RATIONAL CONTRACTARIANISM, CORRECTIVE JUSTICE, AND TORT LAW

RICHARD J. ARNESON*

Jules Coleman's new book revisits topics that his writings have addressed for several years and to which he has made important contributions. In *Risks and Wrongs*,[1] Coleman writes with sustained originality and vigor across a broad range of issues of central importance to political theory and legal philosophy. My aim is to celebrate his achievement by airing a few criticisms.

I. RATIONAL CONTRACTARIANISM

Rational choice theories have been enjoying a vogue in the social sciences in recent years and have been applied to a wide variety of social phenomena beyond the economic marketplace. In light of their power and elegance, and their significance in the social sciences, the question arises whether these theories can provide a foundational doctrine for normative moral and political theory. The view that Coleman calls "rational contractarianism" pursues an affirmative answer to this question.

The contractarian borrows from current economics a conception of rationality in the thin sense. According to this conception, the rational agent is one who makes the most effective use of available means in order to maximize the fulfillment of whatever goals she happens to have. Under conditions of certainty about to what outcome would result from any action she might take, this view stipulates that the rational agent is one who maximizes utility; under conditions of risk, in which the agent can identify, for every action she might take, each possible outcome and the probability of achieving it if she chooses the action, the rational agent is stipulated to be one who maximizes her expected utility. Branches of economic theory such as noncooperative game theory seek to render the idea of rationality in the thin sense more precise and to extend it to conditions of strategic interaction. The contractarian assumes that the notion of thin rationality (as a correct economic theory will

* Professor of Philosophy, University of California, San Diego.

1. JULES L. COLEMAN, RISKS AND WRONGS (forthcoming 1992) (manuscript dated July 1991, on file with author; pages cited to manuscript).

ultimately conceive it) is exhaustive of the content of rationality; thus, to demonstrate the rationality of moral rules or the rules of institutions is simply to demonstrate that an individual motivated by a thin notion of rationality would accept such rules.

"Thick" rationality, in contrast, refers to what is rational for an individual to do, taking into account all the reasons for action that bear on her choice. Rationality in the thick sense may include thin rationality together with constraints on what it is permissible to do to achieve one's goals: mandatory goals that one must pursue irrespective of whether they are included in one's own set of chosen goals and various criteria of fitness to be imposed on potential goals.

The contractarian makes the strong and highly controversial assertion that thick rationality collapses into thin rationality. I find the assertion implausible. It certainly requires a defense beyond the bland citation of the fact that thin, or economic, rationality is central to many contemporary efforts in empirical social science. Coleman himself seems more interested in describing and exploring rational contractarianism than in justifying it, so perhaps he does not owe his readers a full-scale defense of the doctrine.

The controversiality of the economic conception of rationality, taken as exhaustive of the constraints of reason, is perhaps obscured by the fact that the economic conception of an agent's interest is purely formal. An agent's interests are determined by her preferences, which in turn are a theoretical construction from her choice behavior. To say that an agent prefers X to Y is to say that she would choose X over Y if she were to choose between them. If an agent acts consistently to further others' interests at some cost to her own, her acts are not formally inconsistent with utility-maximizing behavior, because her choice to further others' interests partially constitutes her utility. More highly elaborated empirical and normative theories, however, give the economically rational conception of an individual's utility or interests a more determinate content. For example, the economic theory of consumer behavior assumes that each consumer maximizes her utility from her own consumption and leisure; likewise, the economic theory of the firm stipulates that firms maximize profits.

It is useful to start our inquiry with a common-sense re-

stricted idea of a person's interests and consider how we have reasons that are not derived from our own interests for furthering the interests of others. Our capacity to accept as binding on our conduct reasons for action that do not appeal to our interests arises from our capacity to transcend our own standpoint in thought, and to view our own interests, goals, and preferences from a distance. From this impartial standpoint, I am one person among others, and my own concerns matter no more than anyone else's equally-pressing concerns.[2]

The impartial standpoint imposes a strong symmetry requirement on justifications of personal actions, principles, and institutional arrangements. In ordinary life, people appeal to this symmetry requirement when debating the merits of social arrangements in terms of their fairness or justice. The exact nature of this symmetry requirement is obscure and controversial, but my claim is simply that it excludes egoism. The impartial standpoint claims authority over my conception of practical reason, but this authority cannot—even on an extreme utilitarian account—simply take the form of insistence that I should give no favor to myself in deciding how to act. As a human animal I am psychologically disposed to favoritism: To oppose this disposition directly and uncompromisingly is to play King Canute against the tide. Some complex compromise is necessary.[3]

One should allow for the human tendency to partiality when considering the requirements of impartiality. The capacity to occupy in thought an impartial standpoint, however, rules out as a serious possibility the claim that justification could be made solely in terms of my self-regarding interests. The best political arrangement from my standpoint might be everyone else's acceptance of the status of slave to me. Yet being maximally conducive to whatever is best for me cannot count as a justification, because any reasonable conception of justification is bound up with a symmetry requirement that the "slavery of everyone except me" principle flagrantly violates.

The rational contractarian approach does not eschew symmetry and the associated idea of fairness but rather tries to de-

2. This impartial standpoint and its consequences for ethics have been explored in writings of Thomas Nagel. *See* THOMAS NAGEL, EQUALITY AND PARTIALITY (1991); THOMAS NAGEL, THE POSSIBILITY OF ALTRUISM (1970); THOMAS NAGEL, THE VIEW FROM NOWHERE (1986).

3. *See, e.g.*, SAMUEL SCHEFFLER, THE REJECTION OF CONSEQUENTIALISM (1982).

rive it from the description of the arena in which a rational agent seeks to achieve her ends. In the contractarian construction, the rational agent exists in a world populated by other equally rational agents, each of whom seeks to satisfy her own interests. In any description of this hypothetical world that remotely resembles the actual world, these equally rational parties will be unequally endowed with resources and will enjoy unequally favorable opportunities for predation and for retaining the fruits of predation. No set of constraints that may be rational for such parties to accept could be symmetrical.[4]

It is even more certain that if one realistically postulates agents endowed with varying degrees of rationality, any "moral" constraints that a rational agent would accept in order to advance her interests indirectly in such a world would not resemble constraints of morality. The constraints would look like a differentiated ethic: Play square with those smart enough and tough enough to reciprocate; exploit those who are not.

