Liberal Markets

Jeremy Waldron

University of California, Berkeley

Follow this and additional works at: https://repository.law.umich.edu/mlr

Part of the Law and Economics Commons

Recommended Citation
Available at: https://repository.law.umich.edu/mlr/vol92/iss6/9

This Review is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
LIBERAL MARKETS

Jeremy Waldron*


INTRODUCTION

What is the relation between markets and the law? At first sight they seem to be mutually exclusive forms of social ordering. In markets, individuals make their own arrangements. If cooperation takes place, it takes place voluntarily. If resources are transferred from one person to another, both individuals must agree. We pass laws, on the other hand, when for some reason we are unwilling to allow people to make their own arrangements. In a legal ordering, we decide as a society what is to be done in some set of circumstances, rather than allow individuals to sort things out for themselves. Thus law seems to evince a suspicion of the kind of unilateral or bilateral decisionmaking that market transactions epitomize.

The paradox of course is that market societies are among the most rule-governed and legalistic societies we know. They embody rules of property, they are enveloped in commercial codes, and the individuals whose freedom they vindicate seem to engage in almost continual litigation. Whatever we conclude, then, about law and markets, it cannot be that they are contrary or antipathetic forms of social organization.

One view is that legal rules operate as supplements to markets, correcting for the failure of market transactions in the messy circumstances of the real world to produce the efficiency and prosperity that economists associate with perfect competition, full information, and zero transaction costs.

They may work, for example, as follows: In the real world, it is sometimes possible to imagine transactions that would have occurred but for some identifiable imperfection, uncertainty, or transaction cost and to work out what the outcome would have been if those transactions had taken place. On the basis of that counterfactual reasoning, a society may deploy legal rules as a response to market failure, either by directly imposing the outcome that would have resulted in ideal circumstances or by permitting individuals to bring about the outcome unilaterally, without the normal market requirement that the consent

of those affected must first be secured. Thus, if transaction costs prevent individuals from fully specifying the terms of some agreement, the law of contracts may impose on them the terms to which they would have agreed in the absence of transaction costs, and, through the award of damages, the law may adjust their resource holdings accordingly. Or, if transaction costs prevent negotiation altogether, the law of torts may allow something like "forced transfers" at the initiative of one party, requiring only that he pay as compensation to the other person affected the price upon which the two of them would have agreed if they had had the time and opportunity to deal willingly with one another. In this way, the law may respect underlying market values, even while it is coercively imposing transactions on parties who have not in fact agreed to them. To justify such imposition, we can say that the real world in which the rules are coercive is a world where voluntary interaction is blighted anyway by transaction costs, uncertainties, and other imperfections. The world in which all transactions are truly voluntary is the ideal world of perfect competition. Law, then, is really respecting, not violating, our freedom by moving us coercively from the mire of inefficiency in the real world to the efficient outcomes that we ourselves would have agreed to, if only we had lived in an environment more conducive to trade.

In a collection of essays, *Markets, Morals and the Law*, published in 1988, Jules Coleman1 wrote that he was “currently working on a book . . . that explores the extent to which nonmarket institutions [like the law] can be explained and justified as responses to problems of market failure.”2 The working title of that project was “The Market Paradigm,” which is Coleman’s term for the approach I have just outlined. The book that has eventually appeared, however — after many vicissitudes including, we are told, the theft of the only existing version of the copy-edited manuscript (p. xvi) — is entitled *Risks and Wrongs*. It amounts to a comprehensive critique and repudiation of the market paradigm and its replacement by an alternative theory about the role contract law plays and particularly — as the title indicates — the role tort law plays with regard to market transactions.

According to Coleman, the market paradigm, with its emphasis on repairing market failure, neglects the role played by legal rules in creating and sustaining structures of market interaction in the first place. Coleman rejects the view that we need law only when markets do not work. He argues, on the contrary, that we need law in order to make markets work. We need a system of social cooperation to make a system of voluntary transactions and market competition possible (pp. 59-62).

The point is most obvious in the case of property. Without rules to

1. John A. Garver Professor of Jurisprudence, Yale University.
secure particular resources to particular individuals, the idea of market exchange makes little sense. As Coleman puts it, "You and I will have a hard time negotiating the sale of some piece of property if we both believe we own it" (p. 147). Property rights embody the essential principles of a market order: that particular resources are assigned to particular individuals; that an individual’s consent is required before anything is done to the holding assigned to him; and that each person has the legal power to transfer his holding to another on, roughly, any terms they agree to. It goes without saying that the existence and efficacy of these principles is wholly the work of the law and that markets could not possibly exist, even in an ideal form, without them. Law here is constitutive, not just ameliorative.

To put it another way: on issues concerning who gets to control and make decisions about which resources, we are not willing to leave individuals to their own devices, at least not initially. Left to their own devices, individuals might well engage in a Hobbesian war of all against all for the use of the resources that surround them. I might try to use and make decisions about the very piece of land that you are attempting to use or make a decision about, and the result would be chaos at best, violence at worst. The matter therefore is one on which we need a collective decision — a decision that initiates market processes by lending social weight, first, to the assignment of particular resources to particular persons and, second, to the endowment of those persons with the power to alter the assignment of the resources entrusted to them on terms of their own choosing. These collective decisions define a structure of rights that subsequently enables individuals to deal with one another on terms of mutual advantage.

Coleman does not spend a lot of time on property rules. In my view the main importance of property in his argument is the foundation it lays for his understanding of the law of torts. In the paradigm that Coleman attacks, tort law is a response to market failure: when transaction costs are high, liability rules allow individuals to act in ways that affect others’ rights and to compensate them ex post without having to secure their consent in advance. This suggests that there would be no place for tort law absent market imperfections. But of course that is nonsense.

In a perfectly competitive market with zero transaction costs, individuals might still be tempted to act in ways that impact on others’ rights without securing their consent. Even if a consensual action would lead to an efficient outcome, someone might still be tempted to bypass the consent requirement in order to avoid having to pay the price of the other’s consent. Such actions would be simply wrongs — that is, infringements of the rights established by the basic property rules protecting persons and their holdings. Any legal system must have ways of dealing with such wrongs and making good the losses
that they occasion. On Coleman's account, this is the basic task of tort law. To understand it, we need a theory of corrective justice — a theory that operates at a deeper level than the economist's account of market failure. Tort law is not just a matter of solving problems of transaction costs. The aim of this area of the law is to vindicate against loss or derogation the very rights that constitute the fabric of the market order.3

For a market to exist, not only must there be an array of individuals with alienable entitlements, but these individuals must also be disposed to bargain and strike deals with one another. The primary task of contract law, Coleman believes, is to foster this disposition by providing safeguards against violation — safeguards that the parties would otherwise have to spend valuable transaction resources creating for themselves — and default rules to cover uncertainty and unforeseen contingencies (pp. 122-40). By offering these resources, the law may open up opportunities for cooperation that a costly search for certainty and security would otherwise dissipate.

With regard to contract law, then, Coleman's story is close to the account the market paradigm offers, inasmuch as both accounts stress the importance of responding to transaction costs. But there is this difference. According to the market paradigm, the point of contract law is to seek directly the efficient outcomes that the parties have been unable to secure themselves. When contractual difficulties arise, we hold the parties to the contract we reckon they would have agreed to ex ante in the absence of transaction costs. Under Coleman's account, however, the main point of contract law is not efficiency as such but the facilitation of market transactions by reducing in advance the misgivings that people in the real world may have about the uncertainty of dealing with others. In fact, the best way of reducing such misgivings is to announce in advance that, in the event of any dispute, the terms of the hypothetical ex ante contract will be imposed. Any other default rule would have the effect of discouraging one or the other party from entering into the contract. But, though his solution is roughly the same as the one proposed in the market paradigm, Coleman's rationale is quite different. Indeed, differences like this abound throughout the book: Coleman indicts the market paradigm, not so much for what it proposes, but for its casual reliance on a poorly thought through notion of efficiency and for its lack of attention to the other interests that individuals and society as a whole may have in the integrity of market transactions.

In the pages that follow, I shall outline and examine the issues

3. This outline will do as an initial characterization of Coleman's position. Later, however, we will see that his view purports to be somewhat more complicated than this: tort law not only vindicates rights against infringement but also protects certain interests against wrongful action even when those interests do not amount to rights. See infra Part III.
raised in this general overview in more detail, paying attention to qualifications and provisos that are necessarily neglected in an introductory account. In Part I, I shall focus on Coleman’s general view of the importance of markets in a liberal order. I shall look at his critique of theories organized around the idea of efficiency and examine his alternative view that the function of markets is to preserve the stability of large heterogeneous communities. In Part II, I shall look briefly at his conception of contract law as a body of legal rules that offer resources to individuals who might be diffident about dealing with one another in the social conditions in which markets seem necessary, and I shall criticize his account for failing to emphasize the justice issues involved in contractual breach. Finally, Part III will be devoted to a discussion of corrective justice. Here I will argue that the methodology Coleman uses to analyze tort liability leaves many underlying moral questions unanswered in a way that sits ill with his critique of the superficiality of the usual economic approach.

Before I continue, I want to note that *Risks and Wrongs* is not an easy read. It is densely argued and proceeds at quite a high level of abstraction and technical sophistication. It assumes acquaintance with rational choice theory, an ability to grasp game-theoretic matrices, a working knowledge of law and economics, and considerable familiarity with the main alternatives in the growing field of the philosophy of tort law.

In addition, the author’s lingering affection for the positions he attacks makes his argument difficult to follow at times. Coleman is interested in and understands the market paradigm; he has thought through the concepts of rational choice, economic efficiency, and market failure perhaps more deeply than many of their proponents. He is anxious that these ideas not be misunderstood — as they are by almost all of their contemporary critics — and he is confident that nothing but a deep and detailed grasp of the complex structures of these ideas can reveal their inadequacy.

In itself, this is an honorable approach. It has the consequence, however, that Coleman often spends as much time expounding the positions he is attacking and defending them against spurious objections as he does articulating his own critique. Moreover, because he often proceeds without explicit organizational signposts, and because his own view is always complicated, in some instances differing only subtly from that of the opposition, it is sometimes hard to discern whether a given position that Coleman expounds is one that he wants to convince us of or not except by reading on for ten or twenty pages. I suspect that in some cases Coleman himself is not sure. To read his work is to watch theoretical argumentation in progress. An attractive feature of Coleman’s approach to his subject — a feature all too rare in modern jurisprudence — is that he does not attempt to conceal from
the reader when he is torn between two views or when he is in the process of changing his mind.

All this makes an already dense book more difficult to follow. It does not, however, detract in the end from its value. For those who persevere, *Risks and Wrongs* is as rewarding as it is challenging. With some books one comes away with a sense of having been bullied or bluffed; in this case, one comes away from the complexities, uncertainties, and detail of Coleman's account with a sense of having engaged in the honest labor of trying to work out something hard in the company of a scrupulous and formidably qualified companion.

Of course for those readers who know the ropes and are ready and waiting to engage with Coleman on some specific issue, such as tort law and the nature of corrective justice, these aspects of Coleman’s style will not be so much of a problem. These readers are used to hearing him think out loud. They profit from that as a matter of course. For them, the main virtue of *Risks and Wrongs* will be the splendid opportunity it provides to ponder at length positions that have so far been articulated only in relatively short articles.

I. **THE SOCIAL IMPORTANCE OF MARKETS**

The first part of *Risks and Wrongs* is dominated by the following aphorism: “The market paradigm,” says Coleman, “holds that cooperation is a solution to failed competition, but it is equally plausible to view competition as a solution to failed cooperation” (p. 60). The view that Coleman wants to attack presents competition in the marketplace as the basic or primal form of human interaction and holds that cooperative interaction — that is, social decisionmaking through political processes yielding legal rules — is made necessary only when market competition fails. The view he defends takes social decisionmaking as basic and presents market competition as a form of cooperative interaction to which we resort when the costs of social decisionmaking become too high.

Coleman’s alternative has the advantage of sociological realism. Market competition cannot flourish prior to social cooperation, he says, because to a certain extent competition depends upon cooperation.

Competition presupposes a stable, enforceable scheme of property rights. Any such scheme is a collective or public good. So we must assume that in order for competition to exist, at least some problems of cooperation — for example, the establishment and maintenance of an enforceable property rights scheme — have been successfully resolved. [p. 61] Similarly, markets presuppose constraints on the use of force and fraud and, more generally, limits on the practices to which individuals may resort in the scramble for competitive opportunities. The establishment and enforcement of these norms, Coleman argues, is a matter
for society. It cannot be expected to emerge as a by-product of unregulated competition among individuals.\(^4\) Markets, then, must be explained as a special case of social cooperation, rather than the other way round.

Quite apart from what markets presuppose, there are many social forms that are simply unrelated to markets, and Coleman observes quite correctly that it would be nonsense to suggest that these social forms have no part to play in our understanding of law (p. 73). Humans are social beings; they awake to consciousness to find themselves cooperating in rich structures of interaction with others. These structures cannot be grasped using only the bare-bones axioms of rational choice theory. Answering any interesting question about social organization, Coleman argues, "requires detailed understanding of a community's history and its culture" (p. 67). He writes:

No substantive conclusions about the reasonableness or the rationality of the ways in which individuals organize their collective lives together can be derived from an abstraction, from, in particular, the ideal of rationality alone.

... It is of no help to argue that the institutions that emerge or would emerge among rational agents at different places and times in different circumstances are a function of "transaction costs" of one sort or another. For everything that is interesting about a people and relevant to the determination of rational organization among them falls in the category of transaction costs. [pp. 66-67]

Coleman does not mean to suggest that we are at the mercy of local conditions or that local ways of doing things are simply given to us as


Individuals accustomed to preying upon one another may bargain toward what is in effect a peace dividend, recognizing each other's de facto possessions and saving at least some of the costs of defending those possessions against others' incursions. Though this alternative means forgoing the advantages of successful predation, rational individuals might reckon it worthwhile, especially if the era of conflict had settled down into some sort of rough equilibrium, in which the marginal benefits from further investment in predatory activity were equal to the marginal costs that such effort required.


Briefly, Coleman responds in the present volume by denying that it would necessarily be in everyone's interest to agree to an arrangement of the sort that Hume indicates: "[C]ooperation that requires forgoing predation may not be a rational strategy for all agents." [P. 57]. This is certainly true: the Hume-Buchanan theory represents at best a possibility. There is nothing inevitable about it. Ideally, though, Coleman's strategy in the chapters under discussion in this Part of the review would require a demonstration of the impossibility of this mode of emergence of property rights. This he has not provided.
a part of our social being. The particular kind of social being that one finds in a society like ours is the individual with interests and an inquisitive mind of her own. Moreover, the individual's perspective is matched by the acknowledged mutability of social forms. We have developed a useful practice of interrogating social arrangements at the bar of rational individual advantage and changing them sometimes when they fail that test. All the same, even if rational choice theory is relevant to social understanding, it is relevant only as an expression of rational cooperators' grasp of the advantages and disadvantages of particular cooperative forms under particular social conditions. Because, in Coleman's view, markets can be thought of only as socially constructed institutions (p. 5), they are to be understood in accordance with this methodology. We should not explain markets in terms of the primacy of competitive individual behavior. We should ask instead: Under what conditions would rational cooperators find it advantageous to set up competitive markets as a form of social interaction or to sustain such markets if they found that their society already provided them?

The usual answer to this question is that societies set up markets in order to promote economic efficiency — in order, that is, to secure Pareto gains that cannot be secured through deliberative decisionmaking by the society as a whole. Coleman, however, is adamant that efficiency is not a helpful term to use in describing the advantages of markets (p. 102).

His point here is partly a quibble about the casual use of efficiency and inefficiency in law and economics. For Coleman, the most illuminating definition of efficiency is the simplest: Pareto optimality. A situation is efficient if and only if there is no way of changing it that could possibly make one or more persons better off without making somebody worse off. Now, if efficiency as such were a value, any situation with this characteristic would be preferable to any situation without it. But that conclusion is manifestly false. A situation in which I have some apples and you have some oranges is probably inefficient, in view of the possible gains from trade. A situation in which I have all the fruit and you have none is almost certainly efficient because there is no way of making either of us better off without making someone — me — worse off. However, the latter situation is certainly not a Pareto improvement upon the former, and arguably it is not an improvement of any sort. Indeed, according to strict economic theory, the two situations are not comparable: “From an economic point of view, there are no grounds whatsoever for asserting that efficiency is always preferable to inefficiency” (p. 101).

The point is obvious and can be accepted immediately. Normally, however, when people extol the efficiency-promoting properties of markets, they have in mind the ability of markets to generate Pareto-
optimal outcomes via Pareto improvements. Coleman recognizes this point (pp. 101-02), but he quite rightly insists that we should ponder the significance of the qualification.

One way of looking at the requirement that we should secure efficiency only through Pareto improvements is that it is the expression of a very strong notion of cooperation: something counts as a form of cooperation only if all the persons involved in it stand to gain. According to this notion, a situation in which someone gains at another's expense is a case of exploitation, not cooperation — even if the former's gain is greater than the latter's loss. Cooperation in this sense necessarily requires mutual advantage.

If we analyze the matter this way, however, we find that we have said very little in fact about the substantive advantages of markets when we say that they generate efficient outcomes via Pareto improvements. Calling something a Pareto improvement is simply a way of characterizing it as a cooperative move, and efficiency, in this context, means simply that no further such Pareto improvements are possible. So, saying that markets have the advantage that they promote efficiency is saying nothing much more than that they take cooperation, in this strong sense, about as far as it can be taken. The efficiency claim tells us little of interest about what is distinctive — and thus distinctively advantageous — about the forms of cooperation that markets involve.

Let us return to what is structurally distinctive about market transactions. Although each market transaction takes place in a society and affects the overall state of affairs in that society, its distinctive structural feature is that it requires decisionmaking only by the pair of individuals who happen to be involved. For example, a given field is currently planted in carrots, and someone has the bright idea that the land could be better used if it were planted in tulips. In a nonmarket society, peaceful exploration of this option requires a decision by everyone, or at least a centralized decision taken in the name of the society as a whole. The society has to ask itself: Should we grow tulips or carrots in this field? If market structures are available, how-

5. Cf. ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 33 (1974) ("Using one [person] for the benefit of others, uses him and benefits the others. Nothing more. . . . Talk of an overall social good covers this up. (Intentionally?).").

6. As a way of characterizing what markets can achieve, this point is almost trivial. A market transaction is by definition a voluntary transaction on all sides, and the idea of mutual gain is usually not much more than a representation of voluntariness in the language of individual utility. The point becomes important, however, in the context of what Coleman calls the market paradigm, for in this case we are considering legal responses to market failure.

We talk of market failure when processes of voluntary interaction fail to yield the outcome that we think in some sense the market ought to have yielded: we are stuck in $S_1$ and we believe the market ought to have yielded $S_2$. Because, in these cases, $S_2$ is defined as the favored outcome — that is, the outcome at which the law should aim — independently of its being the outcome of actual processes of voluntary exchange, it becomes substantively important to insist that $S_2$ be not only efficient but also a Pareto improvement over $S_1$. 
ever, there will be some individual who is, so to speak, in charge of the carrot field and who is authorized to transfer it to the control of tulip enthusiasts — or anyone else — on almost any terms he pleases. The decision about changing the crop can be made therefore as a matter of bilateral negotiation rather than centralized decisionmaking about resource use.

