

Chicago-Kent College of Law

From the Selected Works of Richard W. Wright

December, 1999

Principled Adjudication: Tort Law and Beyond

Richard W. Wright, *Chicago-Kent College of Law*

PRINCIPLED ADJUDICATION: TORT LAW AND BEYOND

RICHARD W. WRIGHT*

I. INTRODUCTION

Are there moral principles underlying tort law (or law in general)? If so, what are they, and is there a need for judges or others involved in identifying, interpreting and applying the law, or for legislators dealing with issues related to the area of law, to understand and take proper account of those moral principles? If there are underlying moral principles that should be taken into account, when and how should they be taken into account?

These issues are the subject of two recent articles by Professor John Smillie, who addresses them in the broad context of civil adjudication in New Zealand¹ and the specific context of the availability of exemplary damages for personal injury.² Professor Smillie identifies and distinguishes three approaches to adjudication that have been used or advocated by New Zealand judges: formalism, fairness, and efficiency. While noting familiar problems with the supposedly traditional formalist approach, which relies on literal interpretation and deductive application of explicitly stated rules, Smillie ends up strongly advocating it over either of the morality-based approaches: the fairness approach espoused by Sir (now Lord) Robin Cooke and other influential New Zealand judges and the utilitarian efficiency approach advocated by Sir Ivor Richardson. Smillie asserts that there is and can be no consensus on what constitutes fairness, so that the fairness approach results in judges' employing their own personal value preferences and thereby undermines the very idea of the rule of law. Although he finds the usual versions of the utilitarian efficiency approach to be normatively and descriptively implausible, he asserts that a limited, mixed version of that approach furthers the same basic value as the formalist approach but in a less simple and straight-forward, and hence unsatisfactory, manner. He claims that the formalist approach promotes equal individual freedom or autonomy, which he identifies as the fundamental value of the common law, classical liberal philosophy, and modern libertarian philosophy.³

Professor Smillie's discussion is erudite and thought-provoking. He addresses issues that are of great theoretical and practical importance. I agree with much of what he says, including his identification of equal individual freedom or autonomy as the fundamental value that underlies the common law and classical liberal philosophy. However, I disagree with a number of his premises and conclusions. As will be developed below, I believe he understates the problems and indeed the impossibility of the formalist approach, misconstrues the concept of equal freedom or autonomy as narrow self-interested libertarianism, fails to appreciate the equal freedom

* Professor of Law, Chicago-Kent College of Law, Illinois Institute of Technology. Copyright Richard W Wright. All rights reserved. Permission is hereby granted to copy for non-commercial purposes as long as appropriate citation is made to this publication.

1 John Smillie, "Formalism, Fairness and Efficiency: Civil Adjudication in New Zealand" [1996] NZLR 254.

2 John Smillie, "Exemplary Damages for Personal Injury" [1997] NZLR 140.

3 Smillie, *op cit* n 1, at 254-255, 257-258, 260-261, 268-269, 272-273.

norm's status as the foundational norm of the fairness or justice theory and its rejection by the utilitarian efficiency theory, and overlooks the possibility and indeed necessity of a principled approach to adjudication that draws on the detailed principles of justice that underlie and give content to the law in general and tort law in particular.

II. THE IMPOSSIBILITY OF FORMALISM

As Smillie notes, the formalist approach to adjudication is based on a strict positivist conception of law, according to which the law consists solely of the rules explicitly stated in legislation or leading judicial precedents.⁴ The more sophisticated legal positivists (including Smillie) acknowledge that moral issues often do and should affect the creation of law, that certain minimal conditions of the rule of law (e.g., generality, publicity, consistency, and other elements of what Lon Fuller called the 'internal morality' of law⁵) must be satisfied before a legal system can be said to exist, and even that a legal system will not come into being or survive unless it has a minimum substantive natural-right content (protections of life and property), at least for those in power.⁶ However, the legal positivists insist on a fundamental distinction between the (explicitly stated) formal law and any underlying or extrinsic principles or policies.⁷

More importantly for our purposes, the strict legal positivists insist that the law, once created, can and should be interpreted literally and applied deductively in the great majority of cases, without any explicit or implicit resort to morality or any other underlying or extrinsic purpose. This formalist approach to adjudication, which is defended by Smillie, was once defended by HLA Hart, who, although he rejected the formalist label, was the leading modern proponent of positivist formalism. He asserted that legal rules have a 'core' of plain meaning that can be deductively applied to decide the great majority of cases.⁸

Like Smillie, Hart recognized that in some cases the arguably applicable rules may conflict or be ambiguous or silent. He described these cases as being in the 'penumbra' rather than in the core of plain meaning of the system of rules. In these cases, judges must turn to underlying or extrinsic purposes to reach a decision. However, Hart noted, these purposes need not be moral principles. Moreover, these purposes are not themselves part of the law, which is exhausted by the preexisting rules' core of plain meaning. Rather, the judges are acting as legislators by creating new law, and are (retroactively) applying that new law to the existing case, rather than applying preexisting law. As legislators acting in the penumbra, judges should reach whatever result seems best in the particular case, free of any constraint imposed by the preexisting law (which, by assumption, is exhausted by the

4 Smillie says 'largely' rather than 'solely': op cit n 1, at 255. But for the strict legal positivist the law is exhausted by the 'core' of the explicitly stated law. See below, n 9.

5 Lon L Fuller, *The Morality of Law* (Yale Univ Press, New Haven, rev'd ed, 1969), pp 33-94.

6 For example, HLA Hart, *The Concept of Law* (Clarendon Press, Oxford, 1961), pp 167-168, 189-207.

7 For example, HLA Hart, "Positivism and the Separation of Law and Morals" (1958) 71 Harv LR 593, at 614-615.

8 Hart, *The Concept of Law* at 119-120, 122-123, 131-32, 150; Hart, "Positivism" at 607, 610, 615; see Smillie, op cit n 1, at 255-257, 273.

formal core).⁹

The legal positivists' treatment of penumbral cases as instances of retroactive, discretionary, open-ended judicial legislation raises several serious problems for the rule of law – the very same problems that Smillie attributes to the fairness approach.¹⁰ First, there is the problem of the judiciary's institutional role, legitimacy, and competence. Judges are not supposed to create law, but rather to interpret and apply law that is created by a representative, democratically elected legislature. Second, given the assumption that there is no preexisting law to apply in the penumbra, adjudication in the penumbra is governed by what the particular judge thinks would be the fair or otherwise best result, rather than by some general rule of law, thereby undermining consistency, predictability, and the very idea of the rule of law rather than the rule of a particular man or woman. Third, the judgment in each case necessarily is an instance of retroactive creation and application of 'law' to past conduct. Beforehand, at the time the relevant conduct occurred, there was no law to guide or govern that conduct.

Both Hart and Smillie are acutely aware of these problems with the penumbra. Hart's notion of the formal core, which Smillie adopts, is aimed at rescuing the rule of law (in the great bulk of cases) from the problems of the penumbra. The rescue effort succeeds only if there exists a set of identifiable, complete, consistent, unambiguous rules that can be literally interpreted and deductively applied, without conflicts or gaps, to (the great bulk of) particular fact situations. None of these conditions are satisfied.

As the American legal realists delighted in pointing out, it is usually difficult, if not impossible, to identify a coherent, comprehensive and determinative rule in the text of any judicial opinion, or even to derive such a rule from an opinion or set of opinions. This is hardly surprising when one considers that the task and authority of the court extends only to the resolution of the particular dispute before it. Indeed, insofar as the court expounds on issues and situations that are not relevant to the resolution of the particular dispute before it, what it states is not a part of the authoritative holding of the case and thus has no binding precedential force, but rather is mere dicta – a statement which is not authoritative or binding, although it might ultimately be deemed persuasive, in a subsequent case. For any one opinion or set of opinions, a multitude of broad or narrow rules or holdings with varying consequences can be formulated. Moreover, in attempting to derive rules from or apply rules to cases, one must be able to identify the relevant similarities and differences among various cases, which can only be determined in light of the underlying purposes.¹¹

The task of identifying or deriving a set of comprehensive, coherent, unambiguous, determinative rules seems especially problematic for areas of law, like tort law, which traditionally have been developed through court

9 See Hart, *The Concept of Law* at 123-125, 128-129, 131, 200-201; Hart, "Positivism" at 607-615, 627-629; Smillie, *op cit* n 1, at 255. Hart at one point indicated that judges operating in the penumbra should reason by analogy to the core cases and further the purposes that underlie the relevant rules. Hart, *Concept of Law* at 124. Yet, if (as Hart asserted) these purposes are not themselves part of the law, why should judges be constrained to further them rather than other, arguably better purposes in the particular case?

10 Smillie, *op cit* n 1, at 255, 260-261, 266.

11 See KN Llewellyn, *The Bramble Bush* (Oceana, New York, 2nd ed, 1951), pp 38-39, 42-49, 64-69; Steven J Burton, *An Introduction to Law and Legal Reasoning* (Little, Brown & Co., Boston, 2nd ed, 1995), pp 31-41; Hart, *The Concept of Law* at 122, 131, 155.

decisions rather than through statutory enactments. However, the situation is not much better with statutes, which, even when systematically reorganized into codes, are usually disjointed and unclear and always incomplete. This sometimes results from legislative compromises, stalemates, or hasty or otherwise deficient drafting. It is in any event inevitable given the imprecise nature of human language and the impossibility of drafting complete and unambiguous language to cover every possible situation. Moreover, statutes are not self-executing, but rather must be applied to specific situations. In the end, a statute means what the court interpreting and applying it says it means. We are confronted again with the problem of identifying and deriving rules from judicial decisions in discrete cases with limited issues and facts.¹²

Moreover, as Lon Fuller argued, even in the admittedly easy cases, in which the rule clearly applies, there always is an implicit reference to the underlying purpose(s) of the particular rule. For example, Hart assumes that a rule stating 'no vehicles in the park' would always apply to the 'standard' or 'core' instance of an automobile.¹³ But what about an automobile rendered immobile as part of a monument or a play area, or a mobile refuse-collection truck or emergency vehicle? Although a judge could interpret the rule literally as applying to such 'core' instances,¹⁴ why should she? Fuller agrees that most cases are easy (e.g., a private citizen driving through the park in a non-emergency situation), but he insists that they are easy not because of some literal core of plain meaning but rather because of the implicit underlying purposes. Conversely, he states, a hard case should be decided in the light of the underlying purposes of the relevant rules or area of law, rather than being treated as an open-ended decision of what the particular judge thinks would be the fairest or best result in the specific situation.¹⁵

Hart eventually accepted much of Fuller's critique of strict legal positivism, including Fuller's criticism of the formal core. He acknowledged that legal interpretation always requires explicit or implicit resort to the purposes underlying the relevant rules.¹⁶ In addition, he stated that it had never been his view that a judge is free to implement what she subjectively perceives to be the best result in the (still supposedly lawless) penumbra:

[A]mong the features which distinguish the judicial from legislative law-making is the importance characteristically attached by courts, when deciding cases left unregulated by the existing law, to proceeding by analogy so as to ensure that the new law they make is in accordance with principles or under-pinning reasons which can be recognized as already having a foothold in existing law. Very often in deciding such cases courts cite some general principle or general aim or purpose which a considerable area of the existing law can be understood as exemplifying or advancing, and which points towards a determinate answer for the instant case.¹⁷

Thus, legal formalism is indeed a pretence. It is also unwise. Since the process of deriving and applying rules always requires explicit or implicit

¹² See Hart, *The Concept of Law* at 122-126.