The rational contractarian might argue that calculations of individual advantage that adequately register the world's uncertainty will induce the rational individual to accept impartial principles that divide the benefits of cooperation symmetrically. Should I insist on the lion's share of benefits when such insistence appears to be to my advantage? Perhaps appearances are deceiving. The weak and feebleminded, whom I like to imagine I may profitably exploit, might have strong allies. The weak might in time become stronger. Those who seem weak now may be trying to deceive me for complex strategic reasons, or perhaps my eyes are deficient. Symmetry or fairness should thus be understood, according to the rational contractarian, as a hedge against uncertainty. This line of thought is unpersuasive.

Sometimes we do possess the information required to make expected utility calculations. Even when uncertainty is not eliminable from practical deliberations, its presence requires that we ignore all but the best and worst possible outcomes of actions we might take, but does not require singleminded fixation on the worst possible outcomes. It does not counsel a play-it-safe strategy on the part of those who perceive themselves to be well placed to exploit and dominate others. As an empirical

4. *See generally* DAVID GAUTHIER, MORALS BY AGREEMENT (1986); *see also* Jody S. Kraus & Jules Coleman, *Morality and the Theory of Rational Choice*, 97 ETHICS 715-49 (1987).

matter, uncertainty is a sometime thing. The contractarian who introduced an assumption of strong uncertainty about bargaining strength into rational contractarian argument without empirical warrant would be introducing a Rawlsian perspective by the back door.

The ordinary moralist reads symmetrical content into morality by accepting such symmetry directly on the ground that impartiality requires it. The rational contractarian seeks to derive much the same symmetrical content from a construction that posits ideally symmetrical agents, and then asks what constraints each would find it in her interest to accept. This strategy cannot work because these idealizing assumptions have no empirical warrant. Gauthier announces that he hopes to show that "under actual, or realistically possible, conditions, moral constraints are rational."[5]

Yet as we move toward greater realism by relaxing the idealizing assumptions such as the equal rationality of agents, we decisively break the link between self-interest and the constrained maximization morality of the rational contractarian. The rational contractarian, of course, is free to pursue the project of determining whether disposing oneself to accept constraints on utility-maximizing behavior is a rational strategy for agents facing the real world. The project would lose contact with ethics as ordinarily conceived.

I want to mention an aspect of the "symmetry" to which one is perhaps driven by the impartial perspective and that a sophisticated rational contractarianism is unlikely to embrace. Our intuition is that other things being equal, it is morally more desirable to bring about a gain or avoid a loss for someone who is worse off than for someone who is better off. This view is given a famous and somewhat extreme expression in Rawls's difference principle: Institutions should be arranged so as to maximize the resource holdings of the worst-off representative individual.[6]

One might explain this view as reflecting the tendency of resources to yield diminishing marginal utility as more are acquired, though this tendency is weak and flickering. At any rate, Rawls asserts that it is inherently a matter of greater moral urgency to help the worse off, quite apart from the possibility that

5. GAUTHIER, *supra* note 4, at 174.
6. *See* JOHN RAWLS, A THEORY OF JUSTICE 75-78 (1971).

channeling benefits in this way may yield increases in some other value. This general idea, which is more plausible than its expression in the stringent difference principle, is that the moral value of achieving a marginal benefit for a person is greater, the lower that person's benefit level was prior to gaining this benefit.

II. Coleman on Rational Contractarianism

Coleman rejects what he calls the "market paradigm," but as he notes rejecting the market paradigm does not entail rejecting rational contractarianism. I take it that Coleman is officially agnostic on the prospects of rational contractarian moral and political theory. The bulk of his book develops an interpretation of tort law in terms of corrective justice, which he does not show to be either justified or unjustified in rational contractarian terms. The discussion above suggests that Coleman's agnosticism is overly optimistic about the prospects of the rational contractarian program.

Coleman makes interesting but sketchy remarks about possible rational contractarian but non-market-paradigm reasons for favoring economic markets:

> Markets are most attractive when individuals have broadly divergent conceptions of the good, where the relationships among individuals tend to be one-dimensional, discrete, non-repeating, and where the benefits and burdens of cooperation are spread out over persons, time, and geography. They are least necessary where interaction is repeated, where relationships are multi-dimensional and direct, and where there are shared conceptions of the good.[7]

There is a chicken-and-egg problem here. Markets facilitate cooperative relationships among individuals that are one-dimensional, discrete, and non-repeating, and that are expressed in transactions at a distance. That is part of the point of having markets. In societies where market interaction and weak ties are pervasive, individuals' conceptions of the good tend to diffuse; the diverse life experiences facilitated by the broad menu of goods for sale on the market promote diverse views of the good. The conditions under which markets are attractive, according to Coleman, are conditions that the operation of markets tends to produce. By contrast, in a society organized as

7. Coleman, *supra* note 1, at 7.

one big crusade, we might speculate with Coleman that people who agree on the good and engage in multi-dimensional, diffuse, and repeating cooperative relationships tend to eschew markets and find crusade organization attractive. An alternative hypothesis is that over time the society organized as a crusade will tend to suppress diversity and promote agreement on the good and close-knit interaction.

Coleman posits that the benefit of markets may be their tendency to promote social stability under the conditions he cites. My sense is that we should value markets highly to the extent that we place a high value on diversity of the good among citizens and on the other conditions he cites, which markets tend to promote. We value not just order but good order, and we associate good order with markets.

III. Efficiency and Equity in Tort Law

Why might one want to assess the adequacy of tort law solely according to the degree to which it achieves Kaldor-Hicks efficiency? A contractarian might urge that rational persons would unanimously agree *ex ante* that tort disputes should be resolved on the basis of this efficiency norm. Persons who do not know whether Kaldor-Hicks will select them as gainers or losers in their future disputes, and who suppose that they are equally likely to be gainers or losers, will maximize expected utility by embracing Kaldor-Hicks.

Why should this consideration be decisive for a person who strives to maximize her expected utility under *actual* circumstances? I may know that I am a farmer adjacent to a railway, not a railroad owner. In the disputes I expect to face, I will consistently turn out to be either a loser whose losses hypothetically could be compensated or a gainer whose gains are insufficient for hypothetical compensation of losers. Under Kaldor-Hicks, I will lose my suits. It is not clear why determining what I would have agreed to *ex ante* under ignorance should move me any more than I should be moved to comply with rules that are justified on the basis of their rationality given thoroughgoing ignorance in a Rawlsian original position. The appeal to what is *ex ante* rational blurs a line the distinctness of which the contractarian should sustain.