Under what conditions might this be a useful way of running things? Coleman’s answer is that it becomes useful in proportion to the difficulties involved in centralized decisionmaking (p. 64). These difficulties will include the size of the society and the dispersion of its population, but the factor that most interests Coleman is the heterogeneity of values in society. If we have to make the carrot-tulip decision together as a collective, then we all have to grapple with the aesthetic and dietary issues involved. Some of us will regard tulips as frivolous; others will say that life with a high-fiber diet is barely worth living without color and beauty. The conflicts that these differences occasion might well be irreconcilable.

If all forms of interaction required shared commitments, little, if any, mutually advantageous and otherwise socially desirable interaction would occur. Were allocation decisions left to the political process, for example, conflict would abound. In contrast, the market has the special virtue of allowing, even encouraging, socially desirable and mutually advantageous interaction in the absence of broad consensus on matters of fundamental importance to the lives of its citizens. [p. 63]

The point is that the carrot grower’s values need not agree with those of the florist in order to transfer the field to him. In a money economy, the florist may be able to make the transaction worth the carrot grower’s while in a currency with which the carrot grower can pursue any goals he pleases. The florist can be expected to do this whenever he can make it worth the carrot grower’s while and still see some advantage from the transaction for the pursuit of his own florid values.

This is an enormously attractive conception of the advantage of markets from a liberal point of view. It indicates that we should expect market structures to exist and flourish in a large, heterogeneous community in which there is little broad consensus about tastes and values. We should expect markets to be less prominent in “small homogenous communities defined by shared values and interests” (p. 64), for in those contexts it would be much easier to assign allocational decisions to the society as a whole. Of course this is more or less exactly what we find.

Coleman’s account also has the capacity to explain why the voluntariness aspect of market transactions is important from the point of view of social organization. Usually, voluntariness and mutual benefit are represented as two sides of the same coin. Occasionally, however, we may want to say that a person has acted voluntarily to his own disadvantage. Or we may want to say — as we do in cases of “market
failure" — that a mutually beneficial outcome is available in principle, but that somehow voluntary transactions cannot reach it. In either of these cases, we may have to choose whether the voluntariness of market transactions matters more than their yielding mutual benefit or vice versa in order to decide whether social intervention is appropriate. If we determine that it is mutual benefit that matters, then we should encourage intervention by society — to reverse or replace voluntary individual decisionmaking. But, in this situation, those who make the social decision must find some common ground concerning the values involved. In other words, if a "beneficial" outcome is to be imposed involuntarily — so far as the immediate parties are concerned — then those who impose it must do so in the name of a socially authoritative conception of the good. If, on the other hand, we stress that it is voluntariness that matters, we do so partly because of our doubts about our ability to agree on standards of value and benefit that we can impose in the name of the whole society. Because of these doubts, we prefer to let pairs of individuals make their own arrangements.

Of course, as we have already seen, the members of a society could not cooperate through markets if all their values were contested. There must be an initial allocation of property rights, and there must be norms prohibiting force and fraud as well as limiting the excesses of competition. These are necessarily matters for social decision, and there is no way to finesse the value confrontations they involve. Nevertheless, as Coleman points out, "the scope of agreement about values is quite minimal when compared with the extent of fundamental disagreement compatible with market exchange" (p. 63). The market, then, is an artifact of our determination to continue cooperating with one another in society, even in the absence of consensus about the good. We use the values that we do share to create and sustain a structure that enables us to cooperate with regard to values that we do not share.

Unfortunately, Coleman spoils his account a little by his insistence that the key to all this is the value of social stability: "[M]arkets contribute to social stability. That is their attraction to liberal political theory" (p. 5). He writes, "The decision to organize allocation decisions by markets is a political one that has more to do with a desire to finesse or to avoid articulating the nature and scope of fundamental disagreements that may lead to instability than it has to do with the efficiency properties of such arrangements" (p. 64). Coleman's argument is that the fabric of social cooperation might be seriously imperiled by the conflict that centralized deliberation on values would generate. Better for the health of our cooperative arrangements to have allocative decisions made in ways that do not require us all to
address such contentious issues. In opposition to Coleman’s account, I want to argue that social stability, by itself, cannot be all there is at stake because stability could be promoted by means other than markets.

In fact, a society whose members disagree about values could secure stability by simply avoiding most of the decisions that pose potentially divisive value issues. Allocative decisionmaking could be minimized by a rule that resources will be devoted, in the name of stability, to the traditional uses to which they have always been devoted. Though each of us has bright ideas about how to better society, we should recognize that such ideas contain the seeds of conflict. A society that fostered this cautious attitude would not be very progressive or prosperous, but it would certainly be stable.

In fact, I doubt whether Coleman believes that social stability is the ultimate value. In at least one formulation of his argument, the language he uses is more complicated: markets “maximize social intercourse without raising the questions about the scope and extent of any underlying consensus, thus contributing to social stability” (p. 67). Here his point is not merely that markets allow interaction without threatening stability; more conservative institutions could do that, but the level and density of interaction would be very low. Rather, the virtue of markets is that they allow a very high level of interaction in society without generating value conflicts. Without markets we would have to purchase stability by forgoing many interactions — exchanges, for example — that were available in principle. I know Coleman is averse to efficiency-talk, but it is hard to see why we should not express this point by saying, not that markets promote stability as such, but that they allow the simultaneous pursuit of stability and economic efficiency.

Coleman might respond by saying that the alternative way of pursuing stability that I have mentioned — that is, minimize allocative

7. This stability-based account has at least the advantage of explaining the distinction between value dissensus and value skepticism. Coleman notes quite rightly that he is not committing himself to skepticism about ultimate values merely by virtue of the observation that people in fact disagree about them. Pp. 447-48 n.19. Even if these disputes have a right answer, the objective existence of that right answer will not by itself end disagreement or the conflict and instability that articulated disagreement might generate. For a related argument about the place of moral objectivity in controversies about adjudication, see Jeremy Waldron, *The Irrelevance of Moral Objectivity*, in *Natural Law Theory: Contemporary Essays* 158 (Robert P. George ed., 1992).

8. This is to say nothing of the profoundly — and often refreshingly — destabilizing effects of markets. See, for example, the apotheosis of market capitalism in chapter 1 of *Karl Marx & Friedrich Engels, Manifesto of the Communist Party* 83 (Samuel Moore trans., International Publishers 1933) (1848), in the famous passage that ends:

All fixed, fast-frozen relations, with their train of ancient and venerable prejudices and opinions are swept away, all new-formed ones become antiquated before they can ossify. All that is solid melts into air, all that is holy is profaned, and man is at last compelled to face with sober senses his real conditions of life and his relations with his kind.

*Id.* at 12.
decisionmaking and reinforce traditional structures — is unlikely to be available in a society that is already heterogeneous. We naturally associate the existence of value diversity with social forms in which people are impatient with the restrictions that traditional ways of life impose. That is a fair point. However, it raises questions, which Coleman largely neglects, about how societies develop the sort of value diversity that he describes. He rightly observes that any account of cooperative forms requires an understanding of local situation and culture (p. 67), but he neglects to mention that it also requires a theory of social dynamics. In fact, the most plausible account of the source of value diversity in society is that it is in part the product of, not just the background to, the emergence of market economies. If so, then markets — or protomarkets — must once have had attractions that were independent of their capacity to secure stable interaction in circumstances of social heterogeneity.

In fairness, I should note Coleman’s own careful proviso: “It would be foolish to claim that the only value of markets is the one I have set out” (p. 67). Markets serve other liberal values, he says, such as individual autonomy; in the concluding chapter of Risks and Wrongs, Coleman suggests that autonomy — or, as he puts it in his concluding paragraph, “political freedom” — may be even more fundamental than stability in our account of the advantages of market interaction (p. 439). The connection between autonomy and stability is a complex one, according to Coleman. On the one hand, autonomy explains why stability is a value: “stable social practices or institutions are a prerequisite of individuals being free to pursue their projects, plans and goals . . . ” (p. 436). On the other hand, as Joseph Raz has argued, autonomy may be impossible without a range of possible ways of life from which to choose — that is, without ethical diversity.9

We have to be careful, however, with this last point. Raz’s argument, as I understand it, is not that autonomy is an independent value that happens to require a diversity of ways of life and thus a stable structure of interaction that is compatible with that diversity. Instead his argument is as follows: first, in a society like our own, one can only prosper by being autonomous; second, autonomy of course requires an array of values from which to choose; but, third, the existence of such an array is normally one of the conditions under which autonomy becomes an option in the first place, rather than something one has to support or defend on the basis of autonomy’s importance.10

If a society were not ethically heterogeneous, it would likely not be the sort of society that valued autonomy. No judgment therefore about the desirability of ethical diversity can be predicated on the value of autonomy because the value of autonomy — at least as Raz defines it

10. RAZ, supra note 9, at 205-07.
— depends in part on the contingent fact that ethical diversity happens to exist. So autonomy, by itself, cannot ground the particular combination of stability and diversity that Coleman invokes as the basis of the attraction of markets.

II. THE ROLE OF CONTRACT LAW

I have dwelt at length on the account of markets presented in Part I of Risks and Wrongs, not only because it is intrinsically interesting, but also because it is crucial to what Coleman wants to say about other aspects of the law. We have seen that markets may be set up as a way of fostering interaction among people who do not have much in common. But people who do not have much in common are often leery of dealing with one another. If a particular transaction involves anything more complicated than the simultaneous handing over of cash and commodities — if, for example, it involves placing an order, laying down specifications, stipulating deadlines, turning up for work several days in succession — then there is room for uncertainty and mistrust: How do I know that he will do what he says? How does he know what “high quality” or “in a few days” means to me? Ex hypothesi, we do not agree about basic values, so how can we trust each other on detailed and delicate matters?

The paradox is an intriguing one: “Markets are most difficult to create and sustain where they are needed most” (p. 105). As Coleman writes:

Markets require contracting or exchange. Exchange is threatened by uncertainty. Uncertainty can be reduced by factors that are endogenous to the relationship between the parties. For example, if potential contractors are involved in repeat play or are members of closely knit communities, then their incentives to defect from transactions will be reduced by reputation effects. But in these sorts of circumstances, pure markets are not as important to social stability as they otherwise would be. [pp. 68-69]

The kind of heterogeneous society that requires the market as a basis of cooperation is necessarily a society of strangers, diffident and unsure of one another. If they can rely only on their own resources, individuals will have to proceed very cautiously in their dealings with one another, making exhaustive inquiries in advance about every potential partner-in-trade, stipulating everything explicitly, exchanging hostages, covering themselves with alternative arrangements in case their current deals fall through, and so on. The costs of this caution to either or both parties are likely, in many cases, to outweigh any possible advantage that may accrue from the agreement. The problem of transaction costs, therefore, is not merely an occasional vicissitude but

something that is endemic to the very situation in which transactions are deemed desirable.

Having posed the problem in this way, Coleman retreats from it a little. It is tempting to describe the transactional difficulties in a society of strangers as though they were single and isolated encounters between people who will never come across one another again. But a single transaction does not a market make, and it is markets that interest Coleman — markets in the sense of relatively enduring frameworks within which millions of transactions take place on a daily basis. The fact that a market comprises a large array of transactions may help to reduce the costs that individual traders face in finding and securing contractual opportunities. If the array is spread over time, for example — if the market comprises a large number of repeat players — then reputations for fidelity or defection may become established among traders. These reputation effects will be similar to those that accrue in a close-knit *gemeinschaft*, with the difference that they are predicated not on shared substantive values but simply on the common importance to all of successful trade or business. As Coleman puts it, “[a] potential trading partner’s noncompliance is the sort of information everyone has an interest in obtaining” (p. 142), whether or not she shares substantive values with those from whom she obtains the information.

Even more interesting are the connections between the structural characteristics that economists lump together under the heading of “competitiveness” and the problems that rational choice theorists associate with “the division of the cooperative surplus.” In a single deal from which both parties stand to gain, there will have to be detailed bargaining about the exact division of the surplus. In a competitive market, however, the bargaining problem is largely superseded by the emergence of market prices. The costs that any one pair of traders has to invest in solving the division problem are accordingly quite low.

Among the difficulties that remain, the main threat to markets is that parties are required to enter into transactions under conditions of uncertainty:

Almost every contract dispute that winds up in litigation turns on a point about an incomplete contract; the traditional reasons for incomplete contracts are matters of information cost: (1) Some contingencies may be unforeseeable; (2) planning for every foreseeable contingency can be expensive; and (3) some contingencies may involve private information. [p. 124]

Coleman does a good job of laying out the decision problems that this sort of uncertainty can create. He insists that we cannot understand uncertainty unless we grasp the game-theoretic structure of contracting, which he presents as, roughly, an impure coordination game embedded in a prisoner’s dilemma (pp. 106-08). A contract becomes possible when two people discern at least one way in which they would
both be better off by cooperating than by proceeding individually. Often, however, there are not just one but many possible ways of cooperating clustered around this contractual opportunity, each involving a slightly different matrix of gain for the individuals involved. Moreover, if the available forms of cooperation are complicated, the danger arises that either of the parties might find it even more advantageous to proceed only so far in the cooperative enterprise and then defect, leaving the other perhaps worse off than if he had not embarked on cooperation at all. Against this background, the parties must figure out ways of securing some particular arrangement, and they must proceed against the background of uncertainty about the factual situation, uncertainty about how the other sees the situation, and uncertainty about each other's bargaining position.

To act in the shadow of this uncertainty is to take a risk, and obviously the existence of that risk limits the range of deals that will seem attractive to each party: a small- or medium-sized potential advantage from cooperation may seem not worth pursuing when discounted by the probability that it will not in fact accrue. At the same time, any attempt by the parties to minimize risk by reducing uncertainty is itself costly. Risk minimization will therefore preclude certain possible deals at the margin by limiting the net advantage that can accrue from the contractual process as a whole.

Can the law — that is, society acting collectively — do anything to reduce this contractual uncertainty? In common with the defenders of the market paradigm, Coleman thinks it obvious that if the law can, it should. He briefly considers the view — which he attributes to James Buchanan — that transaction costs should be regarded simply as part of the real-world framework of exchange — as much a part of it, for example, as the structure of entitlements (pp. 91-98). On that view, what others call market failure are instances of efficiency, not inefficiency: exchange has gone as far as it can and, taking everything into account, it is not possible to make anyone better off without making someone worse off. To impose by law the contract to which the parties "would have" bargained but for transaction costs is tantamount to shifting the goalposts in the middle of the game, according to this view. Coleman has some sympathy for the view he attributes to Buchanan. If the parties' situation is defined by taking transaction costs into account, then imposing the bargain that they would have reached absent transaction costs smacks of the cardinal sin of imposing on them an efficient outcome that is not necessarily Pareto superior to their existing situation. Coleman respects Buchanan's insistence on voluntariness, not mutual benefit, as the sign of a market

---

12. Unfortunately, Coleman provides no reference. In an endnote (p. 450 n.2), Coleman refers the reader to an earlier discussion in COLEMAN, supra note 2, at 133. But, although that discussion also attributes the view to Buchanan, it does not provide any reference either.
transaction, particularly if mutual benefit is defined relative to an idealized bargaining situation that does not exist (p. 97). He also welcomes any approach to transactional difficulties that takes into account the real-world costs of removing them (pp. 134-36) and does not just imagine them away.

There are important differences, however, between contractual uncertainty and other aspects of the framework of exchange. We can see these differences if we analyze the possibility of changes in the framework in rational choice terms (p. 96). Some changes in the framework of exchange — the redistribution of property, for example — would be unwelcome to certain persons and so could not possibly be characterized as Pareto improvements. Others — such as the removal of transaction costs — may make everyone better off because they open up opportunities for Pareto improvements — relative to the actual status quo — that did not exist before. Coleman’s suggestion, then, is that we should think about contract law, not as a way of imposing idealized bargains, but as a way of defining a framework of market cooperation that makes certain deals possible that would not be possible if the framework were differently defined.

Potential contractors face uncertainty. They may be reluctant to enter into an agreement because they are unsure what will happen in the event of a given contingency or they may be unsure what will happen in the event of contingencies that they have not been able even to contemplate. Law can dispel this uncertainty — to some extent — by authoritatively announcing in advance what terms the parties will be held to in the event of their situation coming before a court. But what, in general, should such an announcement contain?

In some cases the content may be unimportant. Coleman offers the case of the mailbox rule as an example in which it does not particularly matter how contractual uncertainty is eliminated so long as it is eliminated (p. 138). In other cases, the default rule that the law provides may make a considerable difference. The position that Coleman ultimately favors is not significantly different from the view associated with the market paradigm. To facilitate and encourage contracting, the law should assure the parties in advance that, in the event of any dispute or uncertainty between them, it will impose on them as far as it can the terms to which it thinks they would have agreed.

13. It is not true of everything that we call a transaction cost that its removal would make both parties better off. The cost to one of the parties of supplementing incomplete information about the other’s preferences is a transaction cost, but removing it may well mean that the parties negotiate toward a deal that is worse for the other party than the one the parties would have struck if the information had not been provided. I think therefore that Coleman is wrong to criticize Buchanan for failing to distinguish between transaction costs as such and other aspects of the exchange framework. P. 93. The important distinction is between aspects of the exchange framework whose alteration would leave both parties better off and aspects whose alteration would prejudice the interests of one or the other of the parties.
That last sentence of course is incomplete: "to which . . . they would have agreed" should be followed by an "if . . ." specifying the exact content of the hypothetical. It is surprising that, in Coleman's discussion throughout Chapter Eight of *Risks and Wrongs*, he invariably allows phrases like "what the parties would have agreed to" to stand on their own (pp. 165-67, 169-70, 175-76, 179, 181-82). As far as one can tell, he means "would have agreed [to] under conditions of full information" (p. 174). That phrase sounds innocuous enough, until one recalls the view he attributed to Buchanan — that transaction costs, including the present distribution of information, are part of the framework of exchange.14

Information may be regarded as a commodity, which can be equally or unequally distributed. If we were actually providing each of the parties with full information, we would necessarily be ensuring that they had equal information, and we would thereby be engaging *pro tanto* in the redistribution of wealth. Offhand it is not clear to me why imposing on them the terms to which they would have bargained if they had both been in possession of full information is not *also* redistributive, at least relative to a situation in which information is unequally distributed in a way that affects the deals to which the parties, respectively, are willing to agree. If you have spent $100 to ascertain that my antique watch is worth $1000 and I am willing to sell it to you for $500 because I am too mean to have it valued, why should the court force upon us the contract to which we would have agreed if we had both been in possession of the information for which only you were prepared to pay?15

I mention this difficulty mainly because Coleman himself contrasts enforcing the terms to which the parties would have agreed with enforcing terms that strike the court as fair or distributively just (p. 181). The latter option can easily appear attractive. Why should the court reinforce the parties' bargaining inequalities? Why not seize the opportunity to improve the background distribution of wealth? Coleman's answer to these questions is, I think, persuasive. If the courts announce in advance that, in any case of contractual uncertainty that comes before them, they will impose whatever terms make the background distribution of wealth more just, then they will inhibit, not facilitate, market transactions. The imposition of redistributive terms may well make one of the parties worse off than he would have been absent any contract. That party then will have reason to avoid

14. *See supra* notes 12-13 and accompanying text.