¹³ *Ibid.* at 123, 125.

¹⁴ See Frederick Schauer, 'Formalism' (1988) 97 Yale LJ 509, 510, 524-528, 532-533, 540-541, 545-547.

¹⁵ Lon L. Fuller, 'Positivism and Fidelity to Law - A Reply to Professor Hart' (1958) 71 Harv LR 630, 661-669.

¹⁶ HLA Hart, 'Introduction' *Essays in Jurisprudence and Philosophy* (OUP, Oxford, 1983), p 7-8.

¹⁷ *Ibid.* at 6-7.

resort to the underlying purposes, good judicial reasoning should explicitly consider these purposes in the hard cases. As Hart himself pointed out, failure to do so merely camouflages conservative or wooden judicial thinking that results in decisions which, although clearly wrong in light of the relevant principles or policies, nevertheless are announced and justified as supposedly being required by adherence to some rule or precedent, even when the conditions that gave rise to that rule or precedent do not apply in the current situation.¹⁸ This, indeed, seems to have occurred in the case that Smillie uses to illustrate the formalist approach, *Ross v McCarthy*,¹⁹ the reasoning of which even Smillie describes as perhaps being ‘excessively cautious and unduly deferential.’²⁰ If the courts had always followed Smillie’s preferred conservative, hands-off, formalist approach, especially in cases like *Ross v McCarthy*, the common law would never have evolved in the first place.

III. THE NECESSITY AND PLAUSIBILITY OF PRINCIPLISM

If there is no formal core, so that all cases are penumbral cases requiring purposive interpretation, we may seem to be driven to the disastrous results for the rule of law that were listed above when discussing the legal positivists’ conception of judicial reasoning in the penumbra. If there are no rules, and it is left completely up to the subjective opinion of the particular judge as to what is the fairest or otherwise best result in each case, we do not have the rule of law, but rather the unpredictable, inconsistent, retroactively applied rule of particular men and women. This, of course, was the articulated view of the extreme legal realists²¹ and is the articulated view of their current successors: the legal pragmatists, critical legal theorists, and postmodernists.²² It is from this legal nihilist hell that Hart and Smillie, understandably and rightly, are intent on delivering us. What alternatives other than (the pretence of) formalism are available?

18 Hart, *The Concept of Law* at 126-127; Hart, “Positivism” at 610-611.

19 [1970] NZLR 449.

20 See Smillie, op cit n 1, at 255-257. Smillie refers to the ‘rule’ subsequently adopted by Parliament, which however seems merely to have listed factors to be considered. The *Ross* court easily could have announced a standard negligence rule, ‘There is a duty to take reasonable care to keep one’s stock off the highway,’ which presumably would have led to consideration of the same factors, whether or not they were explicitly mentioned by the court.

21 For example, Jerome Frank, *Law and the Modern Mind* (1930); Joseph C Hutcheson, Jr., “The Judgment Intuitive: the Function of the ‘Hunch’ in Judicial Decision” (1928) 14 Cornell LQ 279. But see Jerome Frank, *Law and the Modern Mind* (6th ed, 1949) (shifting from rule scepticism to fact scepticism). For powerful attacks on the extreme legal realist position, see Hart, *The Concept of Law* at 133-143; Lon L Fuller, *The Law in Quest of Itself* (Beacon, Boston, 1940) p 52-68.

22 The legal pragmatists, like the other legal nihilists, deny the existence of general rules or underlying foundational principles. However, unlike the other legal nihilists, they assert that law somehow still exists and is normatively grounded in ad hoc, context specific determinations of ‘what works best,’ and many of them still claim allegiance to the values of justice and right. However, they fail to explain how anyone can determine ‘what works best’ in the absence of any explicit or implicit foundational moral principles. What is the definition or test of ‘working’? All the legal nihilists seem to be, as Ronald Dworkin states, merely formal or “external” sceptics who however “play the game internally” by appealing to certain moral values and relying on principled arguments, which almost always are based on the foundational value of classical liberalism: equal individual freedom. See Ronald A Dworkin, “‘Natural’ Law Revisited” (1982) 34 U Fla LR 165, at 175-177; Mark V Tushnet, “The Left Critique of Normativity: A Comment” (1992) 90 Mich LR 2325.

Smillie identifies two alternative approaches employed or advocated by prominent New Zealand judges: judgments explicitly based either on fairness or efficiency.²³ He assumes that each of these approaches consists of a direct appeal to the respective preferred value – fairness or utilitarian efficiency – unmediated by any explicit or implicit anchoring in or reference to the positive law expressed in statutes or judicial opinions. He further assumes that there is no social (or at least judicial) consensus, embodied in the law or otherwise accessible, with respect to the importance, content, or implications of these values. If these assumptions are correct, then, as Smillie notes, all the horrors of the positivist conception of the penumbra, as described above, follow.²⁴

I am not familiar with the relevant New Zealand cases or literature. However, Smillie's discussion of the fairness approach does not persuade me that either of his assumptions applies to that approach. He notes that New Zealand judges employing the fairness approach refer to 'shared social values and expectations' concerning fair or just results that are not 'necessarily reflective of majority public opinion at any particular time' but rather 'take a long-term view reflecting enduring values' that are 'rooted in the community.'²⁵ Although he observes that ordinarily one might expect to find such enduring values 'deeply rooted in existing laws and institutions,' as in England, he asserts that this possibility is foreclosed in New Zealand by the judiciary's progressive and creative rather than conservative and formalist approach to adjudication, which he believes has led to a radical departure from the values previously embedded in the common law. However, this assertion seems to be based on Smillie's (erroneous) libertarian interpretation of the foundational norm of equal individual freedom and its elaboration in classical liberal philosophy and the common law.²⁶

As I will discuss in the next section, the principle of corrective justice, which is based on the equal freedom norm, focuses on the very goal that Smillie finds objectionable in Lord Cooke's fairness approach: 'protect[ing] citizens' 'reasonable expectations' of security for their interests' and 'upholding minimum standards of [nonharmful] moral behaviour in their dealings with others.'²⁷ If your conduct creates a risk of harm to the persons or property interests of others, you are subject to the corrective-justice requirement that your exercise of your freedom be consistent with those others' equal external freedom. As Lord Cooke has explained, this principle is reflected (imperfectly) in the golden rule and *Donoghue v. Stevenson*'s neighbour principle.²⁸ Contrary to Smillie's assertions, neither classical liberal philosophy nor the common law 'recognize contractual consent as the [sole or] primary source of civil obligation.'²⁹ In this statement, Smillie overlooks not only the common law of torts, trusts and restitution, which

23 Although it is not clear, Smillie seems to treat fairness and efficiency as 'competing conceptions of justice.' See Smillie, op cit n 1, at 258-259, 268. While the concepts of fairness and justice are strongly related, the concept of efficiency is in direct conflict with the concept of justice. See below, sections IV and V.

24 Smillie, op cit n 1, at 258-261, 266.

25 Ibid at 260-261.

26 Ibid at 254, 258-259, 272-273.

27 Ibid at 261.

28 Ibid. See below, n 54.

29 See Smillie, op cit n 1, at 262, 272-273.

are and always have been at least as important as contract law in creating and enforcing corrective-justice obligations, but also the common law of property, which is heavily influenced by distributive-justice as well as corrective-justice principles.³⁰ Yet, although the principles of justice may go further than a minimalist libertarian would like, they do not support the 'strongly paternalistic, communitarian sense' of fairness that Smillie infers from Lord Cooke's statements.³¹

Smillie also notes Lord Cooke's general presumption against taking 'economic [efficiency] considerations such as comparative cost, insurance and loss-spreading' into account.³² But if the fundamental value underlying the law is fairness or justice, it is quite correct to subordinate or even ignore such economic efficiency considerations (but not distributive justice considerations), which are in direct conflict with the concept of justice.³³

Finally, contrary to Smillie's assertion, I see no departure from formal justice but rather truisms in Lord Cooke's statements that '[c]oncrete decisions are in the end applications of very broad criteria . . . to a particular set of facts' and that 'facts vary infinitely,' so that (as Smillie concludes) 'decided cases have little *binding* force as precedents.'³⁴ No case is exactly like any previous case. As Hart emphasized, the formal justice requirement of treating like cases alike can only be achieved by comparing cases in terms of relevant similarities and differences, which in turn can only be identified by reference to the relevant purposes. To instead focus on and 'freeze' certain features into a rule that is to be applied through blind formalism, for the sole purpose of attaining certainty and predictability of decision, is to abandon formal as well as substantive justice, rather than to implement it.³⁵

In sum, apart from one quoted statement by Justice Thomas, which asserts the right of judges 'to express that opinion which best serves their intelligence and wisdom and best discharges their personal scruples and conscience,'³⁶ Smillie's discussion does not demonstrate to me that New Zealand judges applying the fairness approach are deciding cases based on their own personal values or conceptions of justice rather than on a shared concept of justice that has been and continues to be deeply rooted in New Zealand law.³⁷

If New Zealand judges who employ the fairness approach do so by resorting (explicitly or implicitly) to a shared concept of justice that is deeply rooted in New Zealand law, they are employing the only approach that can rescue us from the otherwise pervasive perils of the positivist penumbra. We will use the term 'principlism' to refer to this approach, since it is based

30 See Richard W Wright, "Substantive Corrective Justice" (1992) 77 Iowa LR 625, at 709-710.

31 See Smillie, op cit n 1, at 260-261.

32 Ibid at 260.

33 See below, sections IV and V.

34 See Smillie, op cit n 1, at 260 (emphasis added).

35 See Hart, *The Concept of Law* at 126-127, 155.

36 See Smillie, op cit n 1, at 261.

37 However, I share Smillie's concern that New Zealand has followed Australia in lumping actual-causation and proximate-causation issues together as an undifferentiated and unelaborated issue of 'common sense,' rather than distinguishing and clarifying each element. See ibid at 264; *March v E & MH Stramare Pty Ltd.* [1991] 171 CLR 506; Richard W Wright, "Causation in Tort Law" (1985) 73 Calif LR 1735.

on the proposition that there is a coherent set of principles underlying the positive expressions or sources of law, which are considered to be part of and are used to interpret and apply the real or actual law.

Is principlism plausible? There is strong reason to believe that it is. Any attorney in actual practice knows, and any successful law student learns, that, although the law is not rule-determined, it is rule-guided. Incomplete and broadly phrased rules generally provide the basic (and sometimes the detailed) framework for every area of law, which however must be understood and filled out by reference to the underlying purposes. As Hart claimed and Fuller acknowledged, the great bulk of situations that arise in daily life are easy 'standard cases.' In the 'standard case' in which I pull into a petrol station and fill up my car's tank, or go into a store and fill my bag with groceries, my obligation to pay for the goods I have obtained is quite clear. Similarly, there is little or no dispute about the illegality of the 'standard case' of murder, rape, theft, or battery. Yet, as even Hart eventually acknowledged, these standard cases are easy not because they can all be deduced from some literally interpreted, comprehensive, unambiguous rule, but because they are clear in light of the purpose(s) underlying the rule(s).