Leaving contractarianism to the side, we might consider whether the economic analysis of law provides reasons to as-

sess tort law entirely by Kaldor-Hicks efficiency. Using Kaldor-Hicks efficiency as a paramount norm to assess legal policy implicitly assumes that a monetary scale can serve as a basis for interpersonal comparison. If a proposed change in a three-person society would confer on Jones a benefit that she values at $4, a loss on Smith that she values at −$2, and a loss on Black that he values at −$1, the change is Kaldor-Hicks efficient. Claiming that this efficiency gain warrants implementing the change thus implies: (a) In some morally significant sense Jones's gain outweighs the losses of the others; and (b) this is the only relevant policy consideration. I take it as apparent that (b) cannot be defended. Consider, then, the proposition of (a).

One's willingness to pay depends *inter alia* on one's available wealth. Smith might be willing to pay more for a change than Jones would pay to prevent it, even though the change produces less utility gain for Smith than utility loss for Jones. Alternatively, Smith may be a poor transformer of resources into utility, so that a dollar gain to Smith would produce less utility than a dollar gain to Jones. Further, for egalitarian reasons, we might regard bringing about a small utility gain for Jones, whose utility level is low, as more morally urgent than bringing about an identical utility gain for Smith, who is already thriving with high utility. Another possibility is that from a moral standpoint we are disinterested in utility, construed as a measure that uses each individual's own value judgments about the relative importance of benefits to rank their relative importance for her. We are concerned, instead, that individuals are enabled to be and to do with their resources what is objectively valuable on some scale that is independent of individual preferences and valuations.[8]

In the terminology of Amartya Sen, the thought is that we should be concerned with the set of valuable functionings that resources make available to individuals. There is no reason to suppose that a two-dollar gain to Smith will achieve a greater increase in valuable functionings than a one dollar gain to Jones because individuals differ in their capacities to gain from money.

The points made above presuppose the possibility of interpersonal comparisons either of utility or of well-being on an

8. *See generally* Amartya Sen, *Well-being, Agency and Freedom: The Dewey Lectures 1984*, 82 J. PHIL. 185-203 (1985).

objective scale. If interpersonal comparisons cannot be made for purposes of assessing legal policy, Kaldor-Hicks judgments are not thereby enhanced as if they may somehow serve as surrogates for interpersonal comparisons. Willingness to pay can be significant only as a proxy for some other measure of comparison, and there is no good reason to suppose that it is a generally good proxy for any sensible measure of individual advantage. If interpersonal comparisons cannot be made, then Kaldor-Hicks is a proxy for nothing.

The economic analyst of torts usually concedes that in principle, efficiency may conflict with fairness values, and that nothing in the economics of the situation precludes a decision by the tort theorist to give priority to fairness over efficiency. In practice, the argument goes, the concession gives up very little. Take the position of a judge who is deciding a tort law case that will become a precedent for future decisions. There are two types of case to consider. Either the plaintiff and the defendant are involved in a contractual relationship, or they are strangers. In the former type of case, the judge has very limited power to change by her decision the distribution of benefits and burdens between potential plaintiffs and potential defendants. If the judge finds for the defendant manufacturer in a products liability case, potential customers deciding whether to purchase the product will take into account the reduced liability of manufacturers resulting from the judge's decision and adjust the price they are willing to pay for the product accordingly. There is no *a priori* reason to think that the manufacturers of the product will be better off if the liability decision goes their way rather than against them.

On the other hand, if the plaintiff and defendant are not involved in a contractual relationship, then the decision of liability can very well shift the distribution of benefits and burdens as between the class of potential defendants and potential plaintiffs. If someone driving a car strikes a pedestrian and the injured pedestrian sues, a decision in favor of the suing pedestrian that becomes precedent renders pedestrians better off in relation to motorists. But the economic analyst of torts, acknowledging that distributional shifts can be imposed in cases of this type, wonders why they should be. The class of motorists and the class of pedestrians are likely to be heterogeneous with respect to classifications that matter for fairness values.

Perhaps motorists are on average slightly wealthier than pedestrians, but tort liability for automobile accidents is an unwieldy and imprecise method for shifting wealth from the wealthier to the worse off. In modern market economies there already exists a more effective institution, the progressive income tax, which is available as a tool to redistribute wealth. The economic analyst of torts thus concludes that for a mix of normative and positive reasons, tort law ought to be concerned with efficiency alone.

As it stands, the economic argument fails. First, "efficiency" in the argument above must refer to Kaldor-Hicks efficiency, but Pareto efficiency is the only efficiency concept that has clear normative significance. The fact that a particular tort law change would be Kaldor-Hicks efficient would have clear normative significance only if it could be argued successfully that in the long run a policy of implementing Kaldor-Hicks efficient changes can be expected to be to everyone's advantage.[9] Such an argument could only show, not that Kaldor-Hicks efficiency is intrinsically morally valuable, but that its pursuit would be a means to something else that is intrinsically morally valuable.

Another weakness in the argument for treating Kaldor-Hicks efficiency as the sole standard for assessing tort law policies is that it rests on a false identification of "equity" with "the equitable distribution of income." The institutional economist proposes that institutions should be assessed for efficiency—with the sole exception of the income tax system, which can be charged with balancing the values of "efficiency" versus "equity" by using tax policy to adjust the distribution of income.[10]

As it stands the claim begs the question against someone who thinks that there are special equity standards—such as the standards of corrective justice—that are central to tort law, and most effectively implemented only through tort law, or most effectively through tort law. Coleman's book powerfully develops this point.

9. For an argument along this line, see A. Mitchell Polinsky, *Probabilistic Compensation Criteria*, 86 Q. J. Econ. 407 (1972).

10. *See* Steven Shavell, *A Note on Efficiency versus Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Distribution?*, 71 Am. Econ. Rev. 414 (1981).

IV. Coleman on Corrective Justice

I expected Coleman to move from the denial of Kaldor-Hicks efficiency as the appropriate normative standard to the further claim that tort law should be assessed according to the standards of corrective justice, and that conformity to the most defensible standards of corrective justice is at least a necessary condition for the acceptability of tort practice. Coleman's actual view of the relationship between corrective justice and tort law is trickier to discern.