15. After reading and rereading the appropriate sections, I am unsure whether Coleman recognizes this difficulty or not. He certainly makes the connection with the Buchanan point: "What [does] the agreement among ideal agents under ideal circumstances have to do with the circumstances in which the parties find themselves?" P. 178. But he leaves the question hanging and the reader crying out for signposts to indicate whether it is rhetorical or, if not, how he proposes to answer it.
contracts altogether, or, if they are unavoidable, he will have reason to invest very heavily in safeguards and gap-filling measures to avoid any possibility of recourse to the law (pp. 181-82). In either case, we are back where we started as far as contractual transactions are concerned.

However, exactly the same argument will apply to at least some cases in which the courts impose the terms to which the parties would have agreed if they had had full and thus equal information. So Coleman's position needs to be stated more carefully. I suggest this: the courts should impose the terms to which the parties would have bargained absent those transaction costs whose removal would leave both parties better off. Not all uncertainties or transaction costs will satisfy this condition. But, if we ignore the condition, we will be doing exactly what Buchanan condemned: using the legal response to market failure as an opportunity to alter the framework of exchange to the detriment of one of the parties.\(^\text{16}\) If we alter the framework in this way, then, by Coleman's own argument, we make it less likely that parties will reach agreements.

I have one other point to make about Coleman's account of contract law. It seems important for him that legal liability for breach of contract be perceived as just another way in which the law helps to reduce transaction costs. He offers two arguments for this characterization. First, he asserts that "the source of breach is often an unresolved uncertainty" (p. 147); thus, the adjudication of an alleged breach often amounts to filling in the terms of an incomplete or uncertain contract. Second, even when uncertainty about terms is not involved, parties will not enter into mutually advantageous bargains when there is a substantial possibility that their partner's defection from the arrangement will leave them worse off than they were when they began; the cost of safeguarding against such defection can be seen then as a transaction cost in relation to the bargain in question (p. 130). Both points are important. Coleman develops the second point, in particular, in a way that illuminates the relation between the defection problem and the other problems that define the possibility of a bargain — initial coordination and the division of the surplus, including the division of the costs of safeguarding against defection (p. 130).

Yet Coleman's account seems to leave something out. When a person is held liable for a clear breach of contract and either damages are awarded or specific performance is required, there is a sense that an obligation is being enforced and, indeed, that justice is being done, in at least one sense of that protean term. Is there no connection whatsoever between contract law and the moral principle of fidelity to promises? Does the old maxim *pacta sunt servanda* intimate nothing

---

16. See supra notes 12-13 and accompanying text.
in the way of a principle of justice? Is it purely a matter of addressing the problem of transaction costs?

I do not want to be misunderstood. There is no doubt that the existence of "a system of enforcement with which to preclude or to deter noncompliance or to compensate parties injured by others' noncompliance" (p. 118) is one very important way in which law helps to facilitate the operation of markets, just as the existence of a system of tort law can be seen as a way of promoting efficiency. But, whereas in the latter case — as we shall see — Coleman wants to go further and insist that tort law also — indeed, primarily — serves and embodies principles of corrective justice, he is not prepared to say anything of the sort about contract law. The asymmetry is all the more mysterious in that his account of corrective justice identifies not only the infringement of rights but also various forms of wrongdoing relative to local market norms, such as unfair competition and so forth, as a basis of tort liability. Isn't breach of contract the most fundamental form of market misconduct?

Why, then, is there not a form of justice specifically oriented toward a party's obligation to do what he has promised to do, or at least to leave his contractual partner in the position in which he would have been if the contract had been carried out?

I have no idea how Coleman would answer these questions. The asymmetry I have indicated, between his account of tort and his account of contract, is unqualified in Risks and Wrongs by any gesture toward the idea that keeping one's contracts might be regarded as a matter of justice. As a result, Coleman's account leaves much of contract law more or less completely unilluminated — I refer particularly to the doctrine of consideration, the issue of third-party beneficiaries, and other aspects of contract law that can barely be understood except as the working out of principles of justice specifically oriented toward the issue of contractual fidelity.

---

17. See infra Part III.

18. Coleman briefly mentions contractual obligation with regard to the issue of imposing the ex ante agreement. Pp. 166-70. But the issue of the existence or nonexistence of obligations is treated there as though it were solely an issue of how best to facilitate markets, in a manner that Coleman would not accept for a moment in an analogous discussion of torts.

19. The only time the issue is touched on at all is at the end of the book, when Coleman dwells briefly on the distinction between the "forward-looking" approach he takes to contracts and the "backward-looking" approach he takes to tort. Pp. 430-32. He argues that the forward-looking approach illuminates issues in contract in a way that it would not illuminate issues in tort. But this is not sufficient to rule out the possibility that there are other aspects of contract doctrine that would be illuminated by a backward-looking approach.

Two other opportunities — missed opportunities — for discussion of this issue stand out: (i) Coleman talks about the reputational effects of contractual defection (pp. 142, 144), but he offers no account of the fact that such effects depend on contract breaking being regarded in most cases as a form of wrongdoing; and (ii) as we have seen, Coleman talks about distributive justice as a possible basis for default rules (pp. 181-82), but he offers no account of the possible relations between justice-in-contract and distributive justice analogous to the complex and interesting account of corrective justice and distributive justice that he offers toward the end of Risks and Wrongs. Pp. 350-54; see infra Conclusion.
III. TORT LAW AND CORRECTIVE JUSTICE

Though it is presented as a general account of the relation between law and markets, *Risks and Wrongs* is mainly a book about torts. Part III, entitled “Rectifiable Wrongs,” is longer than the other two parts put together. Like the rest of the book, it is heavy going — again, mainly because Coleman invests a lot of energy in explaining exactly how his view differs from those of other tort theorists, notably Ernest Weinrib, Richard Epstein, and George Fletcher, as well as those who approach the matter within the confines of the market paradigm. It is not my task in this review to explore these differences. Coleman’s own position, if I understand it correctly, can be stated quite succinctly.

A market consists of a structure of rights — for example, property rights — and a structure of local norms indicating what the community regards as acceptable and unacceptable ways of doing business. If these norms are breached or if rights are infringed, the individuals whose rights they are — or whose interests are protected by the norms — may suffer loss. There are, of course, other ways of suffering losses in the marketplace — for example, loss of market share through ordinary processes of competition. But such losses are not necessarily wrongful. A wrongful loss, according to Coleman’s account, is either a loss accruing from the infringement of a right or a loss that is someone’s fault in that it accrues from that person’s failure to observe some local norm of market conduct (pp. 330-60).

So far all this gives us is a taxonomy: losses are either wrongful or not; among those that are wrongful, some derive from wrongs (infringements of rights) and others from wrongdoing (violation of norms). The category of wrongful losses, however, interests us for the following reason, according to Coleman. There is something called corrective justice — a distinctive body of important moral truths — that establishes connections of principle between wrongful losses and the agents responsible for them. The connection is that the agent who is responsible for a wrongful loss has a duty to repair it, usually by ensuring that the person who has suffered the loss is compensated. This duty exists whether or not the activity that occasioned the loss — or the activity-plus-compensation — promotes efficiency. The existence of this duty is a matter of justice. It is grounded purely on the basis of backward-looking considerations — that is, purely in virtue of the responsibility connection between the agent in question and the wrongful loss that has been suffered (pp. 361-85).

Two qualifications must be added to round out the picture. Though the connection between wrongful loss and an agent’s liability

is a matter of justice, it is not unconditional. A society may have other practices that make good the losses that people suffer from wrongdoing or rights infringements: a society may have institutions of social insurance to cover personal injury, for example, along the lines of the New Zealand Accident Compensation Scheme. If such a practice exists and if its effect is to make good the losses to which corrective justice would otherwise apply, the justice-based reasons for holding that the person who inflicted the loss has a duty to repair it are superseded or extinguished.

This possibility is matched by another contingency concerning the legal enforcement of duties of corrective justice. Corrective justice, as Coleman describes it, is a matter of moral principle. There is therefore a law-morality issue about the embodiment of corrective justice in legal norms. Coleman argues that liberal principles permit the law to enforce duties that require injurers to make good the wrongful losses they inflict, but do not require the law to do so. The law might do nothing in the matter — leaving the injurer's duty to compensate roughly on a par with his duty to apologize — or it might enforce a New Zealand-style scheme, or it might implement alternative policies promoting the efficient spreading of risk. In fact, according to Coleman's account, the permissibility of either of these last two options is overdetermined. On the one hand, the law need not enforce what are merely individual moral duties arising out of corrective justice. However, if the legal system does decide to deal with losses in some other way, even the moral duties arising out of corrective justice are extinguished so that the question of their legal enforcement becomes moot.

As things stand, however, Coleman believes that the legal systems of the United States do embody something like corrective justice, and that many of the central features of tort law can only be understood in this light (pp. 374-75). He concedes that not all aspects of contemporary tort law make sense in these terms. Our tort law is best understood as a hybrid: while some of it enforces agents' moral duties to repair wrongful losses, "other parts of tort law may reflect alternative approaches to allocating losses, both wrongful and other, in ways that sever the relationship between agent and loss under corrective justice" (p. 434). Coleman concludes that although tort liability in fact embodies corrective justice to a considerable degree, there is no affront to justice or political morality in the commonly heard proposal that we should abandon tort altogether and replace it with a no-fault system of social insurance (pp. 401-06).

What I have outlined is Coleman's position, as I understand it. It is not possible to do justice to its every subtlety, even in a review of this inordinate length. It is possible that I have got parts of it wrong: Coleman's presentation, as I have said several times, makes it hard to
be sure that one has got hold of his view. I want to raise some questions, however, about the *kind* of view that Coleman takes himself to be presenting.

Those questions are best evoked by considering Coleman’s main criticism of the account of torts offered in the market paradigm (pp. 374-80). According to this account, tort liability is merely an instrument in our pursuit of efficiency. Suppose *P* is injured in an accident. In the market paradigm, if we hold *P* to have the right to recover damages and if we impose a correlative duty on the individual, *D*, who injured him, we do so because this is the most efficient way of spreading and influencing the risks associated with activities like the one in which *D* engaged. That, according to Coleman, is a purely instrumental basis for bringing *P* and *D* together as plaintiff and defendant (pp. 203-04). In the market paradigm, he says, there is no reason *in principle* why actions in relation to the imposition of risk should be initiated solely by accident victims, nor is there any reason *in principle* to single out the person who caused the injury as the one to bear the loss or as the one to make an example of to influence risk imposition by others. “In economic analysis,” says Coleman, “the inclusion of injurers and victims is fundamentally contingent and does not in any way depend on their relationship to one another” (p. 380). Since it is obvious, however, that the nexus between victim-plaintiff and injurer-defendant *is* a matter of principle in tort, the economic account fails as an interpretation of this part of the law.

As a matter of fact, I think this criticism of the market paradigm is unfair. Recall the picture of tort law painted on behalf of the market paradigm at the very beginning of this review.21

*D* wishes to engage in some activity — driving, construction, and so forth — that may possibly affect the person or property of various individuals *P*₁,..., *P*ₙ. *D* reckons he stands to benefit so much from this activity that even if it does harm a particular member of this group (*P*ᵢ), *D* will be able to pay enough to make it worth *P*ᵢ’s while and still be better off. In principle, then, *D* could reach a mutually beneficial agreement in advance with *P*ᵢ. Unfortunately, he does not know in advance who *P*ᵢ is — that is, he does not know which of the *Ps* will be affected or if any will. Moreover, it is too costly for *D* to secure the conditional consent of all of them in advance. So he goes ahead with the activity, in effect imposing the risk on the various *Ps* without their consent. On the economic account, the law of torts allows *D* to do this, provided he compensates anyone he does in fact injure.

Now suppose *D* actually injures *P*ᵢ. Then, on the economic account, he owes *P*ᵢ roughly what he would have agreed ex ante to pay

and what \( P_i \) would have agreed ex ante to accept in the event of this injury occurring, if transaction costs and uncertainty had not prevented such an agreement from taking place. That hypothetical agreement — the agreement that in fact was precluded by transaction costs — is thus the basis of the legal relation in tort between \( D \) and \( P_i \). On this basis, they are related as directly as they would be in the hypothetical agreement itself. To be sure, the relation between them is a contingent one. But it is contingent only on the occurrence of the impact of \( D \)'s activity on \( P_i \) — and that of course is exactly the contingency on which the relation is dependent in any plausible account of tort liability. It is thus simply not true that "in economic analysis, there are no principled reasons for connecting injurers and victims in the way in which tort law does" (p. 381).

Where Coleman has been misled is in his understanding of the further effects of this liability on \( D \)'s behavior. Not knowing in advance which and how many of the \( P_i \)s his activity might affect, \( D \) will be inclined to limit and modify the activity in various ways. He will consider the probability of its affecting people as the probability of his having to compensate them when it does, and he will adjust his activity so that either the probability is diminished, the harm done by a particular impact is diminished, or both, so that the expected value of his loss — through having to pay compensation — never exceeds his expected gain from the activity. These calculations amount, in short, to his taking care. Maybe the incentive effect of all this is something over and above the nexus between a particular plaintiff and a particular defendant. But that nexus is still the infrastructure of the theory, and it is not difficult to represent what appears to be a purely external and instrumental deterrence effect as the shadow cast, in a world of transaction costs, by the bargaining that would have led to actual agreements in an ideal world free of such costs.

To summarize, then: The economic account presents tort law as a response to market failure. But it is a response, not to market failure in the abstract, but to particular market failure as between particular individuals. We talk about market failure when there are two persons who could have entered into a mutually advantageous bargain but did not because of transaction costs. They are the ones who are now related in the law's pursuit of that efficient outcome, and they are the ones who may appear in a tort suit as plaintiff and defendant, respectively.

This is, however, not my main point. Suppose I am wrong in what I have said in the last four paragraphs and Coleman is right that the market paradigm cannot provide anything other than a contingent or instrumental account of the nexus between victim-plaintiff and injurer-defendant. Then we must ask: Is the account offered in Coleman's theory of corrective justice any better? In one way, it is obviously
better. Coleman's principles of corrective justice directly entail that victim-plaintiffs have a right to recover from injurer-defendants. Not only do his principles entail this, but the victim's right to recover from the injurer pervades his entire theory. There is no part of the theory that is not focused on this injurer-victim relationship. In fact, to be honest, the claim that the relation between the person responsible for a wrongful loss and the person who suffers it is central, important, and a matter of principle is pretty much all there is to Coleman's account of corrective justice.

Someone — probably Coleman — will object: "That can't be all there is to the theory. One does not need two hundred pages simply to announce that agents responsible for wrongful losses have a duty to repair them and that such duty goes by the name of 'corrective justice.'" Of course there is much more going on in Part III of Risks and Wrongs. Coleman has to distinguish wrongful losses from non-wrongful losses. He has to distinguish wrongs from wrongdoing and explain that the duty to repair the former is not based on fault. He has to give an account of what "responsible for" means in the context of wrongful losses. And he has to set out in detail how all of this differs from the positions expounded by Weinrib, Fletcher, and Epstein.

The fact is, however, that Coleman's performance of these tasks is not guided by any attempt to make sense of corrective justice or to explain why it imposes the duties that it does. Over and over again, Coleman uses phrases like "Corrective justice requires . . ." or "Corrective justice imposes a duty to . . ." without explanation.22 One is left with the impression that Coleman believes there are such duties and requirements — which he labels duties and requirements of corrective justice — but which he is unwilling or unable to explain. It seems a little unfair, then, to castigate the market paradigm for failing to account for something that is characterized in Coleman's theory by bare unexplained assertion.

Coleman is well aware of this point. In his study of tort liability, he is engaged, he says, in "middle-level" rather than "top-down" theory (pp. 6-11). The top-down approach is the sort of theory he uses in analyzing contracts: contract law is explained in Part II of the book in terms of its serving purposes and values, such as the value of decentralized decisionmaking in a heterogeneous society. We are in a position to appreciate that value apart from the way in which contract law serves or embodies it, so the articulation of that value can be deployed as a genuinely independent basis for explaining why the principles of contract law are as they are. The value provides what John Rawls refers to as an "Archimedean point" for making sense of detailed social institutions.23 Middle-level theory, in contrast, does not appeal to

lofty, independent values to explain or justify our practices: "In middle-level theory, the theorist immerses herself in the practice itself and asks if it can be usefully organized in ways that reflect a commitment to one or more plausible principles" (p. 8).

Coleman acknowledges that "middle-level theory" faces the danger that it will simply get lost in the practice that it is characterizing, though he thinks that there is an even graver danger that the principles "discovered" by a middle-level theorist will be "advanced less because they inhere in the practice than because they reflect the theorist's favored foundational view" (p. 9). I think the former danger is the one worth emphasizing. If a given practice is articulate — as most human practices and certainly all legal practices are — there is a danger that middle-level theory will amount to nothing more than an unexplained repetition of various incantations customarily associated with participation in the practice itself. Coleman's formulations of corrective justice often have this character. Many who are involved in the practice of tort law — such as judges, attorneys, or litigants — go around saying things like $S_1$: "A person who is responsible for another's wrongful loss has a duty to compensate him for that loss." Although this statement is typical of practitioners in the area, it is not clear to me how much one illuminates the practice by simply prefacing statements like $S_1$ with "Corrective justice requires . . . " and calling the result a middle-level theory.

Maybe what is illuminating about Coleman's position is the following: Coleman believes that many theories of tort law — such as economic analysis — treat $S_1$ and other statements like it as though they were peripheral or unimportant in the practice. Middle-level theory can be a useful response because it insists that it is impossible to make sense of torts from the inside, as it were, without regarding $S_1$ as central. The middle-level contribution is to show that any good theory must explain the centrality of $S_1$. That may be a protracted and argumentative contribution if there are other middle-level theorists, like Fletcher, Weinrib, and Epstein, going around saying that the truly central proposition that any good theory must explain is not $S_1$ at all but some alternative proposition $S_2$. In the end, however, to do this is to do no more than specify a theoretical agenda: we have identified $S_1$ as a central feature of the practice that needs explanation and defense, but now we must go beyond middle-level theory to actually explain and defend it.24

My sense of the incompleteness of Coleman's enterprise is heightened by the view — which he himself shares (p. 221) — that the pro-

---

24. Coleman, I think, acknowledges this necessity in an endnote, if I understand the following passage correctly: "In order for one's middle-level account of a practice ultimately to be fully satisfying and complete, the theorist must argue that the norms he identifies in the practice could be accepted by reasonable individuals engaged within the practice . . . ." P. 441 n.2.
positions he assembles under the heading of “corrective justice,” although familiar, are not necessarily plausible or attractive. It is not hard to make the basis of tort liability look very arbitrary. As the price of a moment’s carelessness on the road, a driver may be required to pay over a massive sum of money: “[T]he wrong can be very slight and the damage enormous; or the wrong can be grave and the damage minuscule. In either case, making the victim whole will be unconnected to the extent of the wrongdoing” (p. 235). Another driver who was careless in exactly the same way may go free of any liability because of the mere happenstance of his not having collided with anyone: “One might object to luck playing such an important role in allocating accident costs” (p. 286).