While rejecting 'grand theory,' the true 'legal realists,' who were progressive moderates rather than legal nihilists, explicitly or implicitly adopted the principlist approach. The most influential moderate legal realist was Karl Llewellyn. Llewellyn's attack on rules was aimed not at the notion of rules per se, nor at their admitted role in guiding and even controlling judicial decision-making. Instead, it was aimed at the beliefs that the real rules are the literal 'paper' rules articulated in statutes or judicial decisions rather than those derived from careful analysis of the concrete results of the courts' decisions, and that rules can or should be mechanically or deductively applied to decide particular cases without taking into account the purposes that underlie the rules. Llewellyn treated the underlying purposes as being themselves part of the law. He noted that the judge 'takes the rules of the game in the main not from his own inner consciousness, but from the existing practice, and again in the main from authoritative sources (which in the case of the law are largely statutes and the decisions of the courts).' He insisted that in interpreting those rules 'the court must be very slow in consulting its own views of rightness and welfare if they differ perceptibly from those of the community.' And he argued for a return to the 'Grand Tradition' of the common law: the articulation of, and deciding of cases in accordance with, the basic principles of justice that underlie the law.³⁸

Fuller also was a principlist. He emphasized that the creation and application of law is a purposive activity – the subjecting of human activity to the governance of reason through rules to further human ends. He argued, against the advocates of natural law, that legal interpretation must begin with and is constrained by the preexisting rules embodied in statutes and prior judicial decisions. On the other hand, he argued, against the strict legal positivists, that there is a necessary connection between law and

38 See Llewellyn, *op cit* n 11, at 8-10, 12-15, 20-23, 66-69, 73, 81-82, 156-159.

morality, since the articulated rules must be interpreted and applied in the light of their prior applications and their underlying purposes, which he assumed reflected the community's notions of justice and right.³⁹

Fred Schauer argues that going beyond the literal language of a rule to consider its underlying purpose completely undercuts the concept of a rule as a constraint on the decisionmaker.⁴⁰ But the underlying purpose imposes a constraint, and is itself constrained by the language of the rule. As Fuller noted, the necessity of resorting to underlying purposes does not render rules or their language irrelevant. The language of a statutory or common law rule imposes 'structural' constraints on purposive interpretation. Language does convey meanings, and the purposive interpretation must be a plausible one in the light of those meanings. At the very least, it cannot render the language meaningless, or make the rule a complete nullity.⁴¹

The best known modern principlist is Ronald Dworkin. Dworkin elaborates and refines Fuller's arguments.⁴² However, Dworkin's version of principlism contains serious, indeed fatal, flaws. The most serious flaw is that Dworkin departs from the crucial legitimating premise of principlism, which is that the judge must apply the *community's* moral principles as embodied in the preexisting law, rather than her own subjectively preferred moral principles. Dworkin's 'political order' and 'integrity' arguments on behalf of principlism gain what strength they have from their occasional explicit or implicit references to this premise.⁴³ Yet neither argument directly relies on this premise. Indeed, Dworkin disables himself from relying on this premise, since his interpretive methodology focuses on the judge's subjectively preferred moral principles rather than on the community's moral principles as embodied in the preexisting law. The latter principles would seem to be those principles that coherently explain and justify as much as possible of the preexisting law. Yet, rather than having a judge objectively determine which moral theory or coherent set of moral principles has the *best* descriptive fit with the preexisting law, Dworkin states that a judge should choose the most subjectively attractive normative theory that satisfies some *minimum threshold* of descriptive fit. Indeed, Dworkin even states that the more normatively attractive a certain moral theory is (to the particular judge), the less should be required in terms of the minimum threshold of descriptive fit.⁴⁴ Dworkin, in effect, invites a judge to shoehorn her own preferred moral theory into the preexisting law, opening the door wide (although not completely open) to the perils of the positivist penumbra.

The principlist, by treating the purposes underlying the preexisting statutes and judicial decisions as part of the law, and by insisting that such purposes be explicitly taken into account to decide the hard cases, avoids or at least substantially mitigates the problems of the positivist penumbra. Rather than retroactively creating new law, the judge is interpreting and applying

39 See Fuller, *Quest*, op cit n 21, at 3-15, 57-59, 64-65, 92-95, 99-104, 111-114, 121, 129-140; Fuller, *Morality*, op cit n 5, at 82-88, 106-107, 131-132, 224-232; Fuller, "Positivism" op cit n 15, at 632, 644, 646-648, 655-657, 661-668.

40 Schauer, op cit n 14, at 532-535.

41 Fuller, "Positivism" op cit n 15 at 670.

42 Dworkin, op cit n 22, at 165-173.

43 See Ronald Dworkin, *Law's Empire* (Harvard U Press, Cambridge, 1986), p 164-190; Dworkin, op cit n 22, at 183, 185-188.

44 Dworkin, op cit n 22, at 168-173, 179.

preexisting law. Rather than deciding the case in terms of what she personally thinks would be the best result, the judge should decide the case in accordance with the community's moral principles as embodied in the preexisting law.

However, as both Hart and the legal nihilists point out, the problem of the penumbra is avoided or mitigated only to the extent that a coherent set of moral principles, based on some foundational moral theory, in fact underlies the preexisting statutes and judicial decisions.⁴⁵ Does such a foundational moral theory exist, for the law in general or at least for tort law? We will limit our consideration to tort law, focusing on the two principal competing moral theories of law: the justice theory, for which the basic goal is the promotion of equal individual freedom, and the utilitarian efficiency theory, for which the basic goal is the maximization of aggregate utility or wealth. We will first briefly summarize each theory and then investigate the extent to which each explains and justifies some of the major elements of tort law.

IV. JUSTICE

In both ancient and modern times, the law in general and tort law in particular usually have been understood as being grounded on the concept of justice. While few if any are opposed to the idea of justice, some claim that it is an amorphous or question-begging concept, with no inherent substantive content. It is true that the concept of justice rarely is defined as precisely or elaborated as extensively as the competing concepts of utility or efficiency – perhaps because the idea of justice seems to be universally accepted and tacitly understood without the need for such precise definition or extensive elaboration. Yet it is also true that, from ancient times, the concept of justice has had a specific substantive content that places it in direct opposition to the principles of utilitarianism and economic efficiency.

The classical elaboration of the concept of justice and its relation to law and morality appears in book V of Aristotle's *Nicomachean Ethics*.⁴⁶ Aristotle's conception of the good is non-aggregative and assumes the equal moral worth of each individual as a free rational being. He rejects conceptions of the good that are based on wealth, pleasure, or enjoyment (which are the values that are to be maximized in utilitarian efficiency theories). Instead, he elaborates a conception of the good that is intrinsic to each individual: the full realization of one's humanity through activity in accord with a rational principle and with complete virtue over one's life. The goal of politics is the attainment of this common good for each and every citizen of the state, which Aristotle describes as a community of free and equal individuals. Law is the instrument by which the state achieves this common good by enforcing the requirements of justice.⁴⁷

The fundamental moral significance of persons' status as free and equal individuals is emphasized in Immanuel Kant's powerful elaboration of the foundations of classical liberal philosophy.⁴⁸ The foundation of Kant's moral

⁴⁵ See Hart, *Essays*, op cit n 16, at 7.

⁴⁶ See Wright, op cit n 30, at 683-711, and Richard W Wright, "Right, Justice and Tort Law" in *Philosophical Foundations of Tort Law* (OUP, Oxford, David G. Owen ed. 1995), pp159,175 from which most of the discussion in this section is taken.

⁴⁷ Aristotle, *Politics* III.6 at 1279a17-21.

⁴⁸ Immanuel Kant, *The Metaphysics of Morals* (1797) (Cambridge U Press, Cambridge, Mary Gregor trans. 1991); see Wright, op cit n 30, at 647-663.

philosophy is the idea of free will or freedom, by which he did not mean unrestricted pursuit of one's desires, but rather the opposite: fully realizing one's humanity by subjecting one's actions to the universal moral law, in order to free oneself from sensible inclinations in opposition to that moral law. According to Kant, freedom, and the moral personality constituted by its possession, is an inherent, internal, defining characteristic of each rational being. The possession of free will or freedom is what gives each rational being moral worth – an absolute moral worth that is equal for all rational beings:

[M]an regarded as a *person* [rather than a mere animal], that is, as the subject of a morally practical reason, is exalted above any price; for as a person (*homo noumenon*) he is not to be valued merely as a means to the ends of others or even to his own ends, but as an end in himself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world. He can measure himself with every other being of this kind and value himself on a footing of equality with them.⁴⁹

The supreme principle of morality (the categorical imperative) is '[a]ct only according to that maxim by which you can at the same time will that it should become a universal law,' which Kant reformulates as '[a]ct so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only.'⁵⁰

In the elaboration of his moral philosophy, Kant distinguishes between a doctrine of virtue (ethics) and a doctrine of Right (justice).⁵¹ The doctrine of virtue focuses on the internal aspect of the exercise of freedom – one's choosing and acting in accordance with the proper ends, by subjecting the maxim of one's actions to the condition of qualifying as universal law.⁵² The doctrine of Right, on the other hand, focuses on the external aspect of the exercise of freedom – the constraints on action required for the practical operation of freedom in the external world. It specifies which moral obligations are also legal obligations, enforceable through coercion by others. The supreme principle of Right, which is a corollary of the categorical imperative, is 'so act externally that the free use of your choice can coexist with the freedom of everyone in accordance with a universal law.'⁵³

The categorical imperative and its corollary, the supreme principle of Right, are similar to the golden rule, '[d]o unto others as you would have them do unto you,' which appears (in various forms) as a fundamental principle in many religions and moral theories. However, as Kant noted, the categorical imperative is both broader in scope and more demanding than the golden rule. It is morally wrong under the categorical imperative to fail to respect the absolute moral worth of anyone, including yourself, as a self-legislating rational being, regardless of whether you would allow

49 Kant, op cit n 48, at *434-435.

50 Immanuel Kant, *Foundations of the Metaphysics of Morals* (1785) (Bobbs-Merrill, New York, Lewis W Beck trans. 1959), * 421,429.

51 Kant, op cit n 48, at *218-220, 379-380, 396-397, 406.

52 Among the morally obligatory ends Kant includes one's own perfection (including cultivation of one's physical, mental, moral and social capacities) and the happiness of others (through beneficence, charity, respect, friendship, etc.). He distinguishes these morally obligatory ends from one's own happiness, an end which one naturally has and hence for which the concept of duty is inapplicable, and the perfection of others, an end which others can only set and pursue for themselves. Ibid at *385-388, 391-393, 418 n*, 448-473.

53 Ibid at *231.

others to treat you without proper respect.⁵⁴ All persons should be treated as *ends in themselves* (i.e., as free and equal persons seeking to fully realize their humanity), rather than as mere means to be used to benefit others or society as a whole (as is allowed and indeed required under the utilitarian efficiency theory).