Coleman holds that the best interpretation of a substantial fraction of tort law is that this body of law satisfies corrective justice. He also holds that another fraction of tort law is not consistent with norms of corrective justice, but that inconsistency with corrective justice is not necessarily a fault: Other values may trump the values of corrective justice. Moreover, even if the legal theorist confines her attention to those areas of tort law that are best interpreted as aiming to achieve corrective justice, she will notice that corrective justice does not really demand that tort law should conform to it. An example Coleman gives is that, without violating justice, a society could replace tort law accident compensation with a no-fault social insurance plan. Such a change in law might or might not prove efficient, but even if inefficient it need not be unjust. In this imagined no-fault world, when my actions impose wrongful losses on another and hence give rise to a duty to repair founded on corrective justice norms, this duty would be preemptively satisfied by the operation of society's no-fault insurance scheme, much in the same way that (in the actual world) my duty of reparation might be preempted by the payment of appropriate reparation on my behalf by some benefactor.

The norms of corrective justice permit society to establish a tort law compensation scheme but do not morally require that establishment. The explanation of this slack between what corrective justice is and what tort law should be appears to be that the norms of corrective justice are themselves optional for a society. In this connection, Coleman remarks that "whether or not corrective justice imposes *moral* reasons for acting will depend on prevailing legal and social practices."[11]

This sounds disquieting to those of us who have been taught

11. Coleman, *supra* note 1, at 661.

by John Rawls that "[j]ustice is the first virtue of social institutions" and that "laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust."[12]

Coleman's remark may be less subversive than it sounds. He may be saying that considerations of corrective justice may be overridden by other values of justice. It may be, also, that only a part of corrective justice is optional. Sometimes Coleman's view seems to be that society is at liberty to sever the liability connection between injurer and injured and assign liability on a different basis for good reasons provided that the wrongful losses suffered by injured parties are repaired. A shadow of the discarded annulment view of corrective justice is still visible in the mixed view that he now proposes.

The mixed view of corrective justice holds that wrongful losses should be corrected and that those who are responsible for wrongful losses should correct them. Wrongful losses fall into two categories: violations of rights and wrongful causation of loss to legitimate interests. If Coleman adheres to the mixed view, one wonders why he holds that corrective justice is more insistent on bringing about reparation for the losses of victims who suffer wrongful loss than on bringing about payment of reparation for wrongful losses by those who caused them. Coleman mentions that if I have caused wrongful losses and a benefactor pays reparation on my behalf, my obligations in corrective justice are satisfied. This, however, does not establish asymmetry between the two sides of the mixed view. After all, my obligations of reparation would equally be satisfied if the persons on whom I had imposed wrongful losses were voluntarily to forgive all my obligations arising from those losses. The possibilities that someone might act on my behalf to fulfill an obligation that I owe or that someone to whom I owe an obligation might voluntarily waive it do not somehow render the obligation optional. Coleman seems to be supposing that the insurance payments to injured persons under a regime of no-fault accident coverage would be relevantly similar to the voluntary fulfillment of my corrective justice obligations by a benefactor. This supposition ignores the coercive aspect of no-fault. If insurance coverage is mandatory, then persons with in-

12. RAWLS, *supra* note 6, at 3.

surance coverage whose behavior is nonfaulty and who would not purchase insurance at the rates offered in the absence of the legal regime of no-fault are effectively forced to pay compensation to those accident victims whose losses are properly the responsibility of negligent injurers according to corrective justice norms. This state of affairs may be morally acceptable but clearly violates corrective justice.

Coleman does not clarify the reasons for his belief that annulment of wrongful losses by compensation to their victims is a less dispensable element of justice than the requirement that the creators of wrongful losses should make them good by paying compensation to their victims. After all, if the rationale of no-fault alternatives to tort remedies is that in some circumstances no-fault will lower the combined cost of accidents and accident avoidance and accident compensation administration, this same rationale might in some circumstances justify a no-fault no-compensation system under which the losses of injured parties are not compensated. Circumstances can be imagined in which everyone's expected utility *ex ante* is higher under no-fault no-compensation than under feasible alternatives. If there is a special injustice that attaches to failure to compensate injured victims, Coleman does not articulate its basis.

Coleman expounds corrective justice norms but refrains from endorsing them. There are good reasons for his hesitation. The corrective justice norms that could plausibly be thought to support existing tort practices are at most very crude proxies for acceptable principles of justice, which do not themselves plausibly support existing tort practices. If existing tort practices are justifiable, they are so to the extent that they promote efficiency in the less controversial Pareto sense, or efficiency in the more controversial Kaldor-Hicks sense, or utility in the utilitarian sense. Coleman is thus caught in a bind. If he defends corrective justice norms, he defends the indefensible. If he declines to defend corrective justice norms, then he has no normative ground for rejecting the case for current tort law practice that consists of efficiency-driven arguments drawn from the economic analysis of law. On the whole, Coleman tends to shy away from defense of corrective justice norms; his normative critique of the economic analysis of law is radically incomplete and, in my judgment, cannot be completed.

It will prove useful to confine our attention to a single important strand in corrective justice principles. Consider the portion of tort law that is governed by negligence rules, according to which one is liable for the costs of an accident if the plaintiff can show that she suffered a loss, that one's conduct caused the loss foreseeably or proximately, and that one's harmful conduct was faulty. For simplicity, let us suppose we are dealing with cases in which there is no contributory negligence on the part of the injured plaintiff. Does this strand of tort law conform to acceptable standards of corrective justice? In this connection, Coleman has appealed to a principle of "weak retributive justice," asserted by Joel Feinberg, which holds that "if a loss must fall on either of two parties, one of whom is at fault in causing it and the other of whom is faultless, the party at fault ought to bear the loss, all other things being equal."[13] Feinberg's principle has been nicknamed "Fault Forfeits First." This principle is insufficient to rationalize tort law practice.