Coleman takes features like these to be indications of the inappropriateness of applying to tort law standards of fairness or intuitions of arbitrariness that are more at home in criminal law. When approaching tort theory, he insists, “[w]e must begin by forgetting just about everything we have ‘learned’ from studying the philosophy of the criminal law” (p. 222). The principles of justice applicable to tort liability are principles of corrective, not retributive justice; they must be made defensible therefore on their own terms, not on terms appropriated from some other practice or legal domain.

I am sure Coleman is right to stress the differences between tort liability and criminal liability. Apart from anything else, the aim of compensating victims is at best secondary or incidental in criminal law, whereas it is central in torts. It would be wrong, however, to

---


26. This argument might suggest an approach to justificatory theory similar to that of Michael Walzer: different goods are imbued with social understandings that make different kinds of distributive principles appropriate to them, and it is always a mistake to evaluate the distribution of goods of one kind on the basis of principles appropriate to goods of another kind. See MICHAEL WALZER, SPHERES OF JUSTICE: A DEFENSE OF PLURALISM AND EQUALITY 6-20 (1983). Adapting this approach, one might say simply that our understanding of wrongful losses implicates principles that are sui generis, for which it is a mistake to seek any further foundational justification.

Certainly there are aspects of Coleman’s terminology that mirror Walzer’s. Coleman, like Walzer, insists that each social institution is best understood from the inside: “At the heart of middle-level theory is the conviction that one has to immerse oneself into a practice in order to understand the norms that inhere in it.” Pp. 199-200. But three things make me doubt whether Coleman subscribes to Walzer’s full-blooded antifoundationalism and the social ontology that underlies it. First, he does not embrace Walzer’s position explicitly. Second, he does not rule out a foundational theory of tort law, although he doubts whether it would be a theory of the same kind as the one he sketches for contract. P. 12. Indeed, in at least one place he suggests that a defense of his conception will be incomplete until a foundational account has been given. P. 478 n.1. Third, he acknowledges that society must make a choice between institutions embodying corrective justice and other ways of dealing with losses from accidents, and he gives no indication that this choice cannot be made for reasons. If there can be reasons for this choice, then there can be a theory of such reasons. It is the omission of this dimension of theorizing that one feels most keenly in struggling through Coleman’s account.
infer that individual moral desert, broadly construed, has no place in our understanding or assessment of tort liability. Desert and the associated criterion of fairness are not concepts that can be so easily confined. Think of the variety of contexts in which we say intuitively that something has gone awry because there are enormous differences in people's well-being that are unrelated to any sense of their comparative desert. The idea that how things go for a person should bear some sort of relation to how well he has acted informs our assessment of the world on all sorts of fronts. We use it to rail against God, against nature, against disease; we do not even confine it to the operation of human laws and institutions, let alone to any particular legal domain. In fact, it is natural to invoke standards of fairness for assessing social arrangements of all sorts. We commonly argue about whether students get the grades they deserve, about whether the most deserving team won the World Series, about whether entrepreneurs deserve windfall profits, about who deserves the higher salary in a law school, and so on. In none of these cases is an appeal to personal desert the officially determinative factor, and perhaps that is the main contrast with a system of retributive justice. But, in each case, systematic and sustained dissonance between outcomes and deserts would be seen as a problem for the institution in question and an appropriate basis for criticism.

To put it another way, a gaping disproportion between individual outcomes and the morality of individual characters and conduct seems to us an elementary form of injustice. It is true that, since Aristotle, our concept of justice has been divided among various domains — distributive justice, corrective justice, retributive justice, commutative justice, and so on — and we must expect detailed conceptions of fairness and unfairness to observe these divisions.27 Still, we must not let our familiarity with such distinctions blind us either to the connections that exist between the different domains of justice or to the possibility that there are overarching principles or intuitions of fairness and desert that these domains share in common and that entitle both corrective justice and retributive justice, for example, to be regarded as forms of justice. There is therefore a fairness challenge that must be met with regard to tort liability, and a theory cannot count as meeting it if all it does is carefully articulate the principles whose application poses the challenge in the first place.

Much the same can be said about Coleman's characterization of strict liability and of the limited role that excuses play with regard to tort liability (pp. 224-29). A person may be held liable for the wrongful loss that another has suffered even though it was not his fault — in any sense other than that his activity happened to be its cause. Such a

---

case makes tort liability look very difficult to explain or defend from the point of view of fairness.

Coleman does not attempt to evade this challenge. Indeed, he adopts an analysis in which such cases are central rather than peripheral. Strict liability for wrongs — that is, for infringements of others' rights — is for him one of the two main grounds of an agent's justice-based duty of compensation (p. 332). But facing the challenge is not the same as answering it. Coleman's conception of corrective justice certainly requires some agents who faultlessly infringe others' rights to compensate them for the resulting harm. But saying nothing more than that corrective justice requires it is not a way of dispelling the sense of great unfairness that cases like these exhibit.

In the case of strict liability, Coleman does make one attempt to make corrective justice palatable. Though strict liability for the injurer seems unfair, strict liability for his victim seems even worse. In tort law, Coleman argues, "any decision is a liability decision: losses always must fall on someone . . . Not to impose liability on the injurer is to impose it on the victim, and vice versa" (pp. 223-25). Certainly any liability imposed on the victim is faultless liability; and, if we have to choose between these unpleasant options, it may seem right to give decisive weight to the fact that the injurer's faultless conduct, rather than the victim's faultless conduct, was the cause of the loss (p. 224).

It seems odd, though, to describe the alternative to imposing liability on the defendant D as an imposition of liability on the plaintiff P.

28. The hardest cases for a fairness-based theory are those like Vincent v. Lake Erie Transportation Co., 124 N.W. 221 (Minn. 1910), in which the court held D liable for damage to P's property even though D was fully justified in acting as he did on grounds of necessity. 124 N.W. at 222. It seems unfair to require D to make up for something he was justified in doing; but, if Coleman is right, such a requirement is a paradigm application of corrective justice.

Coleman's analysis of cases like Vincent is interesting. He uses the distinction between wrongs and wrongfulness to carefully separate the issue of whether a right was infringed from the issue of whether the agent was justified in infringing it under the circumstances. Pp. 299-302. The potential problem with this analysis is that the existence of a justification for D's conduct threatens to undermine the claim that he infringed P's right. If D was justified in what he did, then he cannot have breached any duty. But we cannot understand the infringement of a right except in terms of the breach of a duty. Coleman recognizes this difficulty and devotes a long endnote to some further reflection upon it. Pp. 475 n.7. I think, however, that his answer is a good one. P's right imposes a very limited duty on D, namely a duty to secure P's consent before using P's property. That duty does not simply evaporate in the presence of D's justification for the immediate use of P's property, although it is fair to say that in the circumstances it takes second place. Still, at every stage in the process, even when the circumstances that give rise to the justification apply, D remains under a duty to seize any opportunity to secure P's consent. D's justification, in other words, does not give him a right to act as though P's stake in the matter did not exist. P's right and the duty correlative to it continue to exercise gravitational force on D's actions even when D is justified in infringing that right. For a general version of this thesis, see Jeremy Waldron, Rights in Conflict, 99 ETHICS 503, 509-12 (1989), reprinted in Jeremy Waldron, Liberal Rights 203, 211-15 (1993).

I assume, then, that Coleman's corrective justice position takes D's duty to compensate as yet another aspect of the gravitational force of P's right that survives D's justification in infringing it. I must stress again, however, that the persuasiveness of this as an analysis does nothing to explain or justify the imposition of liability in such cases.
Liability is a two-term relation: if \( D \) were held liable, he would be held liable to \( P \). To whom, then, on Coleman's account, is \( P \) held liable, if his suit against \( D \) fails?

Logic aside, the claim about symmetry of liability here is quite misleading. A court that refuses to hold \( D \) liable for \( P \)'s losses makes no statement about the desirability or appropriateness of \( P \)'s bearing that loss. Coleman argues that, by saying that \( P \) is held liable for the loss in the absence of a judgment against \( D \), he means no more than that \( P \) has been determined to be in the better position to shoulder the costs occasioned by the accident (p. 231). But, even if this account is true, it confuses the meaning of the normative term *liable* with a particular — and controversial — theory about the proper criteria for its application.29 Anyway, it is wrong to assume that, in every case in which a judgment is not entered against an injurer-defendant, a comparative determination has been made about his capacity versus the plaintiff-victim's capacity to bear the losses occasioned by the accident. \( P \)'s case may fail for many reasons that have nothing to do with this determination.

My inclination is to say that if \( D \) is not found liable — and if nothing else is done to compensate \( P \) — then the loss just lies where it naturally fell as a result of the incident. The law has simply decided to do nothing about that state of affairs. Coleman contests this description. He argues that the fact that the loss falls initially on \( P \) is a normative conclusion of law, not a statement of natural fact. "[T]he loss begins," he writes, "as the victim's 'legal responsibility'" (p. 232). This description turns out to mean no more than that it is up to \( P \) to initiate the process of recovering damages from \( D \) and that, until he acts and makes out whatever case is required, \( D \) will not be held liable. Now that is indeed a judgment about \( P \)'s legal position. The proper way to describe it, however, is as follows: (i) initially, and apart from the grounds for the commencement of any lawsuit, \( P \) has what Wesley N. Hohfeld called a no-right against \( D \) correlative to \( D \)'s lack of any background duty to pay \( P \) compensation;30 in the circumstances, however, \( P \)'s no-right is coupled with (ii) a Hohfeldian power to change that state of affairs by bringing an action in tort, such power being correlative to \( D \)'s liability to have a duty imposed on him.31 Thus, there certainly are legally defined aspects to \( P \)'s situation immediately after the incident. However, neither (i) nor (ii) nor their combination can plausibly be described as \( P \)'s having an initial liability to \( D \) or any-

---

29. For the distinction between meaning and criteria in the case of normative terms, see R.M. Hare, *The Language of Morals* 94-110 (1952).


31. *Id.* at 36, 39.
one else. It is therefore not accurate to say, as Coleman does, that “[a]s tort liability is currently structured, liability may unavoidably fall to either an innocent victim or an innocent defendant” (p. 286).

Coleman is on much firmer ground when he acknowledges simply that “[t]ort law is unjust either to victims or injurers, probably to both” (p. 221). In a case of strict liability, either the victim is left suffering a huge loss that was not his fault, or the injurer is required by law to bear the cost of a huge loss that was not his fault. Moreover, Coleman can repeat his earlier attempt to address the matter in terms of comparative unfairness without all the misleading stuff about plaintiffs’ liability. “There may be . . . even less justification for letting the loss remain the victim’s responsibility than for shifting it to the injurer,” particularly if, as Coleman says, “we frame the question of who is to bear the costs of a particular injury to exclude everyone other than the victim and the injurer” (p. 224).

But why should we frame the question in that way? Given that it seems arbitrary to require either the injurer or his victim to bear the whole loss, and given that the morally relevant difference between them — when they are both faultless — seems vanishingly slight, why not simply abandon the framework that confines our attention to the two of them and seek a way of addressing accident costs that does not involve arbitrary elements of this sort? That, after all, is the crucial question that we expect tort theorists to address. Is the traditional nexus between plaintiff-victim and injurer-defendant, which lies at the heart of tort law, sufficiently fair, sufficiently nonarbitrary from a moral point of view, to count as a better way of addressing the problem of accident costs than, say, a New Zealand-style scheme that abandons the principle of individual liability altogether?

Unfortunately, Coleman has precluded himself from addressing this question. It seems to be part of his conception of corrective justice that there are no justice-based grounds to guide the choice between corrective-justice institutions and other institutions for dealing with the losses that people suffer. As we saw at the beginning of this Part, the duties of compensation to which corrective justice gives rise are understood by Coleman to be conditional on the absence of other ways of dealing with these matters. Because other may mean “fairer” or “less arbitrary” — as well as “more efficient” or “more popular” — there is no basis on which a philosopher of corrective justice can present tort liability in an attractive light relative to the real alternatives that we face as a society. Such a philosopher can tell us what tort liability is and what its main principles are, but he cannot say anything about why it is worth having.

**CONCLUSION**

I could just conclude with this observation. Coleman himself
Liberal Markets

seems unembarrassed by the fact that the account of tort law to which more than half his book is devoted is entirely middle-level theory and is not connected in any systematic way to the liberal values — markets, social diversity, stability, individual autonomy — that dominate Parts I and II of *Risks and Wrongs*. However, he does make one or two gestures toward a top-down account, and it would be churlish to end a long review without attempting some sort of Colemanesque sketch of the liberal significance of corrective justice and tort liability and their relation to the theory of markets and cooperation that dominates the rest of the book.

In an introductory chapter, Coleman says that despite his general belief that law can “help to create and sustain the conditions under which markets can flourish and contribute to stability,” his view is that “tort law is not primarily an institution designed to create or sustain markets” (p. 12). I believe this formulation is misleading.

Think back, for a moment, to what constitutes a market and what it means to create and sustain one. A market exists when there are a large number of people with alienable resources who are ready to make bilateral deals with one another to allocate and reallocate those resources to various uses. As I pointed out earlier in this review, a market cannot exist without rights; creating and sustaining a market is partly a matter of creating and sustaining rights. To sustain a right is to ensure as far as possible that the duties it generates are observed and to do something about the situation when they are not. It seems obvious that tort law, understood as corrective justice, is a way of performing this last task. It is a way of vindicating rights against certain sorts of infringements by requiring rights-infringers to make good the losses caused by their actions.

Coleman accepts, as we have seen, that this is not the only way of vindicating rights against loss: there could be a scheme of social insurance to make good the injuries and damage that people suffer on account of one another’s actions. He insists, too, that it is not necessarily part of the concept of a right that there be a remedy in damages for interference with it (pp. 335-42). In some cases, it may be part of the content of a right that interference or forced transfer is permissible provided compensation is paid. Usually, however, the right simply prohibits a person from impacting a certain resource without the owner’s consent, so that an unconsented-to impact counts

32. See *supra* Introduction.

33. I stress that tort law is a way of vindicating, not necessarily a way of protecting, rights against infringements. Coleman is right to stress that tort liability is not justified primarily on grounds of deterrence. The point is that when a loss is caused, the fact that it is caused to a right means that something must be done about it; otherwise having a right would not be much different from not having a right. I should emphasize, too, that the language of “vindicating rights against infringement” is mine, not Coleman’s.

34. See *supra* Part III.
as a wrong against the right-bearer. In these cases, the right-bearer appeals to whatever social practice we have of dealing with the mess resulting from rights-infringements. Tort liability, embodying principles of corrective justice, happens to be one such practice.

Can we say anything, then, about why we should choose this practice rather than some other as our way of dealing with rights-infringements? Is there any advantage, from the point of view of Coleman's liberalism, for choosing tort and corrective justice as the institutional basis on which rights are vindicated — and markets sustained — in our society?

We value markets, on Coleman's account, to the extent that we want to economize on centralized decisionmaking in society. The virtue of encouraging bilateral transactions as the basis of the routine reallocation of resources is that we do not always have to come together to agree on social purposes every time a resource is moved from one use to another. But accidental impacts and invasions of various sorts are also part of daily life, and it would undermine the advantages of markets if dealing with these events normally required us to make centralized decisions about which allocation of losses and costs would best promote social goals. Better to have a legal framework for coping with these vicissitudes that deals with these matters as far as possible on a decentralized basis.

One could argue, then, on Coleman's behalf, that this is a good reason for rejecting legal responses to individual loss that involve any sort of global determination of social advantage, such as a determination of overall efficiency or a determination of what would count now — in the light of the accident — as a just distribution of resources. The advantage of tort liability, from this point of view, is that it deals with the matter solely in relation to the particular parties involved, without requiring us to agree on any broader picture of what would count as a desirable way of organizing or reorganizing the deployment and distribution of society's wealth.

Coleman comes close to this argument in one extremely interesting passage (pp. 350-60), when he considers the relation between corrective justice and distributive justice. The question he asks is this: If the basic distribution of resources in society is unjust, why should we assume that a person who damages another's holding is under a duty of justice to make good the loss that results? The problem is easy enough to deal with if we are simply talking about tort liability in law: the law of torts takes for granted the legally sanctioned distribution of rights and asks no questions about it. But corrective justice, as Coleman conceives it, is in the first instance a matter of moral duty, and so we face this question: Can one have a moral duty to repair a loss that one has caused to a holding that is unjust?

Coleman is sure that the moral aspect of corrective justice imposes
some constraint on the distribution of resources that it helps to sustain, but he appears uncertain about how to describe that constraint. At one point, he says merely that the losses to which corrective justice is a response "ought to be of the sort it makes moral sense to be concerned about" (p. 350). This could apply to any established distribution, given that we have some concern about people's de facto expectations. A little later, he seems to stake out a stronger position: "The rights sustained by corrective justice must be real rights . . ." (p. 352). But he goes on to say that they "need not be rights that can be shown to be defensible within the best theory of distributive justice" (p. 352). His considered opinion seems to be that an entitlement can be unjust vis-à-vis full-blooded redistribution done in the name of the state, but just vis-à-vis the initiative of an individual tortfeasor or an individual Robin Hood. If a holding satisfies this condition, then it may be protected by principles of corrective justice: "[T]he state may have . . . authority to reallocate resources in a way in which individuals do not, even if certain individual reallocations might have the effect of making a more distributively just world" (p. 352).

This approach has the attraction of offering individuals some guarantee of their holdings and their expectations that could not be offered on the basis of a theory of distributive justice alone. To put it provocatively, individuals have a certain amount of protection against distributive justice: redistribution is to be undertaken only when the whole community is prepared to mobilize; it is not to be undertaken on the basis of individual initiatives, however well intentioned. This protection and predictability in turn serve both social and individual ends. They serve the social end of stability inasmuch as society will not be riven by the need for reassessments of distributive justice every time anyone invades another's established holding. They also serve individual values of autonomy by establishing a stable framework for action and by contributing to "each agent's capacity to anticipate the behavior of others that might be disruptive to their ability to act on plans they have formulated" (p. 360).

I am aware that these remarks do not provide a full account of the connection between corrective justice and the liberal values that Coleman associates with markets. Furthermore, I have said almost nothing about a subtheme of his book that he himself regards as very important: the idea of liability for wrongdoing in relation to local market conventions, even when rights are not at stake.35 It is impossible in any review to do justice to every aspect of a book, and it is certainly impossible to do anything like justice to the complex web of

35. Pp. 357-60. In a recent discussion, Coleman averred that his claim that corrective justice is layered in part upon local conventional norms already existing in a community "is the most striking and controversial feature of the argument in Part III." Jules L. Coleman, Risks and Wrongs. 15 HARV. J.L. & PUB. POLY. 637, 646 (1992).
themes and subthemes in a book as rich as this one. I hope, however, that I have been able to indicate in these concluding remarks that there is more to say about the connection between corrective justice and liberal values than Coleman's predilection for purely middle-level theory might suggest.
LIFE'S SACRED VALUE — COMMON GROUND OR BATTLEGROUND?