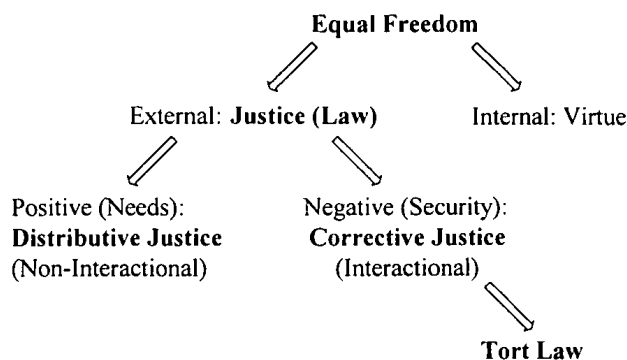
Many scholars overlook Kant's distinction between the doctrine of Right and the doctrine of virtue and consequently misunderstand and misstate the implications of Kant's moral philosophy for law in general and tort law in particular. Kant's doctrine of virtue, while specifying objective moral duties based on the categorical imperative, assesses the moral *blame* or *merit* of conduct in terms of the individual's subjective capacity and effort in attempting to ascertain and satisfy those duties. Kant's doctrine of Right, on the other hand, directly applies the objective requirements of the categorical imperative in determining an individual's moral and legal *responsibility* for the adverse effects that she causes to the person or property of others. Thus, as Kant repeatedly emphasizes, an action's *legality* is judged by its external conformity with the objective requirements of Right, while its *morality* is judged by the actor's internal subjective capacity and efforts to conform her conduct to the objective requirements of the categorical imperative, which grounds all duties of Right and virtue.⁵⁵

There are two distinct types of substantive justice, which were first explicitly distinguished by Aristotle: distributive justice and corrective justice. Distributive justice deals with the public or communal resource-allocation issues. Its aim is a distribution of the community's resources that implements each person's right to equal freedom by giving her an equal or sufficient opportunity to fully realize her humanity as a self-legislating moral being. Corrective justice, on the other hand, deals with private interactions, and it requires that those engaged in such interactions do so in a way that is consistent with the right to equal freedom of the parties to the interaction.

To put it another way, distributive justice and corrective justice encompass the two different aspects of equal external freedom. Distributive justice defines the scope of a person's *positive freedom* to have access to the resources necessary to realize her humanity, and corrective justice defines the scope of a person's *negative freedom* not to have her person or existing stock of resources interfered with by others. Together, distributive justice and corrective justice seek to assure the attainment of the good (full realization of one's humanity) by each person by providing her with a proportionately equal share of the needed resources (distributive justice) and by safeguarding her person and existing stock of resources from interactions with others that are inconsistent with her status as a rational being with equal absolute moral worth (corrective justice).

54 '[The golden rule] is only derived from the [categorical imperative] and is restricted by various limitations. It cannot be a universal law, because it contains the ground neither of duties to one's self nor of the benevolent duties to others (for many a man would gladly consent that others should not benefit him, provided only that he might be excused from showing benevolence to them). . . .' Kant, op cit n 50, at *430 n 14.

55 For example, Kant, op cit n 48, at *214, 218-219, 225-228, 379-382, 390-391; cf. *ibid* at *488 ('All moral relations of rational beings, which involve a principle of the harmony of the will of one with that of another, can be reduced to *love* and *respect*; and, insofar as this principle is practical, in the case of love the basis for determining one's will can be reduced to another's *end*, and, in the case of respect, to another's *right*.').



Given tort's law focus on interactional injuries, it falls within the domain of corrective justice (as do criminal law, contract law, the law of trusts, and the law of restitution). Corrective justice, and hence tort law, protects against actual or threatened interactional injuries to one's person or existing stock of resources regardless of the distributive justice or injustice of the overall division of resources among the parties to the interaction or among the members of the community as a whole. These distributive justice concerns are generally outside the scope and competence of tort law, which comes into play only sporadically and focuses on isolated interactions among individuals.⁵⁶ Tort law as corrective justice requires that each person's external freedom be exercised in a way that is objectively consistent with the equal external freedom of all others whose persons or property might be affected by her conduct.

Misunderstandings about the nature and scope of corrective justice abound in scholarly commentary on tort law. Corrective justice does not assume absolute entitlements, which would be invaded by any adverse effect. It only applies to injuries to one's person or existing stock of resources that are caused by (correctively) unjust conduct – i.e., conduct that is inconsistent with others' equal freedom. Therefore it makes no sense in evaluations of tort law as an instrument of corrective justice to fault it for failing to compensate all injuries or to redress all loss (as if that were even theoretically possible). The purpose of tort law as corrective justice is to provide rights of redress only for *unjust* injuries, not all injuries. Similarly, there is no *per se* failure of corrective justice every time a person who suffers an unjust injury does not obtain redress. Corrective justice rights may be waived by the holder of the right, and very often are waived. There is a failure of corrective justice only when someone who is both entitled to redress and wants redress fails to obtain it. Finally, contrary to a very common misunderstanding, there is no general requirement that a corrective justice duty be discharged personally by the party who is subject to that duty. Such a requirement applies only when the appropriate mode of rectification is punishment. When the appropriate mode of rectification is compensation for the unjust loss, rather than or in addition to punishment, corrective justice merely establishes the duty of the party who caused the unjust loss to see to it that the required compensation occurs. There is nothing in corrective justice which prevents that duty from being discharged voluntarily, on behalf of the party with the duty, by someone else – e.g.,

⁵⁶ See Wright, *op cit* n 46, at 171-174.

that party's insurer or rich aunt. Nor is there any problem from the standpoint of corrective justice if the entity that discharges the duty spreads the cost of discharging that duty to others through voluntary market processes – e.g., by raising the prices of its products.

V. UTILITARIAN EFFICIENCY

Much current academic writing assumes that tort law is or should be based on the principle of utility or its modern economic variant, the economic efficiency theory. While there are important differences between utilitarianism and the economic efficiency theory, the core assumptions are the same, and it will be convenient to refer to these core assumptions as the utilitarian efficiency theory.

Utilitarianism was propounded in the eighteenth and nineteenth centuries by Jeremy Bentham and his followers. The principle of utility or 'greatest happiness' mandates actions which produce the greatest sum of happiness (or pleasure or preference-satisfaction) as added up for the citizenry in the aggregate.⁵⁷ Bentham's principle of utility was intended to be progressive rather than conservative. It was motivated by a reaction against the concentration of wealth and power in the hands of the English aristocracy and was meant to support progressive reform legislation for the benefit of the general citizenry. Nevertheless, there is no independent weight given in the utilitarian theory to the distribution of happiness (or wealth or power) or to the promotion of individuals' equal (positive and negative) freedom. On the contrary, each individual's freedom and interests are subordinated to the maximization of the total happiness or preference-satisfaction for the citizenry in the aggregate. As Bentham's follower, John Stuart Mill, emphasized:

[T]he happiness which forms the utilitarian standard of what is right in conduct is not the agent's own happiness but that of all concerned. As between his own happiness, and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator. In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility. 'To do as you would be done by,' and 'to love your neighbor as yourself,' constitute the ideal perfection of utilitarian morality.⁵⁸

As should be obvious, the conception of equality in utilitarianism is quite different from the equal freedom norm that underlies the justice theory. In utilitarianism, each individual counts equally methodologically only, as an equal and fungible addend in the calculation of the aggregate sum.⁵⁹ Any individual's freedom or interests can and should be sacrificed whenever doing so would produce a greater total of aggregate happiness. It is not permissible to prefer one's own interests or projects, or those of one's family members or friends, over those of any other person, except to the extent that doing so would produce a greater total happiness for the citizenry in the aggregate. Although Mill attempts to equate the principle of utility with the golden rule, the usual (first) form of the golden rule is, as Kant

57 Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (1789), ch. 1.

58 John Stuart Mill, *Utilitarianism* (1863), ch. II.

59 '[O]ne person's happiness, supposed equal in degree (with the proper allowance made for kind), is counted exactly as much as another's. . . . [n.] . . . [T]he truths of arithmetic are applicable to the valuation of happiness, as of all other measurable quantities.' Ibid at ch. V.

noted, merely a universalization rule.⁶⁰ It is very doubtful that many people would agree to the utilitarian aggregate-maximization and impartiality-of-interest requirements (the latter of which Mill equates with the 'love your neighbor as yourself' variant of the golden rule), whether or not they were made universally applicable.

Utilitarians reject the idea of individual autonomy or rights, insofar as those rights are understood (as they usually are) as being independent of or in conflict with the principle of utility. Bentham was quite dismissive of the idea of rights, especially 'natural rights.'⁶¹ Mill similarly viewed rights as being ultimately dependent on utilitarian calculation, although he attempted to develop a utilitarian account of 'that apparent infinity and incommensurability with all other considerations which constitute the distinction between the feeling of right and wrong and that of ordinary expediency [utility] and inexpediency.'⁶²

The dominant variant of utilitarianism in modern legal theory and policy analysis is the theory of economic efficiency. Although the economic efficiency theory departs from utilitarianism in significant ways, its prevailing form shares with utilitarianism the focus on maximizing total aggregate happiness (or welfare or wealth) rather than the promotion of individuals' equal positive and negative freedom.

Classical economists employed a severely constrained criterion of economic efficiency, Pareto optimality. A situation is Pareto optimal if no one could be made better off without making at least one person worse off. If, on the other hand, a shift in resources can be made that will make one or more persons better off without making anyone worse off, the situation after the shift is deemed to be Pareto superior to the situation before the shift since it increases the total aggregate welfare or wealth, regardless of its effect on the distribution of welfare or wealth.

The Pareto criterion is inconsistent with both utilitarianism and the equal freedom norm. It blocks utility-maximizing resource transfers that would be approved (required) by the principle of utility. Although it is often assumed, by Smillie and others, that the Pareto criterion is consistent with and secures individual freedom or autonomy, it does not. Contrary to a common assumption, the Pareto criterion does not require that any action affecting others be done only with the consent of all the affected parties.⁶³ A forced exchange is permissible (mandated) if at least one person is made better off and (through compensation or otherwise) no one is made worse off, regardless of lack of consent or palpable unfairness in the distribution of the net benefits generated by the transfer. Conversely, the criterion's insistence that no one ever be made worse off is inconsistent not only with the promotion of equal positive freedom under the distributive-justice principle, but also with proper establishment of the equal-freedom-

60 See op cit n 54.

61 Bentham, op cit n 57; Jeremy Bentham, "Anarchical Fallacies" in *Nonsense Upon Stilts: Bentham, Burke and Marx on the Rights of Man* (1838-1843) (J Waldron ed. 1987), p 53.

62 Mill, op cit n 58, at ch. V. For a powerful critique, see Will Kymlicka, *Contemporary Political Philosophy* (Clarendon Press, Oxford, 1990), pp 21-44.

63 Contra Richard A Posner, *Economic Analysis of Law* (Little, Brown & Co., Boston, 3rd ed, 1986), p 12; Smillie, op cit n 1, at 271-272.

maximizing level of security (negative freedom) under the corrective-justice principle, since imposing liability for all risk-creating activity, no matter how slight the risk, undoubtedly would lower everyone's equal individual freedom rather than enhancing it.