The standard of fault that is in play in the determination of tort law negligence is an objective standard of due care. The issue is whether one behaved as a reasonably watchful competent person would have behaved under similar circumstances. Whether one was capable of meeting the objective standard of due care, or was blameworthy for failure to meet it, is not relevant to the determination of negligence. Suppose that my clumsiness causes you to slip and break your leg. My conduct is faulty, but my clumsiness may not be attributable to me as my fault. Perhaps I tried to the best of my ability but an inherited lack of skill caused me to slip. We might then wonder whether I am at fault for placing myself where my innate clumsiness could cause harm to others; but let us stipulate that I suffer from poor judgment that I have done my best to overcome without notable success. In this example my conduct is objectively faulty but not morally faulty. I am not blameworthy. The principle "Fault Forfeits First" is plausible only if "fault" is glossed as "moral fault," but the principle is consistent with tort law practice only if "fault" is glossed as "objective fault."

The second and more important reason for denying that "Fault Forfeits First" can justify tort law practice is that the principle assumes that the number of parties who might be

13. *See* JULES COLEMAN, MARKETS, MORALS, AND THE LAW 181 (1988) (discussing JOEL FEINBERG, *Sua Culpa, in* DOING AND DESERVING (1970)).

held liable for a loss has been reduced to two—a faulty causer of harm and another—but does not defend that assumption. In fact, the assumption does not generally hold. Nothing in principle prevents society from shifting the costs of an accident from the injured party, not to the injurer, but to some third party such as the party (if different from injurer) who could have acted at lowest cost to prevent the accident. The loss might also be shifted to an at-fault insurance pool composed of all those who behaved in a faulty way in comparable situations, each faulty agent's contribution to the pool being proportional to the degree of faultiness of his conduct, and the size of the pool being set to equal the total costs of accidents of that type within some time slice. Or there might be numerous other possibilities. "Fault Forfeits First" does not condemn any of these alternative assignments of liability for the costs of accidents. In tort law practice, injurers confront in court the parties they are alleged to have injured, and the choice in assigning liability is limited to these two parties. Once tort law practices are in place and viewed as setting the limits of possible assignments of liability, the "Fault Forfeits First" principle might be invoked to justify holding faulty injurers liable for costs. It is obvious that such applications of the "Fault Forfeits First" principle *assume* the moral acceptability of tort law practices that limit the scope of possible assignments of liability. The principle cannot then be invoked to justify these tort law practices.

So far I have called attention to a principle of corrective justice that might be claimed to justify tort law practices and noted that the principle cannot do this work of justification. An extension of this line of thought shows that these practices are in fact unjustifiable from the standpoint of corrective justice.

Consider the Proportional Fault principle: Those who are at fault should be made to pay the costs of accidents caused (foreseeably) by their faulty conduct, in proportion to its faultiness. This principle implies that if two drivers are equally at fault in driving drunk, and one causes an accident involving $50 damage and the other causes an accident involving $5000 damage, they should pay equal amounts of compensation toward the parties who suffer the damage. The principle also implies that if two drivers are equally at fault, and one happens to cause an accident while the other does not, both should be made to contribute equally toward the cost of the accident.

Judith Thomson has in effect observed that the principle also implies that regardless of the type of conduct that you engage in, and regardless of the total amount of harm that conduct of that type causes in a given time period, you along with all other faulty agents should pay in proportion to fault into a pool that pays the costs of accidents caused by fault.[14]

Drunk drivers, careless starters of fires, inconsiderate spreaders of disease, and so on would under a regime of Proportional Fault contribute according to fault to a gigantic at-fault pool from which accident victims would be compensated. Liability is severed from causation of injury under this scheme, but the degree of faultiness of any individual's conduct is determined by the expected harm to others of that conduct balanced against its expected gains.

The idea of an at-fault pool is subject to many objections. Its administrative cost might be prohibitively high. In practice it might turn out to be a poor way to deter faulty actions likely to cause harm. Our inquiry is limited to the issue whether the principle of Proportional Fault associated with the at-fault pool is a plausible candidate to be a principle of corrective justice. Notice that, if one brackets the issue of causation, the "Fault Forfeits First" principle is a special case of Proportional Fault—the case in which two individuals are involved, one of whom is faultless and the other faulty. Proportional Fault rules that the party at fault should pay all costs caused by faulty conduct in this limiting case.

Some might regard the at-fault pool as a *reductio ad absurdum* of the proposal that one should be liable for the costs of accidents because one's conduct was faulty, regardless of whether one's conduct caused any harm. If I drive drunk in a remote area of Alaska and by luck injure no one, why is it fair to tag my faulty conduct and require me to pay into a pool that pays for the costs of unrelated accidents in other places? The intuition behind the at-fault pool idea is that justice should correct for the effects of luck, or at least for the effects of luck that are not mediated by individual voluntary choice. I rest my view of corrective justice on this intuition.

Exploring the reasons that might have led Coleman to avoid asserting that the idea of corrective justice that is expressed in

14. *See* Judith J. Thomson, *Remarks on Causation and Liability*, 13 PHIL. & PUB. AFF. 104 (1984).

tort law practices is morally mandatory has led to the thought that a different idea of corrective justice, the Proportional Fault idea that is completely at odds with current tort law practices, is morally mandatory. Justice demands the at-fault pool. But this line of thinking is also unstable.

When the at-fault pool is functioning properly, all wrongful losses are repaired, but injurers do not compensate their actual victims. But what does it mean for an at-fault pool to function properly? We may be drawn toward the idea of the at-fault pool by reflecting that luck determines the actual amount of injury, ranging upwards from zero, which is caused by any given faulty act that imposes risk of wrongful harm. Rather than require injurers to make full compensation to victims, we might instead identify all faulty agents and fine them in proportion to the faultiness of their conduct, under the constraint that total fines must equal total compensation adequate to satisfy all valid claims. Any limited at-fault pool, however, will fail to eliminate the impact of luck.[15]

The faulty in France in a given year may be lucky, and cause few wrongful losses, compared to the faulty in Tibet or Uruguay. Perhaps the at-fault pool should function world-wide in order to smooth out the effects of chance. Chance, of course, also operates over time. The amount of harm caused by faulty conduct doubtless varies randomly from year to year and may well vary nonrandomly: Perhaps with population growth and technological advance, the avoidance of behavior that imposes undue risk on others becomes progressively more difficult, so that identically faulty action becomes ever more dangerous.