Alexander Morgan Capron*


Reading Ronald Dworkin's1 ambitious and fascinating attempt to find a middle way out of the increasingly heated public battles over the legal regulation of abortion and euthanasia,2 I found myself thinking — for reasons that I trust will presently become apparent — of an automobile accident a little more than a decade ago. That accident threw a young woman named Nancy Beth Cruzan into a ditch where rescue workers found her unconscious. After restoring her vital functions, they took her to a nearby hospital, where doctors further stabilized her condition and eventually transferred her to a state rehabilitation facility.3

Over the next several years, as she failed to regain her mental fac-

* Henry W. Bruce University Professor of Law and Medicine and Co-Director, Pacific Center for Health Policy and Ethics, University of Southern California. B.A. 1966, Swarthmore; LL.B. 1969, Yale; M.A. 1975, University of Pennsylvania. — Ed.

1. Professor of Law, New York University; University Professor of Jurisprudence, Oxford University.

2. As some participants in these debates will take exception to Dworkin's definitions of these terms, it is important to keep in mind how he uses them: "Abortion, which means deliberately killing a developing human embryo, and euthanasia, which means deliberately killing a person out of kindness, are both choices for death. The first chooses death before life in earnest has begun, the second after it has ended." P. 3. Some proponents of the right to abortion emphasize that abortion "means" the termination of pregnancy and the avoidance of parenthood; if one could accomplish those goals without killing an embryo, Dworkin's definition would not apply. His definition, however, seems to convey adequately the common usage of the term.

The same cannot be said for his definition of euthanasia, especially because his own use of the term makes clear that by "deliberately killing a person out of kindness" he intends to encompass not only active euthanasia of the sort Janet Adkins achieved with Dr. Jack Kevorkian's assistance but also the withdrawal of respirators or tubes providing hydration and nutrition and apparently even the acceleration of death brought about by narcotics used to relieve severe pain. Although one could argue that all of these practices amount to "deliberate killing" in the sense of being done intentionally with a known likelihood of bringing about death, lumping all together under the heading "deliberate killing" is likely to confuse the issue — and possibly, to distort public policy formulation — whatever one may say for its logical consistency. Furthermore, to say that euthanasia involves causing death "after [life] has ended" either takes a very strong position on the "quality of life" issue — equating a range of impairments with nonlife — or draws a line that would exclude from the purview of euthanasia several conditions — such as AIDS and cancer — that many cite as strong cases for voluntary, active euthanasia or "physician assistance in dying."

ulties and slid into what physicians call a persistent vegetative state, her parents slowly and reluctantly concluded that she would not have wished to be sustained in her condition by feeding tubes or other medical intervention. When hospital officials rejected their request to withdraw treatment, her parents sought — and a court awarded — formal guardianship with authority to order termination of all life support. 4

The Missouri Supreme Court reversed the trial judge’s order on the ground that the state’s “interest in preservation of life” overrode any interest that Ms. Cruzan might have in being free of such death-delaying interventions. The state’s interest, held the court, “embraces two separate concerns: an interest in the prolongation of the life of the individual patient and an interest in the sanctity of life itself.” 5

The U.S. Supreme Court affirmed the Missouri ruling, holding that a state does not violate the Due Process Clause when it refuses to allow the forgoing of a patient’s life support, absent clear and convincing evidence that the patient had expressed such a wish while still competent. 6 The Court also reasoned that the state’s interest in protecting human life was sufficiently strong to outweigh a guardian’s conclusion that further treatment is not in the patient’s best interest. 7 The notion that the state has an interest in the “preservation of human life” independent of the patient’s own interests drew stinging dissents from Justices Brennan and Stevens:

[T]he State has no legitimate general interest in someone’s life, completely abstracted from the interest of the person living that life, that could outweigh the person’s choice to avoid medical treatment. 8

4. 760 S.W.2d at 411-12.
5. 760 S.W.2d at 419.
7. Missouri relies on its interest in the protection and preservation of human life, and there can be no gainsaying this interest. As a general matter, the States — indeed, all civilized nations — demonstrate their commitment to life by treating homicide as a serious crime....

... Finally, we think a State may properly decline to make judgments about the “quality” of life that a particular individual may enjoy, and simply assert an unqualified interest in the preservation of human life to be weighed against the constitutionally protected interests of the individual.

497 U.S. at 280, 282.

8. 497 U.S. at 313 (Brennan, J., joined by Marshall and Blackmun, JJ., dissenting). Disagreeing with the evidentiary standard as adopted and applied in the case, Justice Brennan addressed the claim that the state was justified in its allocation of the “risk of error” in a way that favored “the status quo”:

An erroneous decision [that is, one that did not reflect the patient’s true wishes] to terminate artificial nutrition and hydration, to be sure, will lead to failure of that last remnant of physiological life, the brain stem, and result in complete brain death. An erroneous decision not to terminate life support, however, robs a patient of the very qualities protected by the right to avoid unwanted medical treatment. His own degraded existence is perpetuated; his family’s suffering is protracted; the memory he leaves behind becomes more and more distorted.

Even a later decision to grant him his wish cannot undo the intervening harm.

497 U.S. at 320.
However commendable may be the State's interest in human life, it cannot pursue that interest by appropriating Nancy Cruzan's life as a symbol for its own purposes. Lives do not exist in abstraction from persons, and to pretend otherwise is not to honor but to desecrate the State's responsibility for protecting life.9

Given the degree to which the dispute in *Cruzan* involved the sacred value of human life, one is not immediately encouraged to find that this concept is central to *Life's Dominion*, in which Dworkin proposes to offer an escape from the overheated rhetoric of the abortion and euthanasia debates. At first, Dworkin's alternative analysis — replacing a debate over rights and interests with a discussion of the intrinsic value of life — seems merely to shift the terms of engagement without ending the war. Indeed, it is hard to believe that Dworkin seriously thought that his alternative would produce a new era of calm analysis and rational disagreement. Yet, however much *Life's Dominion* makes him seem like the Don Quixote of public policy, he has once again written an interesting book that is well worth reading for its arguments and insights about the importance of returning to the moral question — “What is life's intrinsic value?” — if not for the armistice it attempts to impose on the right-to-life/quality-of-life wars.10

I. **ABORTION AND THE SACRED VALUE OF LIFE**

Most of *Life's Dominion* is concerned with abortion, first in philosophical and then in constitutional law terms. Dworkin begins by demonstrating that even those who speak in right-to-life language — such as President Bush and Vice-President Quayle during the 1992 campaign — do not actually believe that a fertilized egg from the moment of conception is a person whose abortion would always be murder. As the polls show, most people would allow abortion in the case of rape or incest or severe fetal abnormalities;11 this demonstrates that however much they value a fetus, they do not regard it as a child or other person. As Dworkin points out, even the Catholic Church has not always held its present position (p. 39). Moreover, the Catholic position is at odds with the writings of theologians such as Thomas

---

9. 497 U.S. at 356-57 (Stevens, J., dissenting).

10. I note, but leave for the reader to evaluate, Dworkin's claim that his position will be more likely to improve "the quality of public political argument" than the theories put forward by others because he connects theory and practice "from the inside out" — that is, by beginning with the practical problems of abortion and euthanasia and then asking "which general philosophical or theoretical issues we must confront in order to resolve those practical problems." Pp. 28-29. If Dworkin intends more than practical reasoning and induction, and if his metaphor that in reasoning from the inside out "theories are bespoke, made for the occasion, Savile Row not Seventh Avenue" (p. 29) describes his method more than simply his taste in clothes, then one may well doubt that others should give much credence to theories that are so individually tailored. Does that mean that these theories need not be consistently applied to other settings or meet other tests of integrity? See infra note 20.

Aquinas, who grounded his theory of ensoulment on the Aristotelian notion of *hylomorphism*, or the idea that the human soul only exists in an identifiably human body, which by our present understanding would be one with a functioning neocortex — the organ required for spiritual activity, among other forms of conscious reflection, not functional until late in gestation (pp. 40-42).

Thus, Dworkin argues that the opponents of abortion have misdescribed the moral claim on which their opposition rests. He labels the conventional view the *derivative* objection to abortion “because it presupposes and is derived from rights and interests that it assumes all human beings, including fetuses, have” (p. 11). He believes, however, that the antiabortion position is better understood as being *detached* from any presupposed rights or interests and as resting instead on the claim

\[
\text{that human life has an intrinsic, innate value; that human life is sacred just in itself; and that the sacred nature of a human life begins when its biological life begins, even before the creature whose life it is has movement or sensation or interests or rights of its own. [p. 11]}
\]

Unlike the derivative view, framed in terms of rights and interests, the detached view leads not to conclusions about the claims that people can make on others for action or forbearance but instead to more impersonal conclusions about whether a particular action would fail to respect — or would “disregard and insult,” as Dworkin puts it (p. 11) — the intrinsic value of an entity.12

If conservatives have it wrong when they claim that abortion is immoral because it violates the fetus’s “right to life,” liberals also have it wrong in suggesting that all that is at issue is a woman’s right to control her body. Again, Dworkin suggests that once we set aside political rhetoric, we will see that feminists do not view abortion simply in terms of women’s rights (pp. 50-60). To illustrate his thesis, he examines the positions of leading feminists and finds ways in which their views are — or can be made to be — consistent with his thesis that the fetus is something of intrinsic importance, albeit not a rights-bearer.

For example, Catharine MacKinnon rejects *Roe v. Wade*’s right-to-privacy theory partly because it links pregnancy to other situations in which the more powerful one of two connected entities asserts a sovereign right to end the connection.13 Instead, MacKinnon articu-

---

12. Living human beings are not the only entities of value: A beautiful sculpture can be smashed, and that would be a terrible insult to the intrinsic value that great works of art embody and also very much against the interests of people who take pleasure in seeing or studying them. But a sculpture has no interests of its own; a savage act of vandalism is not unfair to it. Pp. 15-16.

lates the pregnant woman’s perspective that the fetus “is both me and not me. It ‘is’ the pregnant woman in the sense that it is in her and of her and is hers more than anyone’s. It ‘is not’ her in the sense that she is not all that is there.”

MacKinnon further contends that until men’s sexual domination of women is replaced with true equality, fetuses must have lower status than women. MacKinnon frames that view, Dworkin argues, in terms of the fetus-as-person debate, but recast in light of his detached view of the fetus, it leads to the “arresting” conclusion “that the intrinsic importance of a new human life may well depend on the meaning and freedom of the act that created it” (p. 56).

Similarly, in the theories of another legal scholar, Robin West, Dworkin finds further affirmation for his contention that feminists support abortion, not as a manifestation of “a right to be left alone but often to strengthen their ties to others.” Abortion thus involves a conflict among responsibilities, including acting responsibly toward the future child — not the existing fetus. A person thinking in these terms is not thinking of competing rights but of relationships among entities, each of whom has intrinsic value.

All of this seems sound and even helpful, at least analytically, albeit less novel than Dworkin makes it out to be. As Daniel Callahan pointed out several years ago, the continuing battle over Roe has meant that both sides have felt the need to harden their positions. As a result, they have failed to confront the moral issues of abortion in the way that we ought to address such important aspects of our lives. The proponents of women’s choice feel particularly vulnerable; any admission that it is not always right to do what they say women have a right to do may appear to be an open invitation to opponents to restrict or revoke the right itself.

Yet, while Dworkin seems correct that neither side in the prolife-prochoice conflict really holds a pure “rights” perspective, it seems naive to the point of disingenuousness to suggest that both positions

14. Id. at 1316 (footnote omitted).
15. Id. at 1317.
17. Life’s Dominion’s 17 pages of footnotes (pp. 243-59) reveal that Dworkin has relied on a good deal of discussion in the academic legal literature as well as the popular media but show less familiarity with the discussion of the issues by lawyers, philosophers, and biomedical scientists in the bioethics literature, aside from chapter 8 (pp. 218-43), which apparently grew out of a paper that he wrote for a project on Alzheimer’s disease for the Congressional Office of Technology Assessment. P. 222 n.12; see Ronald Dworkin, U.S. Congress, Office of Technology Assessment, Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia (1987), microformed on Philosophical, Legal, and Social Aspects of Surrogate Decisionmaking for Elderly Individuals, CIS No. OTA952-30 (Cong. Info. Serv.).
rest on the notion of the sacredness of life. In political — as opposed to moral or existential — terms, the net effect of Dworkin's clever argument is just to shift the terms of debate. No longer is the issue "Is the fetus a person, possessed of a right to life?" Instead, the question becomes "Does the sanctity of all human life stand in the way of abortion?"

One might think that proponents of choice would be most worried by this way of framing the issue. As the dissenting Justices in the Cruzan case make clear in their anger with the majority's reasoning, invocations of the sacredness of human life readily lead to limitations on individual choice in the name of the state's obligation to protect and preserve human life. Yet they need not fear, for in Dworkin's hands the result is to limit severely how far the state can go in enforcing any particular view of the sacredness of life.

II. Abortion and the Law

To translate respect for the intrinsic value of human life into law, Dworkin would permit states to take action geared toward maintaining "a moral environment in which decisions about life and death are taken seriously and treated as matters of moral gravity" (p. 168). Exactly how each individual actually weighs the sacredness of life in making decisions about such matters as abortion and euthanasia would, however, be beyond the reach of state regulation.

Dworkin believes that any imposition of an officially promulgated position on issues about the meaning of life is not only tyrannical but destructive of individual moral responsibility. Yet the community has in the past taken upon itself the enforcement of a single set of views. Although it is doubtful that most Americans would be comfortable allowing the government to treat as heresy views with which the majority disagrees, it is less obvious that a majority would reject some limitations on conduct that directly threatens human life. For example, as the Chief Justice noted in Cruzan, civilized societies everywhere

19. See supra notes 8-9 and accompanying text.
20. In reaching this conclusion, Dworkin first argues that Justice Blackmun basically got things right in Roe v. Wade insofar as the question at issue involves the "derivative" view that the fetus is endowed with the rights of a person. Pp. 104-11. Next, he takes on the conservative constitutional scholars who challenge the right to privacy as a judicial invention. In the process of arguing for a "constitution of principle" rather than the "constitution of detail" that he attributes to Justice Scalia and other conservative jurists, Dworkin subjects "enumerated rights" and "original intent" theories to withering critiques. Pp. 129-44. He ends — as readers of his 1986 book, Law's Empire, might anticipate — with the plea that the public dismiss further arguments framed in these terms about the suitability of judicial nominees or the soundness of judicial opinions and instead "seek genuine constraints [on judicial power] in the only place where they can be found: in good argument." P. 145. The fact that there is no universally accepted metric for identifying bad judicial decisions ought not to lead us to the critical conclusion that legal reasoning is a waste of time, because we can at least insist that judges "accept an independent and superior constraint of integrity in the decisions they make." P. 146.
treat murder as a serious offense. As Justice Scalia added in his concurrence, both the common law and the statutory law in this country have treated suicide as a crime — although it no longer is — and even private citizens' use of force to prevent suicide is privileged.

Thus, Dworkin faces at least two tasks: (i) to show that in furthering respect for the sanctity of life, the Constitution forbids states from outlawing the ending of the life of a human being — such as a fetus — who does not possess the full rights and interests of a person; and (ii) to convince those who now take a prolife view that, having accepted his reanalysis of their underlying premise — from right-to-life to sacredness-of-life — they must limit themselves to preaching and in the end must respect the choices made by people who understand very differently the limitations on human choices that are inherent in life's sacredness. He is more successful at the first than the second task, despite his optimism that reframing our political and constitutional debates in terms of sanctity of life not only better explains what is at stake but also allows for a greater degree of agreement.

To address the first task, Dworkin directs our attention not to the right to privacy but to the First Amendment's guarantees of religious liberty, which he sees as necessarily protecting "procreative autonomy" (p. 166). The Constitution does not protect all claims to free choice in the face of state regulations and restrictions, even when they involve things of great importance to people. However, some matters — such as reproduction — are so foundational and our conceptions of them are so vulnerable to the intrusive effects of state regulation that Dworkin argues they must be protected against the imposition by the state of a single view of life's intrinsic value. To do otherwise, he claims, would restrict what are essentially religious beliefs — recognizing that not all people would so identify this category of beliefs in their own systems of morality.

One problem for Dworkin's conception is to explain how we can distinguish the "existential question — does human life have any intrinsic or objective importance?" — from "more secular convictions about morality, fairness, and justice" (p. 156). One way to derive the answer is to reason backwards: if an essential responsibility of government is to decide how to reconcile our competing rights, then this

21. See supra note 7.
23. As he notes later, his views on religious liberty are very close in substance to those expressed by the three "center" Justices in Casey. Pp. 171-76. The latter rested their opinion on the Due Process Clause instead, but like Dworkin they concluded that "[a]lthough the heart of liberty is the right to define one's own concept of existence, of meaning, of the universe, and of the mystery of human life. Beliefs about these matters could not define the attributes of personhood were they formed under the compulsion of the State." Planned Parenthood of Southeastern Pa. v. Casey, 112 S. Ct. 2791, 2807 (1992).
must be a secular function, and the guarantees of religious liberty must not limit the state in exercising it. Conversely, there is no need for the state to take positions on the meaning of life; these matters are left to individual convictions under the First Amendment.

If Dworkin is correct, however, in concluding that questions that turn on the value of life are essentially religious because they rest on underlying beliefs about the normative value of the beings that embody life — as endowed by the Creator, or as a result of a natural order, or whatever — and if he is correct that decisions closely tied to religious matters of this sort are protected against state dictates, he is then left with several problems. First, the question of regulation and prohibition: May the state regulate abortion — for example, effectively prohibiting it late in pregnancy? Second, the question of education and advocacy: Should the state be permitted to attempt to influence women to choose not to abort?

To the former question Dworkin replies that a state can justify limitations on abortion after viability because the insult to the sanctity of life increases as the fetus develops and because most women know about their pregnancies from an early point, so that a woman waiting until late in term to abort demonstrates an indifference to the cultural sense of the greater wrong in destroying a more developed fetus — an indifference against which society has a right to protect itself. Furthermore, around the time of viability the brain of the fetus is sufficiently developed that the fetus, though not a person in the constitutional sense, “might then sensibly be said to have interests of its own” (p. 169). The state may protect those interests so long as it does so in a way that respects the rights of constitutional persons (p. 169).

As to the second question, Dworkin is comfortable with the line set out in *Casey*, that is, that the state may be an advocate for fetal life provided that in doing so it does not create an “undue burden” by posing “substantial obstacles” to the choice.24 Of course, reasonable people will disagree about when the state reaches the point of “undue burden,” but interventions aimed at promoting thoughtful deliberation — as opposed to those that merely inconvenience or obstruct a woman’s choice — may well be consistent with respect for women as individual judges of life’s value, each by her own lights. Surely, however, even if the state were to limit itself to urging women to appreciate the gravity of their decisions, some people would be as offended by this official preachment as they would be by the “establishment” of any other view that they would regard as essentially religious — as, in effect, Dworkin has admitted all views about the sanctity or intrinsic value of life to be. By this point, Dworkin is skating along at an

---

24. 112 S. Ct. at 2820.
Olympic clip, but the ice supporting him seems to have become awfully thin.