The requirement that no one be made worse off makes the Pareto definition of efficiency quite conservative and practically useless as a public policy tool, since very few shifts in resources, other than those accomplished through contracts that include all the affected parties, do not make at least one person worse off. Hence modern economists have turned instead to the Kaldor-Hicks definition of economic efficiency, according to which a shift in resources is efficient if those who gain from the shift gain enough so that they *could have* compensated those who lose. However – and this is the crucial departure from the Pareto criterion – no compensation is actually required. The shift is deemed efficient as long as the winners gain more than the losers lose.

Therefore, as with the principle of utility, the focus of the Kaldor-Hicks efficiency criterion is solely on the unconstrained maximization of aggregate welfare, regardless of any adverse impacts on individuals' equal positive or negative freedom. Moreover, the efficiency criterion employed by economists usually departs from the principle of utility in its definition of value. The principle of utility seeks to maximize aggregate subjective utility (happiness or pleasure or preference-satisfaction). Since it is difficult and often impossible to make quantitative or even qualitative interpersonal comparisons of utility, economists usually seek instead to maximize aggregate wealth – the value of society's resources as measured by persons' *willingness and ability to pay* for them (or to accept payment for them, if they already own them). As the economists sometimes acknowledge, this notion of value is highly problematic, since what one is willing and able to pay for something depends significantly on one's pre-existing wealth.⁶⁴

In sum, as Smillie seems to recognize, both utilitarianism and the economic efficiency theory are completely at odds with, and normatively much less attractive than, the equal freedom norm that underlies the justice theory.⁶⁵ The moral good in the justice theory is the full realization of one's humanity as a free and equal being, while the moral good under the principle of utility is pleasure or preference satisfaction, and the moral good under the economic-efficiency theory is resource wealth as measured by one's willingness and ability to pay for those resources. The equal freedom theory focuses on the promotion (maximization) of each person's equal freedom to pursue a morally meaningful life (i.e., gives primacy to equal distribution of the good), while the utilitarian efficiency theories focus on maximizing the total sum of the good (pleasure, preference satisfaction, or wealth) with no independent concern for how that total sum is distributed among individuals.

64 Posner, *op cit* n 63, at 11-12; see *ibid* at 13, 14-15. Yet, as Posner points out, 'Utility in the utilitarian sense also has grave limitations, and not only because it is difficult to measure when willingness to pay is jettisoned as a metric. The fact that one person has a greater capacity for pleasure than another is not a very good reason for a forced transfer of wealth from the second to the first.' *Ibid* at 12.

65 See Smillie, *op cit* n 1, at 269-271. But see *ibid* at 272-273 (indicating a willingness to employ a qualified Kaldor-Hicks analysis in conjunction with the Pareto criterion in a quasi-libertarian approach).

The legal economists usually acknowledge that economic efficiency is a secondary goal that is subsidiary to and constrained by the justice goal. However, they almost always ignore the justice goal in their analyses. They set distributive justice aside as merely involving possible redistributions of wealth that supposedly do not affect wealth-maximizing resource transfers. They have a very hard time perceiving any other notion of justice, especially corrective justice. They interpret appeals to justice as implicit references to efficiency, attempts to redistribute wealth, or irrational pseudo-justice sentiments.⁶⁶ Their primary argument is that, although the efficiency theory may appear to be normatively unattractive, it seems to be the moral theory that actually underlies our law and politics, since it allegedly best (or even solely) explains the content of our existing laws.⁶⁷ From the beginning, the prime exemplar for this last assertion has been tort law.

The efficiency theory of tort law, as of the law in general, considers three (conflicting) subgoals: efficient deterrence of non-wealth-maximizing behaviour before it occurs, efficient compensation of losses that have already occurred, and efficient administration of the systems set up to achieve efficient deterrence and efficient compensation.⁶⁸

The efficient compensation goal relies on the utilitarian conception of value – i.e., utility or happiness rather than wealth (as measured by willingness and ability to pay). Past losses are “sunk costs” that no longer can be efficiently avoided or minimized. However, assuming the declining marginal utility of money (i.e. that the more one has, the less utility one gets from each additional unit), aggregate utility can be increased by shifting the losses to wealthier persons or entities or by spreading the losses thinly across many more people. The loss initially concentrated solely and entirely on the tort plaintiff will cause less total disutility or unhappiness if it is shifted to a wealthier person or spread more thinly across many more people.

However, it is generally acknowledged that tort law does not seek to implement, nor would it be an effective means of implementing, this efficient compensation goal. The parties’ relative wealth or ability to spread losses is generally deemed to be irrelevant to tort liability (despite some flirtation with the loss-spreading rationale in product liability theory). Moreover, the persons best able to bear or spread some loss often will have had nothing to do with the injury and are not before the court. The best method for efficiently spreading losses is general social insurance, not tort law. For these reasons, plus the fact that the efficient compensation goal conflicts with the efficient deterrence goal, efficient compensation is usually dismissed as a plausible goal for tort law.⁶⁹

While some legal economists assert that tort law is an inappropriate tool for achieving either efficient deterrence or efficient compensation,⁷⁰ most

66 See for example, Guido Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (Yale U Press, New Haven, 1970), pp 24-33, 291-308; Posner, op cit n 63, at 11-13, 21-26.

67 For example, Posner, op cit n 63, at 23-24.

68 See Calabresi, op cit n 66, at 24-33.

69 Ibid at 27-29, 39-67, 278-285; Posner, op cit n 63, at 186-187, 189-190.

70 See Calabresi, op cit n 66 at 239-277.

legal economists assert that the basic goal of tort law is and should be efficient deterrence – the encouragement of efficient (aggregate-wealth maximizing) behaviour and the discouragement of inefficient behaviour. When discussing efficient deterrence, the economists switch from utility to wealth, as measured by willingness and ability to pay, as the measure of value. A basic distinction is drawn between situations involving low transaction costs and situations involving high transaction costs.

In situations involving low transaction costs, the economists assert that the parties should be forced to engage in consensual resource transfers. The preference for consensual resource transfers is based on the assumption that anyone who voluntarily agrees to a certain transfer must have been made better off by it (or at least no worse off), and thus that such consensual transfers increase efficiency. With perfect information, the government would know the most efficient allocation of every resource and could directly shift all resources to their most efficient use, thus saving the costs of negotiation and compensation. But absent such perfect information, the best evidence of anyone's subjective valuation of a resource is assumed to be what he or she actually is willing (and able) to pay for it in a consensual exchange. However, when the costs of arranging and completing market transactions that would include all the affected parties are so high as to block the efficient transfer of a resource to its highest-valued use, or would constitute a significant "waste" of resources on these transaction costs even if the transfer were not blocked, the economists advise the use of ownership and liability rules that will encourage actors to 'mimic the market' – i.e., achieve the efficient result through a government authorized or ordered nonconsensual transfer (generally without compensation for the forced transfer).⁷¹

Thus the legal economists' basic prescription for tort law is to design legal rules that encourage or require people to engage in consensual resource transfers in situations involving low transaction costs and in forced (nonconsensual) 'market simulating' resource transfers, with or without compensation, in situations involving high transaction costs.

VI. TORT LAW

Is either the justice theory or the utilitarian efficiency theory deeply rooted in current tort law, so that it could and should be used in the principled approach to adjudication (and legislation) as the foundational moral theory that underlies and gives coherence and content to the existing positive law? In a previous essay, I argued that the various major aspects of the standards of care in negligence law – the different risks taken into account for defendants and plaintiffs, the different perspectives applied in evaluating the reasonableness of defendants' and plaintiffs' conduct, and the different criteria of reasonable care applied in different types of situations (depending on who was putting whom at risk for whose benefit) – all could be clarified, explained, justified and illuminated by the justice theory, while none could be explained or justified by the utilitarian efficiency theory.⁷² In the remainder

⁷¹ See for example, Posner, *op cit* n 63, at 13-14, 48-50, 192-194, 203-204.

⁷² See Richard W Wright, "The Standards of Care in Negligence Law" in *Philosophical Foundations of Tort Law* (OUP, Oxford, David G. Owen ed. 1995), p 249.

of this essay, I will address a few more of the major aspects of tort liability, ending with a discussion of the particular issue upon which Professor Smillie focused in his second recent article: the availability of punitive or exemplary damages in tort law. I believe that the discussions in my prior essay and in this essay, although far from comprehensive, provide more than sufficient grounds for rejecting the utilitarian efficiency theory and strong evidence in favor of the plausibility and value of the justice theory.

A. Plaintiff's Consent as a Complete Defence

In tort law, the plaintiff's consent is a complete defence to any tort action, whether it be an intentional tort, a negligence action, or a strict liability action. Although the defence of assumption of the risk in negligence actions has been narrowed in some jurisdictions, it still generally applies to situations involving true subjective consent: a willing agreement to be exposed to a particular known risk.⁷³

The treatment of true subjective consent as a complete defence (or lack of consent as a necessary element of the prima facie case for certain torts) is straightforward under the justice theory, since if the plaintiff has willingly agreed to the injury (or to be exposed to the risk of its occurrence), there has been no violation of his autonomy, dignity, or rights. As is stated in the often repeated legal maxim, *volenti non fit injuria*: to the willing no injury can be done.

It might seem that consent would also be conclusive against liability under the utilitarian efficiency theory, since presumably the person consenting would not consent unless she believed that the expected injury or risk of injury was outweighed by some expected benefit. However, this might not be the case if she misperceives the expected risks and benefits. More importantly, under the utilitarian efficiency theory, we must take into account the risks and benefits to everyone, including those financially or emotionally dependent on her. Although the utility she expects to obtain – e.g., by undergoing a medically risky cosmetic operation or refusing to undergo a medically necessary operation – may outweigh the risks to her alone, they may well not outweigh the expected losses to those financially and emotionally dependent on her.

B. Plaintiff's Contributory Negligence as a Complete or Partial Defence

The plaintiff's alleged contributory negligence traditionally has not been allowed as a defense to an intentional tort action.⁷⁴ This position again is straightforward under the justice theory. The plaintiff should not be required to limit his rightful exercise of his freedom, including his use of his property, in order to avoid foreseeable wrongful intentional injury by another. As was said in a famous opinion by the Supreme Court of the United States:

That one's uses of his property may be subject to the servitude of the wrongful use by another of his property seems an anomaly. It upsets the presumptions of law and takes from him the assumption and the freedom which comes from the assumption, that the other will obey the law, not violate it. It casts upon him the duty of not only using his own property so as not to injure another, but so to use his own property that it may not be injured by the wrongs of another. How far can this subjection be carried out?⁷⁵

73 See for example, John G Fleming, *The Law of Torts* (Law Book Co., Sydney, 8th ed, 1992), pp 79-82, 291-307, 343-344, 362-363.

74 *Ibid* at 281.

75 *LeRoy Fibre Co v Chicago, Milwaukee & St Paul Ry.* (1914) 232 US 340 at 349.

Similarly, a defendant who assaults or batters a plaintiff as she is walking through a deserted area at night should not be able to argue that his liability should be reduced because she was contributorily negligent by being in the area or by wearing 'provocative' clothing.