The problem of defining the limits of moral luck may be illustrated in another way. Imagine two equally faulty drunk drivers who act in ways that give rise to different amounts of harm. It may be the case that within the set of persons equally disposed to drink and drive that day, only some were tempted past their limit and actually became drunk drivers. Examining the set of persons disposed to develop a drinking problem that

15. *See* THOMAS NAGEL, *Moral Luck*, *in* MORTAL QUESTIONS 24-38 (1979). Moral luck is the idea that the moral quality of an agent's choice, the amount of praise or blame to which she is properly liable owing to that choice, can vary depending on factors that lie beyond the agent's power to control. According to Nagel, the idea of moral luck is both incoherent and inescapable. I view the inescapable aspect of moral luck as a forward-looking consideration: Treating people as though they were morally responsible for matters that do not lie within their control can be productive of good consequences.

would culminate in a disposition to drink and drive, we find that some of these people enjoyed moral luck and through no merit of their own never found themselves in circumstances that triggered such a disposition. As Rawls comments in a different but related context, "The idea of rewarding desert is impracticable."[16]

So must be the project of apportioning liability to fault—at least when fault is understood in a morally interesting sense, such that what is my fault must have been within my power to control. It is surely not within my power to control the type of person I am, at some basic level, because my character is initially given by my genetic endowment and early socialization experiences.

The point just made is independent of the possibility that a successful hard determinist argument would undermine the idea of moral responsibility altogether. Take for granted the metaphysical assumption that agents have free will. Even with this assumption in place, the task of sorting out that for which I can legitimately be held responsible, and that which must be attributed to fixed unchosen aspects of my character, is hopelessly unfeasible, if not impossible, given the informational constraints that public officials making and administering law must accept. The project of eliminating the chance variations in the liability incurred by identically faulty agents in a corrective justice system by replacing tort law with an at-fault pool is unrealizable.

The argument that it is morally arbitrary to hold people responsible for outcomes insofar as they are affected by luck can hardly be a decisive objection to corrective justice. Notice that a similar argument from luck could be developed against the element of retributive justice that figures in criminal law practices. From such an argument one should not conclude that it is misguided to try to impose criminal punishment only on those found guilty of a criminal offense by a fair process and to try to make the punishment for each offense proportional to its severity.

The impact of the argument from luck on Coleman's normative rationale for current tort law practice is as follows. First, the relational aspect of corrective justice linking particular in-

16. RAWLS, *supra* note 6, at 312.

jurers to the particular parties they injure is not anchored in any compelling justice norm. The Coleman interpretation of corrective justice should give way to the Proportional Fault principle. Second, when fault is interpreted as moral fault, no remotely feasible version of the at-fault scheme fully implements the Proportional Fault principle. Epistemic obstacles set limits to the extent to which society could or should strive to fulfill Proportional Fault. Third, the fact that Proportional Fault cannot be implemented gives that principle less weight in competition with other values than it would otherwise command. But the importance of this point is diminished by the fact that Proportional Fault can be partially implemented; it is not an idle wheel. Fourth, Coleman's claim that society is not required to implement corrective justice norms is too weak. So far as justice is concerned, Coleman's corrective justice is defective. Only pragmatic considerations might block the conclusion that replacing the fault system with an at-fault system is morally mandatory. Fundamental justice norms undermine the fault system that assigns liability for injuries to the particular faulty parties who caused them.

Consider the principle that one should be held responsible only for what lies within one's voluntary control. Suppose that, having the freedom to choose to drive carefully, I drive carelessly on some occasion and my careless driving causes an accident. The doing of an act that I could have refrained from doing was necessary for the occurrence of the outcome. In another sense, the outcome did not lie within my voluntary control. Given what I did, if further events beyond my control had not occurred, my careless driving would not have caused an accident. In the prevention sense, I am responsible for the accident because I could have averted it; in the control sense, I am not responsible for the accident because what lay within my power did not suffice to bring it about. The assumed fact that I could have prevented an accident from occurring does not allay the worry that it is unfair in an ideal sense to require me to pay more compensation for damages than someone whose identically faulty act happens to cause lesser harm, or no harm at all.

The idea that one ought to be held responsible only for what lies within one's voluntary control is too crude as it stands to express the ideal of eliminating the morally arbitrary effects of luck. Suppose that both Smith and Jones could refrain from

perpetrating identically faulty actions that impose excessive risk of harm on others, but neither does. It might still be true that Smith was barely able to refrain, whereas Jones could have refrained easily and without suffering significant costs. Given this differential cost and difficulty of refraining from the identically faulty act, Smith and Jones should not be held equally responsible for doing it. The idea that one should be held responsible only for what lies within one's voluntary control is a slogan rather than a nuanced articulation of an acceptable principle. Voluntariness admits of degrees along several dimensions.

Coleman's normative defense of corrective justice is severely compromised and his normative critique of economic interpretations of tort law severely attenuated. The backward-looking norms of corrective justice must be reformulated to be credible, but a credible Proportional Fault norm is neither fully implementable nor significantly implemented by current tort law. As the ideal of backward-looking justice fades from the stage, more room is left for forward-looking broadly utilitarian considerations. To the extent that current tort practice can be justified, the efficiency considerations emphasized in economic analyses of the law will carry the burden of justification.[17]

V. Wrongful Losses and Rights Infringements

Corrective justice norms as interpreted by Coleman are framed to explain how liability is reasonably assigned to parties who do not cause injury by faulty conduct. One type of liability not involving fault that corrective justice norms are supposed to interpret is strict liability for injuries caused by engaging in ultrahazardous activities such as blasting a tunnel through a residential neighborhood. Another type of liability not involving fault is illustrated by the story of the desperate backpacker who breaks into a mountain cabin and uses its resources to save his life, thereby inflicting damage on the property. The story is intended to elicit the judgment that the hiker should be held liable for damages to the owner of the cabin even though the hiker has not behaved in a faulty or wrongful way.

Coleman's general scheme holds that corrective justice is concerned with wrongful losses, which may arise either from

17. *See generally* Steven Shavell, Economic Analysis of Accident Law (1987).

wrongdoing or from wrongs. Wrongdoing is wrongful conduct that harms someone's legitimate interests without violating any rights. Wrongs are invasions or infringements of rights. Invasions involve wrongful conduct that violates a right whereas infringements involve nonwrongful conduct that violates a right. As I understand Coleman's view of strict liability for ultrahazardous activities and for appropriations of property as in the backpacker case, he regards these activities as infringing rights, hence as generating wrongful losses.