III. EUTHANASIA

In the nearly twenty years since the New Jersey Supreme Court handed down its landmark decision in In re Quinlan, courts have widely adopted the view that competent patients have the right to refuse any medical care, including life-sustaining interventions such as respirators and feeding tubes. In Cruzan a majority of the Justices recognized that the Constitution guaranteed such a liberty interest to each person. Further, under the common law and under statutes adopted in most states, a person while competent may prepare an "advance directive" that either gives instructions about the extent of life support she wants under specified circumstances — such as "a terminal condition" — or names an agent who is authorized to make treatment decisions if the patient becomes unable to do so, or both. Finally, most but not all jurisdictions allow a court-appointed guardian or an incompetent patient's family or close friends to make treatment decisions in the absence of an advance directive or comparable oral instructions. Such a decision by a "surrogate" may be either a "substitute judgment," based fairly closely on the known values and preferences of the patient, or, in the absence of sufficient information, a determination of what the surrogate believes would be in the "best interests" of the patient, given his or her prospects for survival, improvement, and recovery and the burden that the process would impose — in terms of dignity as well as financial cost and physical pain — on the patient.

To these two hallmarks of the end-of-life cases that are personal to the patient — autonomy and best interests — Dworkin adds a third, impersonal factor: the sacred value of life (p. 194). As noted at the outset, for some people the clear implication of this value is that it is permissible for the state to insist on prolonging a patient's physical


27. 497 U.S. at 278-79.


30. See supra notes 5-7 and accompanying text.
existence even when it might not be in that patient’s best interests. The Court in *Cruzan* upheld a state’s authority to do so, in the absence of clear and convincing evidence that the patient would have wanted treatment to cease under the circumstances. As we have seen in Dworkin’s use of the idea of the sacred value of life in the abortion context, however, *Life’s Dominion* does not support such a simplistic use of the notion that life has intrinsic worth. Indeed, in his final two chapters Dworkin’s major objective seems to be linking the concept of the sanctity of life that he developed in the abortion context to the existing decisional tools — autonomy and best interests.

The analysis in these chapters on critically ill and demented patients is serviceable, albeit much less novel than the earlier material. Many thoughtful writers have in recent years addressed the problems of decisionmaking by and for dying, demented, and comatose patients — when, for example, the present “best interests” of a patient ought to override his or her “autonomous” wishes, whether currently or previously expressed, or to what extent physicians have an obligation to provide or withhold care when doing so conflicts with their professional or personal values. Dworkin would have greatly enriched the discussion had he grappled with contemporary thinking about the place of health and pain in human life and the roles of physicians and others in health care. His philosophical musings on the role played by surrogates — especially when they claim to be able to make a substituted judgment based on what they think an incompetent patient would choose — take no account of studies that suggest that surrogates — and physicians — do not do a very accurate job of making such predictions.

*Life’s Dominion* delves less deeply into euthanasia than abortion; Dworkin is content to set forth a philosophical analysis in this area without exploring its legislative or constitutional details. “None of us wants to end our lives out of character,” he concludes (p. 213). Thus, in assessing a person’s autonomous choices or in calculating his best

---


32. In his final chapter, “Life Past Reason,” in which Dworkin deals with decisionmaking for and about demented patients with special reference to Alzheimer’s Disease, he shifts to using the terms *autonomy*, *beneficence*, and *dignity*. Pp. 218-41. The discussion is a useful elaboration on the earlier examination of these issues and their relationship to the sanctity of life, but the chapter is the least well integrated into the central thesis of the rest of the book.


interests, society should recognize that the sanctity of human life
means more than its mere biological continuation — what Dworkin
terms "the natural investment" (p. 214). Sanctity also means the human investment, what an individual regards as his critical interests, or those things that make up a good life — and hence also a good death. This is a more eloquent and fully elaborated version of the views expressed by Justices Brennan and Stevens in their dissents in Cruzan.35

Because Dworkin eschews wrestling with the "difficult and important administrative questions" (p. 216), he does not get beyond showing that the Chief Justice's opinion failed to distinguish clearly the distinct values of autonomy, best interests, and sanctity of life and, by implication, badly misunderstood the last of these. It is thus opaque whether Dworkin intends his bold conclusion — that respect for the sanctity of life counsels in favor of active euthanasia and assisted suicide in certain cases — to lend support to efforts to legalize physician participation in these practices or merely to be a philosophical argument. This seems to me regrettable, as some of the most telling arguments against legalization come from commentators such as Yale Kamisar who weigh concerns other than those that Dworkin termed religious in his analysis of abortion.36 My own sense is that the risks to the very values Dworkin identifies with the human investment in life are such that society may forbid certain forms of killing, especially by medical personnel, even when carried out at the request of, or to further the best interests of, seriously ill patients, and that doing so does not amount to the "devastating, odious form of tyranny" (p. 217) that Dworkin correctly implies would offend any civilized society, especially one with a formal bill of rights such as ours.

IV. THE VALUE OF LIFE

Dworkin's treatment of the issues at the end of life is briefer, I suppose, because it is less difficult to convince people of his central claim, that is, that the value or worth of human life comes not from mere biological functioning but from how the person living it conceives of his critical interests and lives his life to manifest and achieve those interests. Although Dworkin argues for what he terms the detached view of the intrinsic value of life, in the case of a patient who is or once was a competent adult — or perhaps even a child of an age of reason — this detached view usually coincides with the derivative view, namely that value flows from respect for the individual's rights and interests.

When discussing abortion, in contrast, Life's Dominion faces a

35. See supra notes 8-9 and accompanying text.
harder challenge because the intrinsic value of life is more dependent on the natural investment in that life, the human dimension being still absent. It is thus a mark of the book’s success on an analytic level that it builds such a strong and relatively novel claim for the effect of the sanctity of life on abortion law and practice. Yet, at the same time, the very effectiveness of Dworkin’s argument probably makes futile his professed ambition to bring prolife advocates to the table with their prochoice counterparts for a harmonious discussion carried out with common terminology. Nonetheless, any reminder of the human meaning of our precious lives and any encouragement to reflect on how one should go about living well that is as lucid as Life’s Dominion deserves our applause.
A MORE DEMOCRATIC LIBERALISM

Joshua Cohen*


What do we do when we find the truth? . . . When men learned the Earth was round, did they allow their geographers to continue to teach that it was flat?

. . . .

. . . If you would see the monuments of a society that has come to consider the truths that Jesus Christ taught us as one among an indefinite variety of moral codes by which to live, look around you.

Amen, and Happy Easter.¹

I. SOCIAL UNITY AND MORAL PLURALISM

When Peter Laslett published his first collection of essays on Philosophy, Politics and Society in 1956, he reported that “[f]or the moment, anyway, political philosophy is dead.”² As the book reviews in this annual Survey indicate, things have changed. Political philosophy is back, and its revival owes much to John Rawls’s A Theory of Justice (Theory).³ Published more than twenty years ago, Theory remains the starting point for contemporary work on justice. This fact by itself is sufficient to make the appearance of Rawls’s second book, Political Liberalism (Liberalism),⁴ an important event.

But the intellectual importance of Liberalism reaches well beyond

¹ Pat Buchanan, Tolerance and Truth at Easter, ARIZ. REPUBLIC, Apr. 3, 1994, at E5.
² Peter Laslett, Introduction to PHILOSOPHY, POLITICS AND SOCIETY at vii, vii (Peter Laslett ed. 1956).
⁴ Political Liberalism is presented as a series of eight lectures, which descend in complex ways from earlier lectures and papers. The first five lectures are revisions of previously published articles, but the revisions are substantial even when — as with lectures 4 and 5 — the titles have not been changed. Lecture 6 is a significantly modified version of material presented in public lectures but never before published. Lectures 7 and 8 were published previously and are reproduced without modification. Pp. xii-xiv. Apart from lectures 7 and 8, then, it is a mistake to identify the views advanced in Political Liberalism with positions taken in earlier versions of the lectures.

1503
the biography of its author and the recent history of political philosophy. Rawls's book is a deep and original examination of a fundamental problem of modern politics. Modern societies are marked by manifest ethical, religious, and philosophical disagreements among citizens. Moreover, the disagreements are of a special kind. Although citizens commonly regard the moral, religious, and philosophical views of others as false, they need not regard others as unreasonable for endorsing those views. Because human reason appears not to converge on a single moral outlook, we seem to face "a plurality of reasonable yet incompatible comprehensive doctrines" (p. xvi). What are the implications of these doctrinal conflicts — this "fact of reasonable pluralism" (p. xvii) — for our understanding of the requirements of justice and the possibility of a just society?

A. Democratic Toleration

*Liberalism* addresses this question against the background of the account of justice as fairness advanced in *Theory*. In *Theory*, Rawls proposed an ideal of a well-ordered, democratic society featuring consensus on a conception of justice rooted in the value of fair cooperation among citizens as free and equal persons. But *Theory*, Rawls now thinks, did not take the fact of reasonable pluralism seriously enough. The presentation suggested that justice as fairness depends on a comprehensive liberal philosophy of life — that only people who endorse a view of our nature and of the human good that emphasizes independence, choice, and self-mastery have good reason to endorse justice as fairness.

*Liberalism* asks, then, whether justice as fairness can be freed from this dependence. Can views that disagree about moral fundamentals — some of which reject a comprehensive liberal philosophy of life — nevertheless agree on a political conception of justice rooted in "values of equal political and civil liberty; fair equality of opportunity; ... economic reciprocity; [and] the social bases of mutual respect between

---

5. I say "manifest" because I do not suppose that any society is morally or religiously homogeneous, however much its institutions may suppress the expression of differences by limiting expressive liberty, establishing compulsory forms of worship, or narrowly circumscribing associative liberty.

6. I will say more about the distinction between reasonable and unreasonable later. See infra section IV.C. It will suffice here to note the familiar logical distinction between is true and is reasonable: inconsistent views cannot both be true, but they can both be reasonable.

7. Versions of this question are posed at pp. xviii, xxv, 4, and 133. Rawls does not suppose that the fact of reasonable pluralism taken on its own leads us to a particular conception of justice. The problem of *Liberalism* is generated instead by an apparent tension between the fact of reasonable pluralism and the ideal of a well-ordered society featuring consensus on a conception of justice that articulates such fundamental political values as fairness, equality, and liberty.


9. On the idea of a comprehensive moral conception, see p. 13. For the concern that *Theory* endorses such a conception, see pp. xvi-xvii; see also infra section II.B.
citizens” (p. 139)? Or does the fact of reasonable pluralism imply that we ought to give up on the idea of a consensus of justice, that democratic politics can never be more than a combination of individual calculation, group bargaining, and assertions of discrete collective identities — when democracy works well — and deceit, manipulation, and naked force — when democracy works badly?

In a world full of cruelty, depravity, and grief, we ought not to dismiss the virtues of a politics of group bargaining within a framework of rules that win general compliance — “a mere modus vivendi” (p. 145). Still, Liberalism defends the possibility of doing better: of achieving a consensus on political justice under conditions of fundamental moral, religious, and philosophical disagreement.

The key to that possibility is that political values — for example, the value of fair cooperation among citizens on a footing of mutual respect — are extremely important values and can be acknowledged as such by conflicting moral conceptions, by views that disagree with one another about ultimate values and about the best way to live.10 To be sure, those views will explain the importance of political values in very different terms:11 for example, as rooted in autonomy,12 or self-realization,13 or human happiness properly understood,14 or the appropriate response to life’s challenges,15 or the value of individuality,16 or the equality of human beings as God’s creatures.17 These competing explanations of the political values will in turn manifest themselves in conflicting views about individual conduct and personal virtue.

Still, an affirmation of the importance of political values is not the unique property of a particular moral outlook. For this reason, the different moral views that flourish in a society governed by a conception of justice rooted in the ideal of fair cooperation on a footing of

11. Some views may treat fairness itself as a fundamental value and not as an implication of some deeper moral value. See the “third view” at p. 145.
15. See Ronald Dworkin, FOUNDATIONS OF LIBERAL EQUALITY, in 11 THE TANNER LECTURES ON HUMAN VALUES 1 (Grethe B. Peterson ed., 1990) [hereinafter TANNER LECTURES].
16. See STUART HAMPSHIRE, INNOCENCE AND EXPERIENCE 114, 117-18, 124-36 (1989). Hampshire also explains the value of fair political process in terms of its role in preventing such great evils as “murder and the destruction of life, imprisonment, enslavement, starvation, poverty, physical pain and torture, homelessness, friendlessness.” Id. at 90.
mutual respect may each have good and sufficient reason to support
that conception as the correct account of justice and not simply as a
suitable accommodation to conditions of disagreement. Citizens who
endorse different moral axioms may still arrive at the same theorems
about political justice, and some people may simply endorse a view of
justice without resting that endorsement on a more comprehensive
moral theory.\textsuperscript{18}

In such a society, we have an “overlapping consensus” on a “politi-
cal conception of justice.”\textsuperscript{19} Citizens achieve social unity because
they all accept that conception and so agree to conduct the fundamen-
tals of political argument on the shared ground that the conception
makes available and to set aside for political purposes their deep, ulti-
mate, and persistent disagreements about what we are like, what the
world is like, and how best to face its demands.

This account of the combination of unity and pluralism rests on a
new interpretation of the ideal of toleration — call it “democratic tol-
eration”\textsuperscript{20} — paralleling the new interpretation of the social contract
advanced in \textit{Theory}. In \textit{Theory}, Rawls proposed “to generalize and
carry to a higher order of abstraction the traditional theory of the so-
cial contract.”\textsuperscript{21} The combination of social unity and moral pluralism
captured in \textit{Liberalism}’s idea of overlapping consensus generalizes and
carries to a higher order of abstraction the conventional idea of
toleration.

Conventionally understood, toleration is a substantive political
principle condemning the imposition of an authoritative form of reli-
gious worship or, in a more expansive version, an authoritative form of
personal morality.\textsuperscript{22} Aiming to provide a conception of toleration bet-
ter suited to “the historical and social circumstances of a democratic
society” (p. 154), Rawls’s political liberalism deepens the idea of toler-
ation and “applies the principle of toleration to philosophy itself” (pp.
10, 154). That is, in addition to accepting the substantive requirement
of toleration, \textit{Liberalism} presents toleration as a condition on political
justification, at least when the question concerns “constitutional essen-
tials” and “basic questions of justice.”\textsuperscript{23} Given the plurality of incom-
patible yet reasonable views held by equal citizens in a democratic
society, the ideal of fair cooperation recommends that we free the vo-
cabulary and premises of political justification from dependence on

\begin{itemize}
\item \textsuperscript{18} This possibility plays an important role in \textit{Liberalism}. \textit{See} pp. 155-56.
\item \textsuperscript{19} On overlapping consensus, see pp. 132-72; on the idea of a political conception of justice,
\textit{see} pp. 11-15, 174-75.
\item \textsuperscript{20} Rawls rejects perfectionism in the name of “democracy in judging each other’s aims.”
\textit{Rawls, supra} note 3, at 442.
\item \textsuperscript{21} \textit{Id.} at viii; \textit{see also} p. xv.
\item \textsuperscript{22} On the central role of religious toleration in understanding the value of toleration, see
\item \textsuperscript{23} \textit{See} pp. 137, 227-30.
\end{itemize}
any one view. Put otherwise, Rawls suggests that when we understand political power as "the power of free and equal citizens as a collective body" (p. 136) and take account of the fact of reasonable pluralism, we will want to be sure that political argument on fundamentals proceeds on grounds that are acceptable to citizens generally, not in the terms provided by a particular philosophical or religious tradition (pp. 136-68, 216-18).

To be sure, it may be impossible to gain support for a conception of justice from all views. But perhaps support for the conception, and a willingness to conduct public political argument in its terms, will come from the "reasonable comprehensive doctrines" (p. 59) held by reasonable citizens: the views held by people who are concerned to cooperate on terms that others accept and who recognize that reason itself does not select a single comprehensive view.24

The central line of thought in *Liberalism*, then, is that we can achieve the good of consensus on justice without comprehensive moral agreement;25 the absence of comprehensive agreement does not reduce politics to calculations of individual advantage, interest-group bargaining, or the self-affirmation of discrete collective identities. Instead, because political values are highly important values and are recognized as such within a wide range of moral conceptions, consensus on a conception of justice is possible under conditions of reasonable pluralism and must accommodate those conditions if it is to suit the equal citizens of a democratic society.

**B. Reconciliation Without Metaphysics**

Rawls's project in *Liberalism* bears certain important similarities to Hegel's in his *Philosophy of Right*,26 and it will be instructive to sketch both the commonalities and the differences between their projects.

In his political theory, Hegel aimed to reformulate a classical ideal of political society, which supposed that citizens share an understanding of justice and the human good,27 in light of the post-Reformation idea of unbridgeable differences among citizens on fundamentals. How is it possible, Hegel asked, to achieve the good of shared commitments in the face of apparently ultimate differences in interest and outlook that are so much the focus of the energies of modern civil society? How, in Hegel's terms, can we give stable expression to both the uni-

---

24. On reasonable comprehensive doctrines, see pp. 58-66.
versal and particular aspects of our nature?28

Hegel and Rawls share broadly similar questions and both endorse the hopeful possibility of reconciling apparently competing demands of unity and difference. Their proposals about how to achieve that reconciliation differ profoundly, however, both in substance and in the insight about the reconciliation they expect philosophy to provide.29

Hegel located his answer within a generally antidualistic, logico-metaphysical theory. His philosophical system revealed our nature as free beings,30 showed how our differences are less fundamental than we are prephilosophically inclined to think, linked the expression of our free nature to the institutions of a state whose aim is the realization of the good — understood as the expression of our nature31 — and showed how that expression and those institutions were the natural upshot of historical evolution.32

According to Rawls, evaluative theories are matters of reasonable disagreement, and for that reason we ought not to build a conception of political justice around the view of the good advanced within any one such theory. Moreover, the reconciliation of social unity and moral pluralism cannot proceed on the terrain of metaphysics. Because there are ultimate, reasonable disagreements about metaphysical doctrines, a general philosophical argument against dualisms, for example, cannot provide part of the case for overcoming the specific tension between pluralism and social unity.33 Political philosophy, if it seeks to operate on the shared ground available to equal citizens in a pluralistic public, cannot rest on a metaphysical theory of our true nature, nor can it provide any assurances, grounded in such a theory, about the ultimate expression of that nature in history.34 Its aims must be less ambitious, focused on clarifying how social unity is possible under pluralistic conditions. Such clarification will not yield the

28. On the role of the modern state in achieving this stable expression, see HEGEL, supra note 26, § 260.
30. HEGEL, supra note 26, § 4.
31. "[The good is] realized freedom, the absolute and ultimate end of the world." Id. § 129.
33. In an earlier version of some of the material published in Liberalism, Rawls indicated that "one of Hegel's aims was to overcome the many dualisms which he thought disfigured Kant's transcendental idealism," that Dewey "shared this emphasis throughout his work," and that "there are a number of affinities between justice as fairness and Dewey's moral theory which are explained by the common aim of overcoming the dualisms in Kant's doctrine." John Rawls, Kantian Constructivism in Moral Theory, 77 J. PHIL. 515, 516 (1980). My point is not to deny this common aim. I want only to emphasize that the presentation of justice as fairness as a political conception implies that its resolution of the apparent tension between social unity and moral pluralism cannot draw on a general antidualistic metaphysical view.
34. Later I will discuss some reasons for operating on shared grounds. See infra notes 114-16 and accompanying text.
assurances of unity associated with a historical theodicy;\(^\text{35}\) at best it will lead to an understanding of why the hope for reconciliation is not unreasonable.