The efficiency theorists cannot rely on this direct appeal to the respective rights involved. They instead argue that in all such situations the efficient level of care for the victim is zero since there is zero or negative utility generated by the defendant's wrongful intentional conduct.⁷⁶ But, for the true utilitarian or efficiency theorist, the utility the defendant gains from his conduct cannot be ignored or assumed to be less than the disutility suffered by the plaintiff. This is always an empirical question, as is the question of whether the expected aggregate benefits would have been greater or the expected aggregate costs would have been less if the plaintiff had behaved differently.

The plaintiff's contributory negligence is generally recognized as a defence in a negligence action. At one time, the plaintiff's contributory negligence theoretically was a complete defence to the defendant's liability. Currently, in almost every common law and civil law jurisdiction, it is only a partial defence, which merely reduces the plaintiff's recovery in proportion to his comparative responsibility for the injury, except in 'modified' (rather than 'pure') comparative responsibility jurisdictions, in which the plaintiff's contributory negligence still operates as a complete defence if the plaintiff's comparative responsibility exceeds the defendant's. Jurisdictions other than the United States and courts in the United States generally have preferred pure comparative responsibility, while legislatures in the United States have usually preferred modified comparative responsibility.⁷⁷

The relevance of the plaintiff's contributory negligence to the defendant's negligence liability is fairly clear under the justice theory. Although one cannot be unjust to oneself, the plaintiff who fails to properly respect his own humanity by exposing himself to unreasonable risk is morally responsible, along with the negligent defendant, for the resulting injury, and such moral responsibility properly can and should be taken into account to at least reduce the plaintiff's recovery in proportion to his comparative responsibility and perhaps even to completely bar his recovery if his comparative responsibility (substantially) exceeds the defendant's.

The prior practice of treating the plaintiff's contributory negligence as a complete defence, even if the plaintiff's comparative responsibility was less than the defendant's, has often been criticized as an overly zealous application of the principle of individual moral responsibility, according to which the plaintiff with "unclean hands" was not allowed any recovery, no matter how slight his negligence was compared to the defendant's. However, both the "unclean hands" explanation and another popular explanation (the supposed desire to subsidize developing industry and commerce)⁷⁸ fail to account for the formal doctrines and informal practices that permitted many contributorily negligent plaintiffs to

76 See for example, Posner, *op cit* n 63, at 194.

77 See Fleming, *op cit* n 73 at 269-275.

78 *Ibid* at 269-270.

recover all or part of their losses from negligent defendants. A more plausible explanation, explicitly stated in some cases and implicit in many others, is that it was deemed conceptually or at least practically impossible to assess relative degrees of negligence or negligent causation.⁷⁹ In the absence of an explicit yardstick or apportionment formula, the formal doctrines of negligence, contributory negligence, last clear chance, active versus passive negligence, slight versus gross negligence, and so forth were employed as very crude measures of comparative responsibility to shift the entire loss back and forth between the plaintiff and the defendant. In addition, the contributory negligence issue was usually left to the jury, and it was widely understood that juries frequently found for the plaintiff, awarding full or reduced compensation, even when the plaintiff was contributorily negligent. Eventually, as it became increasingly clear that juries were quite capable of assessing comparative responsibility despite the lack of any explicit yardstick or formula, and that the lack of explicit authorization of this practice undermined the integrity of and respect for the law, the practice was explicitly authorized by statutory and judicial adoption of liability based on comparative responsibility.⁸⁰

Thus, both the former and current negligence liability rules are fairly easily explained under the justice theory, with the current rules being a truer and more explicit elaboration of the moral principles of liability that also underlaid the prior practice.

The utilitarian efficiency theory cannot explain the existence of the contributory negligence defence, especially as a partial rather than complete defence. Under the typical efficiency analysis, the defence is unnecessary to achieve efficient conduct by both the defendant and the plaintiff. The defendant's potential liability if she is negligent will cause her to behave efficiently (nonnegligently) to avoid such liability. The plaintiff, knowing this and realizing that he will have to bear any injury which nevertheless results, will employ the efficient (nonnegligent) level of precaution that minimizes his expected total costs (precaution costs plus injury costs). The legal economists thus attempt to explain the existence of the contributory negligence defence by referring to administrative costs. They argue that the negligence rule with a complete defence of plaintiff's contributory negligence would be cheaper to administer than the negligence rule with no contributory negligence defence, given the probable greater number of claims that would be brought and liability payments that would have to be made under the latter rule. However, they fail to take into account the higher administrative costs per claim under the former rule due to the need to evaluate the

79 See for example *Needham v San Francisco & SJRR* (1869) 37 Cal 409 at 419 ('there is no standard by which the law can measure the consequences of [defendant's] fault, and therefore, and therefore only, he is allowed to go free of judgment').

80 See Joseph W Little, "Eliminating the Fallacies of Comparative Negligence and Proportional Liability" (1989) 41 Ala LR 13 at 20-27 (summarizing the history and noting the courts' reliance on arguments of justice rather than efficiency); Richard W Wright, "Allocating Liability Among Multiple Responsible Causes: A Principled Defense of Joint and Several Liability for Actual Harm and Risk Exposure" (1988) 21 UC Davis LR 1141 at 1156-1158.

plaintiff's alleged contributory negligence, and it is not at all clear which of these offsetting costs would predominate. Moreover, they admit their inability to explain or justify the current comparative responsibility regimes, under which the plaintiff's contributory negligence is only a partial defence. Under the comparative responsibility regimes, more claims will be brought (and perhaps litigated), more liability payments will have to be made, and the resolution of each claim will involve not only the costs of evaluating both the defendant's and the plaintiff's negligence but also the additional significant costs of determining comparative responsibility.⁸¹

C. Intentional Torts versus Negligence

There is a basic distinction in tort law between intentional and unintentional injury. For unintentional injury, the defendant is generally liable only if she was negligent (i.e., behaved unreasonably in light of the foreseeable risks to others) and her negligence caused actual harm to the plaintiff's person or property.⁸² However, unless the defendant was justifiably defending against aggression by the plaintiff or had the plaintiff's consent, the defendant is liable for an intentional injury even if it was reasonable (e.g., occasioned by a reasonable mistake or private necessity). For some of the intentional torts, those involving trespasses against the plaintiff's person or property, the defendant is liable even in the absence of any actual harm to the plaintiff's body or property.⁸³

The justice theory, based on the equal freedom norm, easily explains tort law's distinct treatment of intentional and accidental injury. Intentional intrusions on one's person or property are morally qualitatively distinct from accidental intrusions. As Oliver Wendell Holmes said, "[E]ven a dog distinguishes between being stumbled over and being kicked."⁸⁴ Persons' external freedom, autonomy, and dignity are highly dependent on the security of their persons and property, especially from physical intrusion or injury. Intentional (purposeful or knowing) invasion not only often demonstrates the most serious lack of respect for another by the intentional actor, but also is much more easily avoidable than accidental intrusion. Thus, the equal freedom and dignity of all is promoted by requiring consent for many types of intentionally inflicted harm (but not, e.g., for the known or even desired financial effects of fair business competition), or even for nonharmful intentional physical intrusions on a person's body or land or significant interferences with his dominion or use of his personal property. On the

81 See Posner, *op cit* n 63, at 156-158; Wright, *op cit* n 80, at 1169-1175. When discussing the similar issue of the relative administrative costs of implementing strict versus negligence liability regimes, Posner notes that although strict liability's more liberal recovery rule would seem to increase the number of tort claims, it would reduce the administrative costs per claim by eliminating the costs of determining negligence, might or might not encourage settlements rather than going to trial due to its greater predictability, and might even reduce the number of accidents and hence claims due to its greater predictability. See Posner, *op cit* n 63, at 164, 528-529.

82 There are pockets of strict liability, for example, for ultrahazardous activities. However, in this article I will limit discussion to the two principal types of tort liability: liability for intentionally or negligently inflicted injury.

83 See Fleming, *op cit* n 73 at 15, 24-30, 39-41, 52-53, 76-98.

84 Oliver Wendell Holmes, *The Common Law* (1881) (Little, Brown & Co., Boston, Mark DeWolfe Howe ed. 1963), lect. I, at 7.

other hand, requiring consent for conduct which only risks accidental intrusions, or imposing strict liability for every accidental intrusion, especially if the defendant's conduct was deemed reasonable, would seriously erode everyone's (equal) freedom.

In certain circumstances involving private necessity or the creation or preservation of important public benefits, intentional intrusions without consent are sometimes allowed, but only subject to strict conditions that prevent any significant impairment of the plaintiff's equal freedom. In situations involving private necessity, the defendant in an emergency situation may intentionally enter on or even damage the plaintiff's land or personal property, but not the plaintiff's person, to save life or much more valuable property, but the defendant is required (at least in the United States) to compensate the plaintiff for any actual harm to the plaintiff's property.⁸⁵ Similarly, the government may exercise the power of eminent domain to take the plaintiff's property without his consent for some important public use, but only upon payment of just compensation to the plaintiff. Under the tort law applicable to the intentional creation of a private nuisance, a private entity may be allowed to acquire, without consent, a private servitude on a neighbour's land to continue operations which have been determined to constitute a private nuisance, but only if the operation has an important public rather than merely private benefit, the public benefit greatly outweighs the harm to the plaintiff's interest in the use and enjoyment of his land, there is no significant health threat to the plaintiff, it is not feasible to abate the nuisance without shutting down or reducing operations to such an extent as to lose or significantly reduce the public benefit, and the plaintiff is fully compensated for the loss in the use and enjoyment of his property.⁸⁶

In each of these three situations, the defendant's conduct is deemed reasonable and justifiable only if there is no significant health risk to the plaintiff, and the defendant is required to compensate the plaintiff although the defendant's conduct is deemed reasonable and justifiable. In the first, private necessity situation, the exigency justifies a qualified privilege to trespass on the plaintiff's property. However, since the justification is not based on the plaintiff's aggression against the defendant, but rather on external circumstances for which the plaintiff is not responsible, the plaintiff is entitled to have his equal negative freedom protected by requiring the defendant to compensate him for the intentional damage to his property – otherwise the defendant would be being allowed to use the plaintiff (through intentional damage to his property) as a mere means to the ends of the defendant.

Similar considerations apply in the two takings situations. In each situation, the important public benefit gives rise to a distributive-justice justification for taking all or part of the plaintiff's property. If there were no public benefit, the defendant would not be allowed to force a nonconsensual taking upon the plaintiff, but rather would have to obtain the plaintiff's

85 See for example, *Vincent v Lake Erie Transp. Co* (1910) 124 NW 221 (Minn); Fleming, op cit n 73, at 95-98.

86 See for example, *Boomer v Atlantic Cement Co* (1970) 257 NE2d 870 (NY); *Estancias Dallas Corp. v Schultz* (1973) 500 SW2d 217 (Tex Civ App); cf. Fleming, op cit n 73 at 418-424, 445-446.

consent to the transfer of the plaintiff's property right. However, unlike the usual distributive-justice situation, the distributive justice claim in these takings situations is not that the plaintiff has more than his fair share of society's resources or, even if he does, that he is the only person who has too many resources.⁸⁷ Indeed, he may actually have less than his fair share of society's resources. Rather, the distributive justice claim is that his property is required to accomplish the distributive-justice objective – typically the creation or preservation of an activity that will increase the total amount of resources in society so that everyone's distributive share can be increased – and that a nonconsensual (but compensated) taking rather than a voluntary sale is required to prevent him from demanding excessive compensation for his property and thus appropriating an undue share of the public benefit from the project to himself. To take his property without compensation in this context would be to use him merely as a means to the ends of others.