Coleman's taxonomy is analytically helpful but misleading in at least one respect. The Procrustean element in his schema is the attempt to find wrongful loss in the backpacker and blaster cases. In these cases nothing that should count as violative of anyone's right occurs so long as full compensation for losses caused is forthcoming. In the case of ultrahazardous activity, the agent is not enjoined from blasting. She is doing what she has a right to do, provided that she conducts her operation nonnegligently and pays for any damages caused by her activity. (Let us assume that if the agent blasts negligently and so causes harm, she would be behaving wrongfully and would be subject to punitive damages.) If the agent has a right to engage in the activity that causes harm, it is confusing to add that if she causes harm, she violates a right of the injured party. If the blaster's activity threatened to violate someone's right, one would suppose it would be permissible to force her to desist. But no such permission is granted to those who might be harmed by the blaster's nonnegligent ultrahazardous activity. Similarly, if the cabin owner had a right that the backpacker not trespass on her property in these circumstances, one would suppose she would have the right to use reasonable force to block him from trespassing. But if the owner happened to be present when the backpacker chanced upon the cabin, the owner would not have the right to deny use of the cabin to him. In establishing tort liability rules for this type of case, society limits the property right of the cabin owner so that it reduces to a right to compensation for damages. It seems mistaken to interpret the backpacker and blaster cases to be instances of wrongful loss giving rise to corrective justice obligations to pay compensation.[18]

18. I would not wish to deny that there can be cases of innocent infringement of a right. In the cases in which a right is violated in the absence of wrongful conduct, the

VI. Interpreting Tort Law

Coleman hopes to advance a normative critique of the economic analysis of law, at least insofar as this analysis ignores the corrective justice norms that Coleman finds central to tort law practice, but he leaves the critique undeveloped. Coleman neither definitively endorses the corrective justice norms he identifies nor gives reasons for endorsing them. He does articulate a descriptive critique of the economic analysis of law, on the ground that it does not fit the broad outlines of tort law practice. Coleman's aim is less to defend the mixed conception of corrective justice as a principle of justice than to use that conception to interpret tort law.

The economic analysis of the law is a large tent under which many activities fit. Let us say that in its normative dimension, the economic analysis of law posits that the law should promote Kaldor-Hicks efficiency (or perhaps Pareto efficiency) in an *ex ante* sense, and investigates how to achieve that goal;[19] in its positive dimension, the economic analysis of law tries to demonstrate that the law does achieve efficiency, and perhaps tries to show that the explanation of the law's having the content that it does is that it promotes efficiency. Coleman wishes to show in his positive analysis that the law achieves fairness rather than efficiency, and perhaps that the explanation of the law's having the content that it does is that it promotes fairness. "Fairness" here refers to fairness as the mixed conception of corrective justice conceives it.

These last claims perhaps overstate Coleman's aims. He does not really discuss the extent to which the pursuit of efficiency is incompatible with the pursuit of fairness under the model of corrective justice. Nor does he offer an investigation of the extent to which tort law does pursue fairness rather than efficiency, where these goals conflict. He does not try to develop a complete set of substantive norms of corrective justice and at-

---

rights violator is acting in ignorance, or is in some other state of incompetence that reduces the voluntariness of her act. The backpacker and the blaster may well be acting on fully voluntary choices, so the terminology of rights infringement seems misplaced.

19. A change from one state of affairs to another is *ex ante* Pareto superior if someone's expected utility is increased by the change and no one's is decreased. Whether a change improves or worsens my prospects is relative to a specification of the change. If a change benefits farmers by $2 and hurts ranchers by $1, the change increases my expected utility if I have equal chances of being a farmer or a rancher, but lowers my expected utility if it is known that I am a rancher.

tempt to interpret the law against the background of these substantive norms. He is working at a higher level of abstraction, he tells the reader. His professed aim is to articulate the analytic categories of corrective justice, which particular theories of corrective justice would fill out in different ways. The question arises: What are the proper criteria of a successful interpretation of tort law, as Coleman conceives the enterprise? "A successful interpretation must not only fit but also justify the practice it interprets," writes Ronald Dworkin.[20]

Interpretations so conceived are not amenable to precise assessment, for they pursue two goals at once—fit and justification—and there are no doubt many different ways to weigh their relative importance. Coleman aims to exhibit a partial but significant fit between corrective justice and tort law. He indicates his belief that those tort law practices that fit corrective justice are justifiable but does not explain why.

A Dworkin-style interpretation is a hybrid creature—part justification, part explanation. An interpretation successfully "fits" the phenomena when it offers a set of principles that explain the phenomena in the sense that agents who wished to implement these principles in the world would produce exactly these phenomena. A Dworkin-style interpretation is a sympathetic explanation; the interpreter seeks to find an explanation that best justifies what he is trying to explain, other things being equal. Of course, the best interpretation of a practice may not justify it, because the practice may be partly or wholly unjustifiable.

The explanatory principles that figure in an interpretation need not correspond to the intentions of those who created the object that is being interpreted. Perhaps important tort doctrines were established by corrupt Nineteenth-Century judges who intended to benefit railroad owners come what may, but that does not render the tort doctrines unjustifiable. The doctrines are justifiable to the extent that they can be hypothetically explained by principles that also normatively support them.

Under a regime of tort law, one who suffers accidental loss can sue for compensation only someone who caused the loss to occur—by faulty conduct if this type of case is governed by neg-

20. RONALD DWORKIN, LAW'S EMPIRE 285 (1986).

ligence liability, by nonfaulty conduct if this type of case is governed by strict liability. Opposing the economic analysis of tort law, Coleman celebrates the structure of tort law insofar as it allows only those who cause injury to be tagged with liability:

> [W]e can understand that structure as embodying the ideal that in torts a victim seeks to show that the loss he has suffered is a wrongful one, one which requires recompense as a matter of right, not utility. The heart of the claim reflects a relationship in the world between injurers and victims. In general, the structure of tort law reflects that relationship.[21]

Coleman might be saying that the economic analysis does not "fit" tort law, but if this is his claim, he offers no evidence for it. In particular he does not analyze carefully the extent to which tort law secures relational rights as opposed to utility. My suggestion is that this passage might be read as 1) suggesting a third test of a successful interpretation along with fit and justification, and 2) asserting that the mixed conception of corrective justice scores high on this third test. This third criterion of a successful interpretation of a practice is that it should cohere with the self-understanding of the agents whose activity constitutes the practice. To see what this coherence test amounts to, consider a socio-biological explanation of tort law practice in terms of the maximization of inclusive reproductive fitness. Imagine that this socio-biological theory accounts for past judicial decisions and has good predictive success. Although it succeeds as social science, it is foreign to the judges' own understanding of their activity; even after considered reflection they disavow it. A successful interpretation, in contrast, must re-describe the agents' aims and intentions in ways that they eventually can avow as an account of what they are trying to do. If an interpretation of a practice aims to satisfy all three tests, it straddles the fields of social science, moral theory, and hermeneutics.