Once we understand how the stable combination of shared principles and conflicting faiths that defines an overlapping consensus is possible, then we can see — Rawls thinks — that it is reasonable to adhere to the ideal; the conditions of its possibility are not so demanding as to condemn it. In this way Liberalism offers a "defense of reasonable faith in the possibility of a just constitutional regime."\(^\text{36}\) It argues for the reasonableness of that faith by revealing the commitments it requires as minimally demanding, emphasizing in particular that people within different moral and religious traditions can reasonably endorse those commitments. If Liberalism is right, then it is possible to combine fundamental moral pluralism — to take seriously one sort of difference — with consensus on a conception of justice suited to the equal citizens of a democratic society. But while philosophy can provide that service in a democratic society — that defense of reasonable faith — it can deliver no greater assurance of the rationality of what is actual.

C. Consensus? Really?

Those are the aims of Liberalism. They are likely to meet with skeptical response. The idea of combining disagreement on fundamentals with consensus on political principles suited to free and equal citizens may strike us as nice work if you can get it. In particular, it is natural to suspect that the demands of consensus are less minimal and the faith in its possibility correspondingly less reasonable than Rawls claims.

There are at least four reasons for skepticism about the ideal of consensus, and I will discuss them in detail in Part IV of this review.\(^\text{37}\) As background for that discussion, I want first to explore more fully Rawls's new view, tracing the route from Theory to Liberalism — in Part II — and outlining the strategy of Liberalism itself — in Part III. Before getting to the route and the strategy, however, I want to enter a caveat.

\(^{35}\) History, Hegel says, is the "true theodicy." HEGEL, supra note 32, at 457.


\(^{37}\) See infra Part IV. One basis of skepticism that I will not explore below endorses the possibility of combining political consensus and moral pluralism, but only if the political consensus is confined to questions of just procedure. I explore and criticize this view in Joshua Cohen, Pluralism and Proceduralism, 69 CHI.-KENT L. REV. (forthcoming Summer 1994).
D. A Different Book

Liberalism is a very abstract book, in ways that contrast sharply with Theory. Much of the excitement of Theory derived from its claim to argue from relatively weak, abstractly stated assumptions to powerful, controversial, substantive claims about justice. Here was an egalitarian and liberal account of justice, concerned both with the protection of basic civil and political liberties and with assuring a distribution of resources that would enable people to make fair use of those liberties, and supported by premises arguably much less controversial than its conclusions.38

Moreover, Theory’s many polemical edges helped to sharpen its central claims. Utilitarianism had dominated the field of systematic moral and political philosophy, and Theory aimed to displace it.39 In addition, Theory proposed an alternative to the ideal of natural liberty — sharp libertarian limits on the legitimate actions of the state — and to a liberal pluralism that would ensure fair process but would leave questions of substantive justice to bargaining in political and economic markets.40 To be sure, Rawls devoted stretches of Theory to the nature of justification, rationality, and goodness. But the discussion of these matters was never longer — or shorter — than necessary, and one felt that the discussion was never very far from first-order issues of justice.

By contrast, Rawls’s presentation of political liberalism puts substantive questions of justice aside. Here, Rawls does not focus on the content of justice but on whether justice as fairness can provide shared political ground given conflicting comprehensive moralities.

Moreover, Liberalism lacks the well-defined opponents of Theory. To be sure, Rawls contrasts the ideal of overlapping consensus on a political conception of justice with the communitarian aspiration to achieve social unity through a shared conception of human nature and the human good.41 But communitarianism lacks the sharp definition of utilitarianism, libertarianism, or liberal pluralism, contributing to the relentlessly abstract character of Rawls’s presentation.

38. A central claim in Theory is that we will be led to surprising, egalitarian conclusions about the limits of legitimate socioeconomic inequality by reasoning from the same fundamental ideas — about the equality of moral persons and our basic interests — that support familiar and settled convictions about the injustice of religious intolerance and racial discrimination. See RAWLS, supra note 3, at 19-20, 150-83. To make his case, Rawls gathers the less controversial claims and convictions together in the original position, thus requiring our reasoning about socioeconomic issues to conform to principles and ideas to which convictions about fairness and basic liberties already commit us. See Joshua Cohen, Democratic Equality, 99 ETHICS 727 (1989).

39. RAWLS, supra note 3, at vii-viii.

40. On natural liberty and liberal pluralism, see id. at 65-75. For an argument against natural liberty and liberal pluralism, see BRIAN BARRY, THEORIES OF JUSTICE 217-34 (1989). I contrast liberal pluralism with Rawls’s view in Cohen, supra note 37.

41. See pp. 42-43, 146, 201.
Because it pays less attention to substantive issues of political justice and lacks such sharply defined opponents, Liberalism is unlikely to generate either the excitement of Theory or the same interdisciplinary ferment. But these are caveats, not criticisms. Liberalism is a book of very great depth and importance. In due course it will likely change the shape of political philosophy, sharpening political philosophy's autonomy by increasing its distance from moral philosophy, and perhaps will have similarly salutary effects on political argument itself.

II. LIBERALISM: A PHILOSOPHY OF LIFE?

Liberalism, Rawls says, addresses "a serious problem internal to justice as fairness" (p. xv) — the view presented in Theory. In general terms, the problem arises from a lack of realism engendered by inattention to the fact of reasonable pluralism (pp. xv-xviii). More particularly, the difficulty emerges in the account of stability advanced in Part Three of Theory. To locate the difficulty more precisely, and to see why it is so troubling, I will first sketch three main elements of Theory, then present an objection that many commentators have raised about the main line of argument in Theory, and finally restate that difficulty as a tension internal to justice as fairness.

A. Three Elements of Theory

Theory presents, first, an attractive ideal of a just society — a well-ordered, democratic society, featuring a consensus on norms of justice. The content of the consensus is given by two principles:

[First Principle:] Each person has an equal right to a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all.

[Second Principle:] Social and economic inequalities are to satisfy two conditions. First, they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greatest benefit of the least advantaged members of society.

A society satisfying these principles achieves, Rawls proposes, some measure of "reconciliation of liberty and equality." Suppose that the real value of the freedom guaranteed to a person by the protection of basic liberties is fixed by that person's command of resources, rather than by her position relative to others. Then the two

42. I take the formulation of these principles, first stated in Theory, from Liberalism. P. 291.
43. RAWLS, supra note 3, at 204.
44. Rawls does not think that the worth of political liberty to a person is fixed by that person's absolute command of resources. Because the political process has "limited space," the value of political liberty also depends on relative position. See pp. 328-29. For this reason, Rawls imposes a special requirement of the "fair value" of political liberty: roughly, that people in different social positions have equal chances to hold office and influence the political process. See pp. 327-31, 356-63; RAWLS, supra note 3, at 224-27. For a discussion of relative positions, see id. at 530-41.
principles together require that a society "maximize the worth to the least advantaged of the complete scheme of equal liberty shared by all." This requirement of maximizing the minimum worth of liberty, Rawls says, "defines the end of social justice.

Second, Rawls offers a contractual defense of this egalitarian-liberal conception of justice. Carrying the social contract idea "to a higher order of abstraction" (p. xv), he argues that the two principles would be chosen in an initial situation of choice — the "original position" — in which the parties are assumed not to know anything particular about themselves — about their position in the distribution of alienable resources, their position in the distribution of native endowments, and the determinate aims, attachments, or views of the world that comprise their conception of the good. Required to choose under conditions of severe ignorance, they are uncertain of the effects of their choice on their own lives. Concerned to assure that they can live with that choice wherever they end up, the parties would choose to provide themselves, Rawls argues, with the strong downside protection assured by the two principles.

Third, Rawls proposes that the various constraints on knowledge imposed in the original position represent requirements that strike us, on reflection, as reasonable to impose on norms of justice or on their justification. Concerns about fairness, for example, and a conception of individuals as equal moral persons with a conception of the good and the capacity for a sense of justice fuel these constraints.

B. Original Position: A Liberal Philosophy of Life?

None of these three central elements of justice as fairness has won general acceptance. But criticisms of Rawls's claims about the reasonableness of the conditions imposed in the original position have been especially sharp among moral and political philosophers. Though the details of the criticisms take many forms, the central ob-

45. RAWLS, supra note 3, at 205.
46. Id. In Liberalism, Rawls says, less strongly, that maximizing the minimum worth of liberty "defines one of the central aims of political and social justice." P. 326.
47. RAWLS, supra note 3, at 17-22.
48. Id. at 136-42.
49. Id. at 150-57, 175-83.
50. See id. at 18, 587. For a complete list of passages in Theory that state the idea of the original position as expressing reasonable requirements on arguments for principles, see p. 25 n.28.
51. "If the original position is to yield agreements that are just, the parties must be fairly situated and treated equally as moral persons." RAWLS, supra note 3, at 141; see also pp. 23-27.
jection to Rawls's construction is that the design of the original position presupposes a particular conception of the good. It does not, contrary to Rawls's claims, provide a reasonable device for addressing controversies about justice among people with different conceptions of the good, because it will only be found attractive by people drawn to a liberal philosophy of life — one that holds that individual independence, choice, and self-mastery are the fundamental values that ought to govern our lives.

Critics have localized the offending bias in different places. Thomas Nagel criticized Rawls's assumption that all the parties in the original position want "primary goods" — in particular, income and wealth — as unfairly biased in favor of individualistic conceptions of the good.53 Brian Barry objected to the individualism implicit in Rawls's contractual method of justification, which proceeds from individual judgments about what is best for me, all else equal, to judgments about how society ought to be arranged.54 According to Michael Sandel, Rawls assumed a liberal philosophy of life when he required that we place our conceptions of the good behind a veil of ignorance.55

To see the force of these criticisms, consider Sandel's objection. Reasoning from behind the veil of ignorance requires that we evaluate norms of justice without reference to our own conception of the good. It is a puzzling idea. Why, and how, are we to reason about justice without drawing on our views about the proper conduct and ends of human life? If we hold the sincere conviction that a life of self-realization is a better life — if we think that such a life is genuinely better, not simply the life that we prefer — then what reason could there be for bracketing that conviction when we assess principles of justice?

One reason for such bracketing is that we cannot agree on terms of


55. Michael J. Sandel, Liberalism and the Limits of Justice (1982). Sandel sketches the liberal philosophy of life — a conception of "[t]he deontological universe and the independent self that moves within it." Id. at 177. William Galston also criticizes Rawls for failing to acknowledge his reliance on a conception of the good. See William A. Galston, Liberal Purposes 118-62 (1991). Galston argues, however, that such reliance is no embarrassment. On the contrary, liberalism must openly avow its dependence on a view of the good, albeit a "deliberately thin" view, "a kind of minimal perfectionism." Id. at 177. Galston's view is puzzling. It is not controversial that some account of the good is required for an account of justice. See pp. 173-211; Rawls, supra note 3, at 395-99. Moreover, Galston's account of the good is itself constrained by a concern "to provide a shared basis for public policy." GALSTON, supra, at 178. This constraint suggests that Galston's account of the good may not comprise part of a comprehensive perfectionist conception but may instead be part of a political conception of the good in the sense defined by Rawls at pp. 174-76. I say that Galston's account "may be" political because it is not clear what he means by a "shared basis of public policy" or how the concern to provide such a basis — as distinct from concerns within an account of the good — constrains the role of ideas of the good in his presentation of liberalism.
cooperation for a pluralistic society if the rationale for those terms premises a particular conception of the good. Peaceful cooperation requires agreement, and agreement requires that citizens put aside "the contingencies that set them in opposition." But Rawls's reasons in Theory are not simply a matter of securing social peace. He argues instead that fairness to citizens as moral persons requires that we not rely on any particular conception of the good in justifying principles that all will have to live by. Instead, fairness demands that "[t]he arbitrariness of the world . . . be corrected for by adjusting the circumstances of the initial contractual situation."

But why is it unfair to people as moral persons to treat them in accordance with principles of justice chosen on the basis of an account of the best life? Why correct for the "arbitrariness of the world" by abstracting from convictions about the best life? Why not correct for that arbitrariness by encouraging everyone to endorse the truth about the best life? To be sure, conceptions of the good sometimes set people in opposition; but why are they "contingencies"? According to Sandel's objection, Rawls's answer to these questions itself relies in the end on a particular account of the best life and a particular view of the person that goes with that account. We will only take an interest in what is chosen behind the veil of ignorance if we deny that our fundamental aims and attachments are good indicators of who and what we are. Moreover, we will be drawn to that denial only if we regard ourselves as, at bottom, agents unencumbered by fundamental attachments to our actual ends, as essentially choosers of values rather than as carriers and renewers of the values of particular traditions and communities — only if we are attracted to the idea that our basic allegiances themselves are elements of the arbitrariness of the world and that the unchosen life is not worth leading.

This will do as a statement of comprehensive liberalism, and we can understand a theory of justice built on these foundations as presenting the political implications of such a liberal outlook. But,

56. RAWLS, supra note 3, at 137 (emphasis added).
57. Id. at 141.
58. Sandel identifies two key assumptions in Theory: that we are essentially choosers — the priority of the self with respect to its ends — and that we are not essentially members of a community — "[t]he priority of plurality over unity." SANDEL, supra note 55, at 50-59. Notice that it is possible to deny the first proposition — thus affirming that our identity is fixed by our ends — without denying the second — that is, without affirming that we are essentially members of a community. I might regard myself as standing in an essentially personal relationship with God and as bound by obligations arising from that relationship, or as a locus of artistic creativity, or as essentially a seeker of truth. In each case, I might treat my relations with others as instrumental for those deeper purposes, rejecting the ideal of community. To put the point in historical terms, both Hegel and Nietzsche rejected the conception of the self as essentially a choosers of ends. But, not to put too fine a point on it, they had very different views about community. For criticisms of the conception of the self as chooser, see HEGEL, supra note 26, §§ 15-20, 105-141; and FRIEDRICH NIETZSCHE, ON THE GENEALOGY OF MORALS 44-46 (Walter Kaugmann ed., Walter Kaufmann & R.J. Hollingdale trans., Vintage Books 1969) (1887) (reprinted with ECCE HOMO).
according to the criticisms, the original position's dependence on such specific commitments disqualifies it from serving as a shared or neutral basis for settling on principles of justice in a democratic society whose equal citizens disagree sharply about liberal ideals of autonomy and individuality.

Sandel goes further. He thinks that Rawls's implicit commitment to a conception of the self as an essentially unencumbered chooser of ends is not merely morally sectarian; it is also inconsistent with Rawls's avowed aim of avoiding obscure and controversial Kantian metaphysical commitments and with our experience of both ourselves and our connections with our commitments. In short, Sandel is concerned not simply to demonstrate Rawls's own reliance on a view of the good but also to undermine the liberal conception of justice by exploding the views of the good and the self on which it depends. These further points are not, however, essential for our current purpose.

C. The Internal Problem: Congruence and Stability

Earlier, I mentioned Rawls's claim that Liberalism addresses a problem "internal" to justice as fairness. Thus far, however, I have presented an objection to the original position that might be thought to operate externally. I propose now to show how claims about the objectionable dependence of the original position on a particular philosophy of life can be turned into the internal tension in justice as fairness — the problem in Theory's account of stability — that Liberalism aims to address.

In characterizing the ideal of a well-ordered society and presenting an account of its stability, Rawls makes essential use of the idea of normative consensus. In a well-ordered society, "[e]veryone has a similar sense of justice and in this respect a well-ordered society is homogeneous. Political argument appeals to this moral consensus." Moreover, this shared sense of justice plays a "fundamental role" in ensuring that "the basic structure is stable with respect to justice."

To be sure, some idea of agreement figures in any contractual theory of justice. But the "moral consensus" Rawls refers to is not sim-

---

59. SANDEL, supra note 55, at 94-95.
60. Id. at 179.
61. The project of undermining liberalism by excavating and exploding its psychological and metaphysical commitments traces back to HEGEL, supra note 26. The most ambitious modern effort along these lines is ROBERTO M. UNGER, KNOWLEDGE AND POLITICS (1975).
62. See the introduction to Part II of this review, supra.
63. The following discussion draws on Joshua Cohen, Moral Pluralism and Political Consensus, in THE IDEA OF DEMOCRACY 270 (David Copp et al. eds., 1993).
64. RAWLS, supra note 3, at 263. On the role of consensus in the ideal of a well-ordered society, see p. 35. See also RAWLS, supra note 3, at 5, 453-58.
65. RAWLS, supra note 3, at 458.
ply an ex ante agreement on institutions and relations of authority of a kind associated with Hobbesian and Lockean social contracts.66

Closer in this respect to Rousseau, Rawls supposes that citizens in a just political society share a conception of justice and that politics is openly guided by that conception.67 Justice as fairness aims to specify the appropriate content for such a conception, the content of the general will for a society of free and equal persons.

This emphasis on the role of consensus in the ideal of a well-ordered society is understandable. A moral consensus on political fundamentals is a basic good for at least four reasons.

First, for any conception of justice, the existence of a moral consensus on it increases the likelihood that social order will stably conform to the conception.68

Second, a moral consensus promotes a variety of specific values of considerable importance. Assuming that norms of justice are not motivationally inert, consensus on them increases social trust and harmony, supports social peace, simplifies decisionmaking, reduces monitoring and enforcement costs by encouraging a willingness to cooperate, and — if public debate and decisions reflect the consensus — reduces alienation from public choices because citizens embrace the norms and ideals that guide those choices.

Third, a consensus on norms of justice provides a way to reconcile the ideal of an association whose members are politically independent and self-governing with an acknowledgment of the central role of social and political arrangements in shaping the self-conceptions of citizens, constraining their actions, channeling their choices, and determining the outcomes of those choices.69 When a consensus on norms and values underlies and explains collective decisions, citizens whose lives are governed by those decisions might nonetheless be said to be independent and self-governing. Each endorses the considerations that produce the decisions as genuinely moral reasons and af-

---


68. See Rawls's "third general fact" at p. 38.

69. See the discussion of full autonomy at pp. 77-78. Rawls distinguishes there between endorsing full autonomy as a political value and affirming autonomy as a comprehensive moral value, to be realized in all aspects of life and conduct. The concern to reconcile self-government with interdependence is central to Rousseau's project, though Rousseau's own presentation suggests that he thinks of self-government or moral liberty as a comprehensive moral value, tied to an account of our true nature. On moral liberty, see Jean-Jacques Rousseau, On the Social Contract (1762) [hereinafter Rousseau, Social Contract], reprinted in Basic Political Writings of Jean-Jacques Rousseau 139, 144-46 (Donald A. Cress trans. & ed., 1987); on our nature as free beings, see Jean-Jacques Rousseau, Discourse on the Origin and the Foundations of Inequality Among Men, in The First and Second Discourses 137, 189-90 (Victor Gourevitch ed. & trans., Harper & Row 1986).
firms their implementation.\textsuperscript{70}

Finally, under conditions of political consensus, citizens achieve a form of mutual respect. Each offers as reasons for a decision only considerations that others who are subject to political power take as reasons, and state power is exercised only within the bounds set by these reasons.\textsuperscript{71} The force of this point as a basis for mutual respect is increased by recalling the distinction I noted earlier between a unanimous, ex ante agreement and an ex post consensus on norms of justice that frame political debate.\textsuperscript{72} In a Hobbesian contract of subordination, everyone agrees to submit to a common agent, accepting the will and judgment of that agent as authoritative.\textsuperscript{73} Nothing in the content of the agreement — nothing manifest in political experience itself — directly expresses mutual respect.\textsuperscript{74} With a political consensus, by contrast, the authorization of power proceeds in terms that all citizens accept ex post — in accordance with reasons that are shared and therefore accepted by all who are subject to the power. That does provide a basis for mutual respect.