The utilitarian efficiency theory is unable to explain tort law's distinct treatment of intentional and accidental injury. The fact that a certain consequence is desired or known to be bound to occur, rather than merely being possible or probable or highly likely, does not matter under the efficiency theory, which is concerned solely with the aggregate net benefits or costs. Thus, the legal economists' attempted explanation of intentional tort liability ignores the presence or absence of intent and instead focuses on the distinction between situations involving low and high transaction costs.

According to the economists, the intentional torts are situations involving low transaction costs, in which the efficient approach is to force persons to engage in consensual transactions, while negligence liability applies to situations involving high transaction costs, in which the consent requirement (and the practical necessity of paying compensation to obtain consent) is abandoned and persons are allowed (encouraged) to inflict nonconsensual and uncompensated injuries on others whenever doing so would increase expected aggregate wealth.⁸⁸ To force persons to engage in consensual transactions in situations involving low transaction costs, we must impose extra-compensatory (punitive) damages if the defendant goes ahead without the required consent. To obtain consent beforehand, the defendant would have had to pay for any expected losses plus a bit extra, as well as the transaction costs involved in obtaining the consent. Thus, to remove her expected savings from going ahead without consent, we must impose liability not only for the actual losses but also for the 'bit extra' that would have been required to obtain consent and for the avoided transaction costs. Moreover, we must apply a 'punitive multiple' to these damages, equal to the inverse probability of her being caught and successfully sued, to make up for her discounting her expected liability due to the possibility of not being caught and being successfully sued.⁸⁹

In sum, the economists argue that all intentional torts involve low transaction cost situations (and vice versa), that all negligence actions involve high transaction cost situations (and vice versa), and that all intentional torts give rise to punitive as well as actual damages, with the punitives being calculated by applying a multiplier to the defendant's avoided costs

87 See Wright, *op cit* n 46, at 171-174, 177.

88 *Op cit* n 71.

89 See for example, Posner, *op cit* n 63, at 191-195, 203-204.

equal to the inverse of the probability of her being caught and successfully sued.

It should be obvious that none of this fits actual tort law. For example, assume that you decide to use, without consent, your neighbor's house and car while he is away and unreachable on a month-long trek in the Himalayas. You will keep an eye on, fix up and maintain everything, pay all related bills and reasonable rent, and leave all his property in better condition than when he left. Despite the impossibility of locating and negotiating with your neighbor and the apparent net benefit to you and him of your using his property, you will be liable for intentional trespasses to his land and his car, and likely for punitive as well as nominal damages. Conversely, assume your neighbor is home and you play baseball in your backyard, with a significant but not substantially certain risk that a ball will be hit through his bay window. If a ball actually is hit through his window, your liability will be determined under the negligence action and will be limited to the actual damages, although the costs of negotiating with your neighbor in this situation would be very low (as low as in any situation). Finally, consider a malicious assault, battery or rape, done for the pure pleasure of threatening or inflicting violence on another without her consent. Transaction costs here are essentially infinite, because to require consent in such a situation would remove the very pleasure you expect to gain, yet there is no question you will be liable for punitive as well as any actual damages no matter how much the value to you (measured in terms of utility or willingness and ability to pay) exceeds the harm to the victim.

Similarly, the liability rules in situations involving private necessity, government takings under the eminent domain power, and continuing private nuisances do not vary depending on whether the situation involves low or high transaction costs, but rather according to the criteria discussed above. For example, in the private nuisance cases, the criteria discussed above⁹⁰ have been applied to deny injunctions (which would have had the effect of requiring the plaintiff's consent for the continuation of the nuisance) even though there were few potential plaintiffs and thus arguably low transaction costs,⁹¹ and, conversely, injunctions have been granted even though there were many potential defendants and thus arguably high transaction costs.⁹² (In each case, there was liability for damages due to the intentional nonconsensual nature of the injury.)

Finally, most intentional torts do not involve liability for punitive damages, which are awarded only for malicious or similar morally egregious behavior. Moreover, when punitive damages are awarded, they are not based on the factors mentioned by the economists, but rather primarily on the moral reprehensibility of the defendant's conduct. To the extent that the ratio of punitive to actual damages is considered, the ratio is not calculated or evaluated based on the probability of detection and successful imposition of liability, but rather on other, non-efficiency-related considerations.⁹³

90 Op cit n 86.

91 See for example, *Boomer*, op cit n 86.

92 See for example, *Morgan v High Penn Oil Co.*, (1953) 77 SE2d 682 (NC); *Estancias*, op cit n 86.

93 See for example, *BMW of North America, Inc. v Gore*, (1996) 517 US 559. The US Supreme Court in *BMW v Gore* overturned the punitive damage award in that case despite a brief from law and economics scholars supporting the award.

D. Tort, Crime, and Punitive Damages

At a higher level, there is a basic distinction in the law between criminal liability and tort liability. Yet there often may be criminal liability as well as tort liability for the same conduct. What distinguishes criminal liability from tort liability, and how can there be both criminal and tort liability for the same conduct?

First, there generally is no liability in tort law, no matter how dangerous or heinous one's conduct, unless that conduct has caused a legally recognized injury to the individual plaintiff. The injury ordinarily must be some actual harm, but also includes certain types of dignitary injuries. Moreover, as was discussed above, there is no tort liability if the plaintiff consented to the injury (or to being exposed to the risk of its occurrence). In contrast, there is liability in criminal law in the absence of causation of a discrete injury to any individual. Many crimes – e.g., treason and mail fraud – are complete regardless of whether anyone suffers a discrete injury. Although other crimes, such as theft, murder, and rape, are not complete unless the defendant has caused a particular type of injury to another, there is criminal liability for attempting to produce such an injury regardless of whether the attempt is successful. Consent of the victim is ordinarily not a defence to criminal liability. In earlier times, criminal and tort liability claims often were not clearly distinguished and were prosecuted in a single action by the individual victim. In modern times, crimes are generally prosecuted by public officials rather than by the victims of the crime (if any) or other private citizens, while tort claims are litigated by the individual victim.

Second, while criminal liability traditionally has been predicated on conduct that is deemed to be morally faulty or blameworthy,⁹⁴ tort liability does not require that the defendant's conduct have been morally blameworthy (except for a possible award of punitive damages). One of the traditional basic elements of a crime is the *mens rea* requirement, which focuses on the subjective state of mind of the criminal defendant. Criminal liability generally is not imposed if the defendant did not have the required culpable state of mind due, e.g., to insanity or other mental incapacity, mistake of fact, or other subjective excuses. Tort law generally refuses to recognize such excuses. Tort law imposes liability on conduct which is only legally (objectively) 'faulty,' rather than morally faulty, and in certain situations imposes 'strict liability' for conduct that was neither morally nor legally faulty.⁹⁵ A social stigma of blameworthiness usually attaches to a criminal conviction, while no such stigma accompanies ordinary tort liability.

Third, the principal remedies in criminal law are sanctions that are designed to punish the defendant. The usual sanctions, beyond the stigma of being declared a criminal, are imprisonment or a fine, or both. Although

94 This traditional focus of criminal law has been weakened considerably by the growing use of criminal law as a regulatory device and the related growth of strict liability crimes. Both of these aspects of modern criminal law have been criticized as inconsistent with and even detrimental to its basic function of demarcating and penalizing socially harmful, morally blameworthy conduct. See for example, Paul H Robinson & John M Darley, *Justice, Liability and Blame: Community Views and the Criminal Law* (Westview Press, Boulder, 1996).

95 See Fleming, *op cit* n 73, 22-24, 76-79, 106-108.

in some jurisdictions some compensation is afforded to the victim (usually quite limited, and from the state rather than from the criminal), this remedy is incidental and secondary to criminal law's primary purpose, which is punishment of the defendant.⁹⁶ In contrast, the primary remedies in tort law are designed not to punish the defendant, but rather to undo (or avoid) the loss that was caused to (or anticipated by) the individual plaintiff as a result of the defendant's tortious conduct. This is usually accomplished by an award of damages, but may also in appropriate cases involve a return of misappropriated property or an injunction to prevent imminent injury. In tort cases involving egregious conduct, 'punitive' or 'exemplary' damages may be awarded.

Fourth, the procedural obstacles to imposing liability are much higher in criminal law than they are in tort law. In a criminal prosecution, the prosecutor must establish the defendant's guilt 'beyond a reasonable doubt.' In contrast, in tort and other civil (noncriminal) cases, the plaintiff need only establish the defendant's liability by a 'preponderance of the evidence' – i.e., to the point where the fact-finder has formed a minimal or 'more likely than not' threshold belief, despite possibly substantial doubts, that the defendant is liable. There are also other procedural hurdles in criminal law that make it more difficult to establish criminal liability as opposed to civil liability: e.g., the ban on double jeopardy, the privilege against self-incrimination, and (in the United States, at least) the unilateral and affirmative obligation of the prosecutor to disclose all relevant evidence to the accused before the trial (as opposed to the mutual and less proactive disclosure obligations of the parties in a civil case).

All of the above distinctions are consistent with and justified by the justice theory. From ancient times, there has been a distinction in the justice theory between the 'private wrong' done to an individual whose discrete interests in person or property have been impaired by another's nonrightful conduct and the (perhaps concurrent) 'public wrong' done to the community or society as a whole by a person who consciously chooses to flout or disregard the law.⁹⁷ Tort law protects individuals' rights in their persons and property against discrete injuries which result from conduct by another that is objectively inconsistent with the equal negative freedom of all. As we noted above, right is defined objectively rather than subjectively.⁹⁸ Thus, unlike the typical crime, the typical tort is a 'wrong' not in the sense of a morally blameworthy deed, but rather in the sense of a harmful infringement of the plaintiff's rights, for which the defendant is *morally responsible* but not necessarily *morally blameworthy*. In contrast, the wrong to the community as a whole consists of the non-discrete injury to the dignity and security of each member of the society, which results from (or is constituted by) the criminal's subjective conscious flouting or reckless disregard of the society's norms of public peace and order, by which he declares himself to

⁹⁶ Ibid, 35-38, 82.

⁹⁷ See Aristotle, *Nicomachean Ethics* III.5 at 1113b23-1114a3, V.11 at 1138a9-14; Kant, op cit n 48, at *331; John Finnis, *Aquinas* 210-211 (OUP, Oxford, 1998); Fleming, op cit n 73 at 35, 82; Wright, op cit n 46, at 175-176.

⁹⁸ Op cit n 55.

be outside the law (literally, an 'outlaw').

