) [*Citizens United*]; *Burwell v Hobby Lobby Stores, Inc*, 134 S Ct 2751 (US 2014).
200. Henry A Giroux, *Neoliberalism's War on Higher Education* (Chicago: Haymarket Books, 2014) at 54.
201. Brown, *supra* note 20 at 38, 64, 110, 134, 213.

perspective of the winners. To achieve the desired efficiency, those at the bottom are knowingly and deliberately left unprotected. And, rather than offering them safety via the law, the state demands that they will either protect themselves by not signing exploitative contracts or sacrifice themselves on behalf of the preservation of an efficient economy.

V. Conclusion

In his highly influential book *Commentaries on Equity Jurisprudence as Administrated in England and America,* Judge Story wrote:

> [t]here may be such an unconscionableness or inadequacy in a bargain, as to demonstrate some gross imposition or some undue influence; and in such cases Courts of Equity *ought to interfere* ... such unconscionableness, or such inadequacy should be made out, as would (to use an expressive phrase) *shock the conscience*.[202]

Although written in 1836, Story J's words have gained much eminence throughout the years and have been quoted, with or without adequate reference to him, by numerous judges and scholars.[203] Notably, Story J's message has two separate prongs, both still relevant today. One relates to people's behavior and highlights that an unlimited freedom of contract may create bargains that are so unfair and immoral that they "shock the conscience". The other relates to the role of the judiciary and insists that when asked to enforce such unconscionable contracts, courts "ought to interfere".

Comparatively, outside of the American legal world, recent works have described and analyzed a rising agreement with Story J's influential ideas. For example, Irit Samet has argued that "[o]ne of the most interesting and controversial developments in the recent jurisprudence on

202. Joseph Story, *Commentaries on Equity Jurisprudence as Administered in England and America* (Boston: Little, Brown and Co, 1866) at 268 [emphasis added].

203. Since Story J's words were quoted in full in the known case of *Eyre v Potter*, 56 US 42 (1853), some have attributed them to the US Supreme Court: see *e.g.* Friedman, *supra* note 6 ("[t]he 'unconscionableness or inadequacy' must be such as would 'shock the conscience' – an 'expressive phrase' that retains a hold on the current unconscionability doctrine" at 339, and *ibid* at 60).

equity is the increasing use of conscience categories to discourage overly selfish behaviour among parties to commercial relationships".[204] In the United States, however, the Supreme Court has recently developed — as this article has sought to demonstrate — an approach to agreements that directly conflicts with each of the prongs of Story J's theory.

First, the Court has adopted an unfettered version of the freedom of contract that allows and encourages drafters of contracts to behave selfishly, while classifying conscience-related problems inapplicable, or "unrelated". And second, the Court has strongly disagreed with the notion that courts "ought to interfere" in cases of unconscionability. To the contrary, according to the new arbitration jurisprudence courts "*must*" do the opposite: they "must rigorously enforce arbitration agreements".[205] Accordingly, courts are no longer allowed to "interfere" by using equitable tools such as the unconscionability doctrine.

Again, it is urgent to recognize that this new approach, which this article refers to as neoliberal-*Lochnerism*, creates a second rise of the freedom of contract that operates not only in the domain of arbitration. Rather, *all* the rights of weaker parties — whether produced by regulations or by contractual terms — are rendered meaningless when stronger parties are permitted to use their freedom in a manner that obstructs access to remedies. The genus of arbitration agreements now allowed by the US Supreme Court thus represents a larger assault on fairness, morality, and justice than the eye can see at first glance.

The result, I have argued, is "law without equity" and it presents severe risks to the legitimacy of law and to the future of the social contract that necessarily holds subjects together. Law without equity allows — and even incentivizes — humans who have accumulated enough power to act opportunistically. Without equity's restraining power, those possessing a combination of economic means, political influence, and intellectual sophistication can and will exploit the legal rules of private ordering in order to gain even more power and influence.

But is there any potential to counter the rising and spreading neoliberal

204. Samet, *supra* note 26 at 13.

205. *American Express*, *supra* note 14 at 2309.

rationality? The task is extremely difficult, especially given the growing powers of corporations not only in the market, but also in the political arena.[206] Indeed, despite strong evidence and a dramatic study produced by the Consumer Financial Protection Bureau,[207] there is currently little to no hope that efforts to overcome the new arbitration jurisprudence by revising the *FAA* will be successful.[208] However, as any other quest for hope,[209] the ability to imagine things differently is key. Restoring old images of equity seems to be a good starting point for envisioning a better justice system. Courts must be allowed to see themselves — and act as — "courts of conscience"; they cannot be ordered to participate in exploitation conspiracies. Being moral actors themselves, judges simply "have no business coming to the aid of immoral business practices".[210] And, despite the argument that no one knows precisely what "conscience" meant centuries ago,[211] or how its meaning is supposed to be applied in our days,[212] there must be a way to know what conscience *is not*.

Judges ordered by the Supreme Court to hold against their conscience and to approve manipulations of the law disguised as exercises of the freedom of contract are not only at risk of personal frustration. Given the expressive power of the law judges serve a social role of moral leaders and thus their role in limiting freedom by conscience is of utmost

206. Brown, *supra* note 20 at 152-73 (offering a critical analysis of *Citizens United*, *supra* note 199).

207. US, Consumer Financial Protection Bureau, *Arbitration Study: Report to Congress, Pursuant to Dodd-Frank Wall Street Reform and Consumer Protection Act § 1028(a)*, (2015) online: <http://files.consumerfinance.gov/f/201503_cfpb_arbitration-study-report-to-congress-2015.pdf>.

208. Glover, "Disappearing Claims", *supra* note 16 (explaining why "amending the FAA is an extremely difficult political task" at 3086).

209. Kathryn Abrams & Hila Keren, "Law in the Cultivation of Hope" (2007) 95:2 California Law Review 319.

210. Robin L West, "The Anti-Empathic Turn", in James E Fleming, ed, *Passions & Emotions* (New York: NYU Press, 2013) 243 at 265.

211. See *e.g.* Richard Hedlund, "The Theological Foundations of Equity's Conscience" (2015) 4:1 Oxford Journal of Law & Religion 119.

212. See *e.g.* Rohan Havelock, "Conscience and Unconscionability in Modern Equity" (2015) 9:1 Journal of Equity 1.

importance.[213] For courts to effectively fulfill this leadership role we ought to conceive conscience as "a metaphor for the dynamic interaction between changing social norms and shifting individual beliefs",[214] and not as an arbitrary measure comparable to the Chancellor's Food. On this view, courts are a crucial mediator between society and the self. Armed with equity principles, and especially the unconscionability doctrine, they have the power — and the duty — to restrain the freedom of contract, restore contractual justice, and protect the integrity of the law.

213. Keren, "Guilt-Free Markets", *supra* note 30.
214. *Ibid.*

Canadian Journal of Comparative and Contemporary Law

Volume 2 | No 1 | 2016

Equity in the 21st Century: Problems and Perspectives