If Coleman is claiming that a corrective justice account coheres with judges' understandings of what they do in rendering verdicts, his account probably succeeds. Does the economic analysis of law fail the test of coherence? This is less clear. Judges do sometimes speak the language of economic rationality. Note that the question is not whether judges would immediately recognize their aims as coinciding with efficiency but

21. COLEMAN, *supra* note 1, at 469.

whether they would eventually arrive at this recognition, after reflection. At any rate, the coherence aspect of Coleman's enterprise as I am interpreting it is probably not directly in competition with the economic analysis of law, because economic analysts do not generally seek interpretations that cohere.[22]

Those who argue that tort law ought to seek economic efficiency might hold that efficiency should become the explicit goal of judges. For all that has been said so far, however, it might equally well be the case that judges will attain efficiency better if they do not consciously strive to reach it but instead aim to achieve some other goal that happens to coincide to some extent with efficiency. In this event, the economic analysis of law would better remain an esoteric doctrine.

VII. Tort Law and Distribution

Coleman offers interesting and insightful comments about the relationship between distributive justice and corrective justice. If the distribution of entitlements in society is not just, why is it morally worthwhile to restore those entitlements when people impose wrongful losses on one another that call for correction according to corrective justice? At the extreme, if the existing distribution were determined entirely by acts of theft and aggression, one might wonder why any moral significance should attach to maintenance of corrective justice. The interesting puzzle here is that many people view existing distributions in many societies as unjust, yet do not advocate abandonment of corrective justice pending the full achievement of distributive justice. Coleman offers the conjecture that distribution in a society must pass a threshold level of distributive justice, but that once this level is passed corrective justice norms have full authority even though the distribution is not perfectly just.

An alternative view is that the greater the extent to which distribution satisfies the requirements of distributive justice, the greater is the moral importance from the standpoint of justice of satisfying the norms of corrective justice. But even if the distribution of entitlements is massively unjust, people can expect to be better off if tort law serves the functions of spreading

22. *See generally* William M. Landes and Richard A. Posner, The Economic Structure of Tort Law (1987); *see also* Shavell, *supra* note 17.

risks and providing appropriate incentives to avoid accidents and other social harms. Alternatively, if distributive justice consists of consequentialist norms such as making the worse-off better off, perhaps tort law should be viewed as one among several instruments for achieving distributive justice.

The principle of giving priority to the worse-off does not always conflict with a straight efficiency analysis. Consider an example from products liability law. Suppose that bread can be baked with greater or lesser hygienic precautions, associated with higher or lower production costs. Suppose also that there are just two types of bread consumers, wealthy and poor; the wealthy consumers prefer to pay the extra cost for healthier bread, whereas the poor prefer to pay the lower cost and accept the higher risk of disease. If the government mandates high-quality hygiene in all bakeries in order to protect the consumer, low-quality bread will not be sold, with a resultant utility loss to the poor. (If there are scale economies from combining the two markets, high-quality bread may now be a bit cheaper but likely not a good substitute for the low-quality bread sought by low-income consumers; the end result is that the government regulations will benefit rich bread-eaters and burden poor bread-eaters.)

The discussion above assumes that manufacturers and consumers are perfectly well-informed about the risks and benefits of the actions they might choose, and perfectly able to utilize their information so as to make rational choices. Of these two assumptions, the second is more crucial for legal policy. If consumers lack information about risks of harm from products they buy, and manufacturers can obtain this information, government can mandate appropriate warning labels, which will be utilized effectively by consumers. If consumers lack the ability to process this information and modify their choices appropriately, mandated provision of information may produce consternation rather than utility-maximizing choice. Decisionmaking defects in economic agents may derive either from "cold" cognitive inability to make correct inferences from data to beliefs or from "hot" interest-driven distortion of beliefs. The extent to which such disabilities seriously deflect agents from utility-maximizing choices is a controversial and undecided matter.

The failure of agents to exhibit utility-maximizing rationality has distributional consequences if legal policy is based on the

assumption of full rationality, and if this assumption is fulfilled to markedly different degrees among the agents affected by the policy. Discussing products liability reform proposals, Alan Schwartz asserts that policy should be guided by the norm of consumer sovereignty, which holds that "the law should reflect the preferences of competent, informed consumers regarding risk allocation."[23]

This means that if consumers are competent and informed, the law should follow their preferences. If, on the other hand, citizens are heterogeneous in decisionmaking competence, the rule that serves the competent may not serve the less competent. While it is not true that when rational and rational individuals interact, the equilibrium that results will always yield more favorable outcomes to the rational, the tendency is obvious. Suppose that we seek to find whether a defense of contributory negligence should be added to a rule assigning strict liability to a class of manufacturers. Where care by customers affects the likelihood of accidents, a contributory negligence rule is attractive in that it gives the right incentive to minimize total accident costs provided the due care level is set properly. If customers wrongly estimate the risks of accidents stemming from product use, better provision of information can be arranged. If consumers vary in their capacities to integrate risk information sensibly into their decisionmaking, though, the rule that is best for smart customers will cause foolish customers to incur losses stemming from too little care, incompetent care, or excessive care. On the average one expects that customers prone to bad decisionmaking will be less well off than others over the course of their lives. Failing to hold them responsible for their innocently faulty decisionmaking gives perverse incentives against competent decisionmaking, while ignoring their incompetence effectively punishes them for factors beyond their control. The problem is exacerbated by the fact that governments cannot reliably distinguish innocent from non-innocent incompetence, and innocent from non-innocent faulty conduct. The distributional conflict between the interests of better and worse choosers is pervasive in tort law.

23. Alan Schwartz, *Proposals for Products Liability Reform: A Theoretical Synthesis*, 97 YALE L.J. 353, 355 (1988).