The utilitarian efficiency theory has an extremely difficult time explaining the existence of both criminal law and tort law, especially the possibility of both applying to the same conduct, and most especially the possibility of a defendant being subject to both criminal liability and liability for punitive damages in tort law. Under the utilitarian efficiency theory, there is no conception of a private wrong. Rather, all wrongs are 'public,' in the sense that they are always evaluated in terms of the aggregate impact on everyone. Tort law, criminal law, and all other areas of law have the same goal: encouraging efficient conduct and discouraging inefficient conduct. What need or reason, then, is there for both criminal and tort liability, especially for criminal liability in addition to punitive damages in tort law? The economist's ultimate answer is that criminal law and its imprisonment sanction are needed and exist for the sole purpose of deterring the poor, who would not be deterred by potential liability for punitive damages in tort law.⁹⁹ Once again, the lack of descriptive fit (as well as normative appeal) should be obvious. Criminal law often employs fines rather than imprisonment, does not always involve even potential imprisonment, and is not employed only against the poor.

Is there a justification for punitive damages in tort law under the justice theory? When a person injures another through conduct that is objectively but not subjectively inconsistent with the equal negative freedom of the other, there is no basis for punishment or retribution but only for requiring her to see to it that the unjust loss is rectified. However, when her conduct is subjectively as well as objectively inconsistent with the equal negative freedom of the other, and sufficiently so to be considered malicious or otherwise morally egregious, then in addition to any nondignitary loss the victim may have suffered (which itself requires rectification) there also is a significant discrete dignitary injury to the victim, which can and should be rectified through the imposition of private retribution in the form of punitive damages that she must pay to the victim. If, as is often the case, her action also constitutes a conscious flouting or reckless disregard of the rules of public peace or order, then there is an additional nondiscrete injury to the dignity and security of each and every member of the civil society (or to "the state itself"). This nondiscrete injury is rectified by imposing public retribution on her in the form of a criminal sentence, which she herself must satisfy.

These two distinct types of retribution, one a public retribution accomplished through criminal law for the nondiscrete wrong to the public as a whole and the other a private retribution accomplished through tort law for the discrete dignitary injury to the individual plaintiff, have been recognized in the law since ancient times.¹⁰⁰ Both types of retribution are rectificatory or compensatory in nature and fall within the domain of corrective justice.¹⁰¹ Compensatory private retribution for the discrete

99 See for example, Posner, *op cit* n 63, at 205. The utilitarian efficiency theory also cannot explain the significant differences between the substantive and procedural elements of criminal liability and tort liability.

100 *Op cit* n 97.

101 See Wright, *op cit* n 46, at 175.

dignitary injury inflicted on an individual plaintiff is available in both common law and civil law jurisdictions outside the United States – e.g., as ‘aggravated damages’ in England, Australia, and New Zealand¹⁰² and as ‘satisfaction damages’ in a number of countries in Continental Europe (in Germany as a component of pain-and-suffering damages). Either of these terms is preferable to the term ‘punitive’ damages, which has led to confusion of tort damages with criminal-law goals in the United States.

In all but a few of the United States, punitive damages in tort law have come to be viewed solely as a backstop to criminal law, and thus as a ‘windfall’ to the tort plaintiff, who acts merely as a ‘private attorney general’ in redressing the wrong to the public as a whole (rather than any discrete wrong to himself), and who should therefore (it is argued) share much or most of the punitive damage award with the state. This conception gives rise to a host of constitutional and theoretical problems: e.g., why are the much stricter criminal procedure protections not required, why is the double jeopardy bar not applicable, why is there not an explicitly stated maximum penalty, why allow private parties to make end runs around prosecutorial discretion not to prosecute, why allow punitive damages when there has been a successful criminal prosecution, and why not fix any gaps or deficiencies in the criminal law by revising or providing additional resources to the criminal law? The formalist response to all these legitimate objections – that they apply only to actions that are formally ‘criminal’¹⁰³ – rings very hollow to me, if indeed the purpose of punitive damages in tort law is merely as a backstop to the criminal law.

If punitive damages in the United States were reconceptualized, renamed and returned to their proper function as ‘aggravated’ or ‘satisfaction’ damages for the discrete dignitary injury that was inflicted on the particular tort plaintiff, all of the constitutional and theoretical problems would disappear. The damages would be limited to an amount necessary to redress the dignitary injury that was inflicted on the individual plaintiff by the defendant’s morally egregious injurious conduct. Arguments based on the injury or threat to society as a whole, or the number of others also injured as in the *BMW* case, would not be appropriate, at least directly (as a multiplier), but might be relevant to showing the degree of the defendant’s lack of respect for individuals as persons with rights to personal security and autonomy. On the other hand, since each such award would be limited to redress for the discrete dignitary injury that was inflicted upon the particular plaintiff, there would be no theoretical problem (as there is now) with multiple punitive damage awards as a result of claims being made in multiple suits by different plaintiffs.

VII. CONCLUSION

The preceding discussion of punitive damages in the United States hopefully illustrates the manner in which the principled approach can properly and usefully be employed to clarify and render coherent a currently confused

¹⁰² See Fleming, *op cit* n 73, at 241-243

¹⁰³ See for example, *Tuttle v Raymond* (1985) 494 A2d 1353.

area of law. As Smillie's second article¹⁰⁴ and other recent writings¹⁰⁵ make clear, similar confusion has recently arisen in New Zealand with respect to the nature and availability of aggravated and exemplary damages in tort law, a confusion which seems to be due in part to confusion with respect to the basic principles of tort law itself.

I have argued in this article that there is a principled basis for tort law – the justice theory based on the foundational norm of equal freedom – that coherently explains and justifies the major aspects of tort liability. According to the justice theory, the basic goal of tort law is not and never has been efficient deterrence, efficient compensation (spreading of all losses), retribution or deterrence of public wrongs, or the identification of morally faulty injurers, but rather the prevention and rectification of injuries to individuals' persons and property that are the result of nonrightful conduct by others, which is conduct that is *objectively* inconsistent with the equal freedom of all. As such, tort liability is based on individual moral *responsibility*, rather than moral fault or blame, for having injured another.

Several writers have expressed puzzlement over the nature and principled basis of aggravated damages in tort law. However, there is a firm principled basis for such damages, once they are properly understood. Ancient, classical, and modern law and philosophy have recognized a private rectificatory action, retributive in nature, for redress of the discrete dignitary injury to the plaintiff that occurs when the defendant inflicts injury on the plaintiff through morally egregious behaviour that is subjectively as well as objectively inconsistent with the plaintiff's right to equal freedom. Aggravated damages perform this function in New Zealand. They are compensatory but also punitive in nature. However, they are not punitive in the same sense as a criminal sanction or as exemplary damages, both of which are meant to be retribution for a public rather than a private wrong. Rather they are punitive in the sense of being private retribution for the discrete dignitary injury that was inflicted on the plaintiff by the defendant. Rectification of this dignitary injury to the plaintiff can only occur through such private retribution, which takes into account the egregiousness of the defendant's behaviour as well as the impact of that behaviour on the plaintiff, just as the rectification of the wrong to the public can only occur through public retribution that takes into account the egregiousness of the criminal defendant's conduct as well as the impact of that conduct on the public as a whole. The dignitary injury can roughly be described, as it has been, as an injury to the plaintiff's feelings, but it is distinct from mental or physical pain and suffering as those items have traditionally been understood.

I agree that exemplary damages do not belong in tort law, insofar as they are aimed at retribution or deterrence of a nondiscrete wrong to the public as a whole, rather than retribution of a discrete dignitary wrong to the plaintiff or removal of unjust gains that were obtained through a wrong to the plaintiff. As so conceived, exemplary damages belong in criminal law, as criminal fines, with all the appropriate substantive and procedural restrictions that exist in criminal law. However, if they are aimed at retribution of a discrete dignitary wrong to the plaintiff – i.e., are actually

¹⁰⁴ See Smillie, *op cit* n 2.

¹⁰⁵ For example, Stephen Todd et al, *The Law of Torts in New Zealand* (Brooker's, Wellington, 1997), pp14-17, 1225-1226, 1229-1237; Fleming, *op cit* n 73, at 241-243.

aggravated damages – I see no basis in the tort cases or in the justice principles underlying tort law for limiting them to the intentional torts. After all, *Wilkinson v Downton*,¹⁰⁶ which Smillie cites,¹⁰⁷ was not strictly an intentional tort, since there was no purpose or knowledge of a substantial certainty of causing severe emotional distress, but rather only a reckless disregard of a substantial possibility (perhaps a likelihood).¹⁰⁸ As in *Wilkinson*, conduct which falls short of being malicious or even intentional can constitute contumelious or conscious disregard for the plaintiff's legal rights and thus support an award of aggravated damages.

Yet, as Smillie and other writers have pointed out, recent New Zealand cases have awarded exemplary damages in tort cases where there was no such contumelious or conscious disregard for the plaintiff's legal rights. Moreover, the 'exemplary' damages in these cases seem to have been designed to compensate the plaintiff for her injury. Yet tort actions for compensatory damages related to personal injury, including aggravated damages, have been deemed to be barred since the enactment of the New Zealand Accident Compensation Act (although it perhaps could have been decided that the legislature did not have in mind and did not mean to preempt tort actions for aggravated damages, despite their compensatory nature, given their distinct retributive function). Under the principled approach to adjudication, according to which a judge applying the law must adhere to clear requirements of the positive law even when they are contrary to the underlying principles of a relevant body of law,¹⁰⁹ these decisions are improper.

Nevertheless everyone seems to have sympathy for the concerns that drove these decisions, although not for their results. I certainly do. The New Zealand Accident Compensation Scheme from the beginning lacked a coherent principled foundation. It overthrew the tort system for personal injuries, perhaps under the mistaken impression that it was merely a (very inefficient and incomplete) loss compensation scheme or a fictitious morality play. Yet it did not replace it with a comprehensive efficient-compensation scheme, much less a (normatively more attractive) distributive-justice scheme for assuring minimal or moderate levels of needed resources. It threw out not only whatever (just or efficient) deterrence is attributable to the tort system, but also and more importantly the principle of individual responsibility. In my travels around New Zealand a few years ago, I often heard complaints, on the one hand, that people were not being held liable for their irresponsible behavior that injured others and, on the other hand, that people engaging in irresponsible behavior (e.g., drunken driving) were able to collect from the compensation scheme.

Some of the 1992 amendments to the compensation scheme were motivated not only by the nemesis of all such schemes – the desire on the part of government, employers, and industry to reduce their financial burdens under the scheme – but also by complaints by employers that they were being made to pay for injuries for which they were not responsible and, conversely, concerns by others that medical practitioners were not being made to pay for injuries for which they were responsible. The 1992

106 [1897] 2 QB 57.

107 See Smillie, op cit n 2, at 141, 155.

108 See Fleming, op cit n 73, at 32-33.

109 Op cit n 41; see Dworkin, op cit n 22, at 169.

amendments are generally acknowledged to have produced even more of a tangle of incoherent principles than before, as well as even more serious gaps and anomalies in the compensation scheme and the law of torts.¹¹⁰

In sum, it is important not only for judges to take a principled approach to adjudication, but also for legislators to take a principled approach to legislation, which should include taking seriously the justice values deeply rooted in the law in general and tort law in particular when drafting legislation dealing with responsibility for personal injuries.

¹¹⁰ See for example, Richard S Miller, "An Analysis and Critique of the 1992 Changes to New Zealand's Accident Compensation Scheme" (1993) 52 *Maryland LR* 1070.