Consensus, then, has its virtues. But not every consensus is attractive. Those attractions depend on the content of the consensus and on the conditions under which it is sustained. Suppose, for example, that a moral consensus is attractive because it provides a way to make self-government — or association on terms of mutual respect — consistent with the unavoidable chains of political connection (see the third and fourth reasons stated above). Then the consensus must be freely sustained and not simply a form of enforced homogeneity. A consensus is free only if it is arrived at under conditions that ensure the possibility of individual reflection and public deliberation — for example, conditions that protect expressive and associative liberties.

Here we arrive at the internal problem of \textit{Liberalism}. Assurances of expressive and associative liberties — necessary if the consensus that defines a well-ordered society is to be free and attractive — are bound to be associated with moral, religious, and philosophical pluralism.\textsuperscript{75} But can the value of substantive consensus on justice survive such pluralism? Let us say that a society is liberal only if it strongly protects expressive and associative liberties. Then, to restate the question: Can there be political consensus and social unity, given the inevi-

\textsuperscript{70} We also need to add that everyone believes with good reason that the decisions express the shared norms and values.

\textsuperscript{71} See the discussion of legitimacy at pp. 136-37, 216-19.

\textsuperscript{72} \textit{See supra} notes 66-67 and accompanying text.

\textsuperscript{73} \textit{See Hobbes, supra} note 66, at 120-21.

\textsuperscript{74} \textit{But see} Rousseau, \textit{Social Contract, supra} note 69, at 197 ("Once the populace is legitimately assembled as a sovereign body . . . the person of the humblest citizen is as sacred and inviolable as that of the first magistrate.").

\textsuperscript{75} See the “first general fact” at p. 36.
table pluralism of a liberal society? Why, in particular, ought we to expect — as Rawls suggests — that the members of a well-ordered society regulated by Rawls's principles of justice will find the conditions imposed on the original position reasonable? According to the criticisms I referred to earlier, the original position assumes a liberal philosophy of life and presents the political extension of that philosophy. If these criticisms are right, then the comprehensive views that some members of a just society find attractive will likely lead them to reject the original position.

The discussion of stability in Part Three of *Theory* suggests that the criticisms are right. Because it does, Rawls concludes that his account of the stability of a well-ordered society is in trouble: that it is "not consistent with the view as a whole" (p. xvi).

In Part Three, Rawls advances a two-stage case for the stability of a society regulated by his principles of justice. The first stage focuses on the acquisition of a sense of justice — "an effective desire to apply and to act from the principles of justice [the two principles chosen in the original position] and so from the point of view of justice." Rawls sketches how the members of a just society could be expected, through membership in a series of institutions — from family, to the associations of civil society, to citizenship in the state — to acquire an understanding of and an effective desire to act from a sense of justice to which Rawls's principles give content.

The second stage shifts attention from the acquisition of a sense of justice to the *congruence* of that sense with a person's conception of the good. Here Rawls argues that the members of a just society would, with reason, regard the regulation of their conduct by their sense of justice — as given by the two principles — as itself good for them: that is, they would find their sense of justice congruent with their good, rather than regarding it as an unwelcome constraint on the pursuit of their good. If this claim about the good of a sense of justice is right, then we have an important force for stability in a just society.

Moral pluralism causes troubles for this happy picture. Consider one of the arguments for congruence: "acting justly is something we want to do as free and equal rational beings. The desire to act justly..."

76. I do not mean to suggest that other societies are not pluralistic. See supra note 5.
77. See supra section II.B. In his Tanner Lecture on "liberal equality," Ronald Dworkin defends a version of liberalism on the grounds of its continuity with a more comprehensive liberal outlook on life. See Dworkin, supra note 15, at 20-22.
78. Rawls presents the first part of the case in RAWLS, supra note 3, at 462-96, and the second part in id. at 513-77.
79. Id. at 567.
80. Id. at 462-96. See also Hegel's theory of the formation of the will through the various spheres of ethical life — family, civil society, and state. HEGEL, supra note 26, §§ 142-329.
81. See RAWLS, supra note 3, at 499, 501.
and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire.” 82 The claim that these desires have the same content rests on the argument from the original position. Or, as Rawls indicates elsewhere, the “sentiment of justice” is — for anyone who “understands and accepts the contract doctrine” — the very same desire as the desire to act on principles that would be chosen “in an initial situation which gives everyone equal representation as a moral person,” and also the same as the desire “to act in accordance with principles that express men’s nature as free and equal rational beings.” 83 In the original position, we are represented as free moral persons, so to act from the principles chosen there is to express our nature as free and not to “give way to the contingencies and accidents of the world.” 84

Moreover, the argument from the original position not only selects principles of justice but also requires that those principles take priority in regulating our conduct. To express “our freedom from contingency and happenstance,” 85 then, we need more than a sense of justice given content by the principles chosen in the original position. We must also give priority to our sense of justice, assigning it an authoritative role in the regulation of conduct.

A central element in the case for congruence and stability, then, is that members of a well-ordered society will develop a conception of their nature as free beings, will regard the expression of that free nature in their own conduct as a fundamental good, and will understand — because of their “lucid grasp of the public conception of justice upon which their relations are founded” 86 — that such expression requires acting from the principles of justice that would be chosen in the original position, giving those principles a special regulative role. 87

The case for the two principles, then, depends upon the case for stability; the case for stability depends in part upon the case for congruence; and the case for congruence depends upon an account of our “nature as free moral persons” 88 and the desire to express our nature as free. 89 But this line of dependence strongly suggests that the argument for congruence, and so the case for stability, depends upon a set

82. Id. at 572 (citation omitted).
83. Id. at 478.
84. Id. at 575.
85. Id. at 574.
86. Id. at 572.
87. The condition of “full publicity,” defined at pp. 66-67, requires public availability of the conception of justice and the full rationale for it.
88. RAWLS, supra note 3, at 572.
89. Rawls ties this argument to the Kantian interpretation of justice as fairness. See id. The argument is one of four he offers in support of congruence. It might be interpreted as an argument addressed to those who endorse a comprehensive Kantian view, rather than as one of four arguments that citizens generally will find persuasive. But Theory clearly offers it as the latter.
of moral commitments and self-understandings that some members of
a well-ordered society will reasonably reject. 90

For example, some citizens may think that their nature consists in
the possession of various natural, human powers, that the human good
consists in a perfection that fully realizes those powers, and that the
requirements of morality set out the conditions for such perfection.
Others may think of themselves as creatures of a God who imposes
obligations that bind their moral freedom. Such citizens accept moral-
ities that are, to use Kant's term, heteronomous. They, too, wish to
express their nature and not to give way to the contingencies and acci-
dents of the world. But it is unclear why they should find the original
position a plausible way to specify the content of their expression.
With Locke, they may suppose that their fundamental powers are the
capacity to understand and to act from the Creator's requirements and
that they express their nature by acting from those requirements. 91 To
be sure, adherents of such a view might reject the imposition of a reli-
gious establishment and affirm the importance of the free exercise of
religion. But they would do so because forced religious practice does
not fulfill basic religious duties and so provides no route to salvation, 92
rather than because a regime of religious toleration expresses their
"nature" as free moral persons. They do not acknowledge themselves
to have such a nature.

Some people, then, may reject the characterization of our nature as
free; they will be drawn neither to the reasonableness of the original
position as a rendering of their nature, nor to acting from the prin-
ciples selected there because such action expresses their nature. 93 Thus,
Rawls concludes that the conception of a well-ordered society
presented in Theory "is unrealistic ... because it is inconsistent with
realizing its own principles under the best of foreseeable conditions"
(p. xvii). Under the best of foreseeable conditions, a society that satis-
fies the two principles will be a society in which some citizens reject
the conception of our nature used in Theory to underwrite the original
position and the account of congruence. "The account of the stability

---

90. I have concentrated on the problem for congruence. The account of acquisition, how-
ever, faces a parallel difficulty. An account of the acquisition of a desire to act on principles must
explain why that desire, which is not instrumental, does not reflect a strange affection for rules.
In Theory, Rawls responds to this concern by explaining that moral principles can "engage our
affections" in part because acting on them expresses our "nature as free and equal rational be-
ings." Id. at 476. But this explanation leaves us with a gap in the account of acquisition in the
case of those citizens who do not see their nature in such terms. In this connection, see pp. 82-86
on principle-dependent and conception-dependent desires.

91. See Locke, supra note 17, at 310-11.

92. See John Locke, A Letter Concerning Toleration 18-20 (Patrick Romanell ed.,

93. They may, of course, be attracted to those principles and to the original position itself for
other reasons. See infra notes 110-12 and accompanying text.
of a well-ordered society in Part III is therefore also unrealistic . . .” (p. xvii).

How, then, is it possible to achieve consensus on a conception of justice suited to a democratic society of equal citizens and to reap the benefits of that consensus, given the pluralism of comprehensive moralities that inevitably marks such a society? More particularly, can the presentation and defense of a conception of justice for a democratic society be freed from the unacceptably narrow premises of a comprehensive moral liberalism? That is the question of political liberalism.

III. POLITICAL LIBERALISM

Rawls’s answer to the question of political liberalism contains two parts. The idea of a political conception of justice plays a central role in the first part; the idea of an overlapping consensus is the key to the second.

A. A Political Conception of Justice

Given the plurality of comprehensive moralities, the claim that consensus is possible faces a threshold problem. A conception of justice can win general acceptance only if it can be suitably formulated. Its formulation must be understandable to citizens with competing views of the good and must not itself preclude acceptance by some citizens.94 Some conceptions of justice would, however, on their face, be unacceptable to some citizens — for example, if the conceptions appeal to values that are not implicated in public institutions or that reasonable people might reject. Suppose an account of justice requires a distribution of resources that ensures equal pleasure, or suppose it mandates a distribution that enables each citizen to come equally close to achieving his aims. Both views face troubles because citizens reasonably disagree about the relative value of pleasure and of relative achievement. So these conceptions would be, on their face, unacceptable.95

Rawls calls a view that is suitably formulated a “political conception of justice” (p. 11). Three features — each necessary if the conception is plausibly to provide the focus of agreement, given the fact of reasonable pluralism — define such a conception:96 it must have limited scope, extending only to issues about the basic structure of society

94. This condition is necessary but not sufficient, because a view that is formulated without reference to any comprehensive moral view may nevertheless be attractive only to those who hold a particular view. See infra notes 109-12 and accompanying text.


96. See pp. 11-15.
and not to norms of personal conduct or ideals of life; it must draw on
ideas familiar to citizens from the political culture of a democracy, not
on ideas belonging exclusively to particular traditions of moral
thought that are not available to all; and it must be presented as free-
standing, not as depending for formulation or justification on its roots
in a comprehensive morality. In short, a political conception of jus-
tice is formulated as autonomous from comprehensive conceptions of
the good with respect to scope, content, and justification. Each of
these three forms of autonomy should contribute to the possibility of
its general acceptance.

To see how these kinds of autonomy help to address the problems
about the original position and veil of ignorance I sketched earlier,
consider the aspect of the political conception that Rawls refers to as a
"political conception of the person" (pp. 18-20, 29-35, 48-54, 86-88).
The original position isolates certain features of people as relevant to
its problem of justice, setting aside other features as irrelevant — and
so to be excluded by the veil of ignorance. The relevant features in-
clude certain basic moral powers: the capacities for a conception of
the good — to form, pursue, and revise such a conception — and for a
sense of justice. The irrelevant features include sex, race, natural
abilities, and determinate conceptions of the good.

As my earlier discussion of the original position indicates, Theory
was not entirely clear about the basis of this distinction between rele-
vant and irrelevant characteristics. This lack of clarity contributed to
the impression that justice as fairness was the political expression of a
comprehensive moral liberalism. Thus, Rawls often referred to the
morally relevant or irrelevant as if to say that the distinction derives
from a comprehensive moral doctrine. Sometimes he referred to the
irrelevant characteristics as "contingencies," as though to suggest a
metaphysical foundation for the distinction. Sometimes — as I in-

97. A political conception is presented as freestanding in the way that logic or number theory is. A presentation of logical laws — for example, the law of excluded middle — proceeds without tying the laws to a theory of truth or issues in the theory of meaning; a presentation of number theory proceeds without reference to questions about the ontological status of numbers.
98. Pp. 18-20; RAWLS, supra note 3, at 561.
99. See supra notes 79-93 and accompanying text.
100. For example, Rawls refers to the constraints in the original position as "conditions that are widely recognized as fitting to impose on the adoption of moral principles." RAWLS, supra note 3, at 584. Similarly, the criticisms of natural liberty and liberal equality refer to social circumstances and natural assets as features that are "arbitrary from a moral point of view." Id. at 72, 74-75.
101. I say "sometimes" because some passages in Theory strongly suggest that apparently metaphysical notions should be interpreted morally. Take, for example, the following remark: "Our moral sentiments display an independence from the accidental circumstances of our world, the meaning of this independence being given by the description of the original position and its Kantian interpretation." Id. at 475 (emphasis added). To say that the description of the original position gives the meaning of independence is to say that independence is a matter of the irrele-
vance for moral purposes of certain features of the person, rather than a matter of the metaphysi-
cal contingency of those features.
dicated in my discussion of congruence and stability\textsuperscript{102} — he suggested that the distinction is rooted in an account of “our nature,” permitting both metaphysical and moral interpretations.\textsuperscript{103}

Liberalism draws the distinction between relevance and irrelevance in the same place: the power to form, pursue, and revise a conception of the good and the power to form and act from a sense of justice are relevant; and sex, race, natural abilities, and determinate conceptions of the good are irrelevant (pp. 29-35). But the point of the distinction, according to Liberalism, is to present a conception of the person that will play a role in a political conception of justice, and so Liberalism underscores that the conception of the person is itself political in each of the three ways noted earlier: scope, content, and justification. Thus, irrelevant should not be understood absolutely, metaphysically, or in terms of a general moral view, but only as implying that a feature of a person is not important for the purposes of political argument — in particular, not important for political argument aimed at specifying the requirements of justice for a society in which members are understood as free and equal. Contingent ought similarly to be given a nonmetaphysical rendering, as implying that a feature is not relevant to political argument.

We can, then, determine which features are “irrelevant, politically speaking, and hence [to be] placed behind the veil of ignorance” (p. 79) by systematizing and extending reasonably familiar ideas about the justification of political arrangements in a democratic society. This basis is appropriate for the distinction given the question that justice as fairness sets out to resolve: “What is the most appropriate conception of justice for specifying the terms of social cooperation between citizens regarded as free and equal, and as normal and fully cooperating members of society over a complete life?” (p. 20). The conception of citizens as free and equal represents a familiar element of the political culture of democratic societies. The problem is to determine more precisely what that political conception involves and to address a longstanding controversy about what account of justice is best suited to citizens as free and equal.\textsuperscript{104}

Thus, we look to settled ideals and convictions about basic democratic institutions, and to settled understandings about the justification of public norms in a democratic society, and then draw the relevant-irrelevant distinction by reference to the characteristics of persons that play a role in those ideas, convictions, and understandings. One may then call the irrelevant features “contingencies,” but with no intention to affirm — or to deny — that an individual could exist without the feature in question, or to say — or to deny — anything about the

\textsuperscript{102} See supra notes 78-93 and accompanying text.

\textsuperscript{103} RAWLS, supra note 3, at 251-57, 572.

\textsuperscript{104} See pp. 20, 22, 26, 34-35.
importance of irrelevant features in other settings. They are simply unimportant for the purposes at hand, whatever their metaphysical standing and however important they may be for other purposes, including other ethical purposes.

To be more specific, arguments aimed at establishing that certain properties are contingent (irrelevant to the problem of political justification) and that others are aspects of our essential nature (important to that problem) proceed along at least two main lines. The first seeks to show that current ideals—for example, of fairness, religious toleration, and racial and sexual equality—and patterns of political argument—for example, on constitutional matters—treat certain facts as irrelevant. For instance, it is widely agreed that we ought to protect certain basic rights—expression, political participation, conscience, and equal treatment—without regard to social background, sex, or race. Furthermore, social class ought not to restrict opportunity. These are clear cases of unfairness. So in reasonably settled understandings of justice, we treat facts about class, sex, and race as contingencies—matters that are irrelevant to argument about the justice of basic institutions.

Similarly, the constitutional treatment of religious and political ideals suggests the irrelevance of conceptions of the good to such argument. For example, conversion, sin, and religious laxity are not civil offenses. Whatever its implications for a person’s self-conception, being “born again” has no civil consequences; being born again does not, for example, absolve a person of contractual obligations undertaken prior to that rebirth or give a person who is reborn on election day a right to a second vote. Furthermore, in the case of political ideals, endorsing the legitimacy of the political order is not—in principle, at least—a precondition for equal political rights, a point underscored by conventional hostility to regulating expression by virtue of its content and, more particularly, its viewpoint. 105

A second strategy is to show that certain features of people are themselves so dependent on concededly irrelevant facts that to permit them to play a role in political justification would be tantamount to allowing the irrelevant facts to play a role. So they too should be treated as irrelevant. The development of abilities and talents, for ex-

ample, seems closely linked to the social circumstances and aspirations that the entrenched forms of argument fix as contingencies. So talents and abilities ought to be treated as contingencies and not appealed to as fundamental reasons for differential advantages.

I would need to say much more about these matters in order to evaluate Rawls's distinction between relevant and irrelevant, and the associated political conception of the person. I have provided only an outline of the rationale for the distinction. But its force — and limits — as a response to the original position's difficulties should now be clear.

According to the objection, the original position rests on a liberal philosophy of life that places especially great weight on the importance of choice and that sees the self as, in its fundamental nature, a chooser of its own ends. Rawls's claim that "the self is prior to the ends which are affirmed by it"\textsuperscript{106} suggests a commitment to such a philosophy. But the political conception of the person offers a restrictive interpretation of this priority. It neither affirms nor denies that people could, as a metaphysical matter, exist without their aims as pure choosers of ends, as "Kantian transcendent or disembodied subject[s]" who are "shorn of empirically-identifiable characteristics"\textsuperscript{107} or that citizens can imagine their own lives continuing with their final aims different from what they now are; or that they would actually be the same persons if their final aims were radically altered; or that, as an ethical matter, the aims of citizens are worth pursuing only if chosen by them. Instead, the political conception ties both the content of and the rationale for the alleged priority to the aims of a theory of justice for a democratic society and to the public availability of the idea of citizens as equals.

According to the political conception, citizens are prior to their ends in that no particular ends are mandatory from a public point of view, and citizens must be assured favorable conditions for reflecting on and revising their aims, should they wish. For example, obligations that a person has by virtue of her conception of the good do not have public standing as obligations. Moreover, civil standing does not alter with shifts in fundamental aims, no matter how much a person’s self-conception is bound up with those aims. This is not to say, however, that all obligations are matters of self-legislation, or that fundamental values are a product of choice, or that they are only worth pursuing if they are such a product. The political conception of the person does not state a position on these matters.\textsuperscript{108} That conception is simply a statement about how citizens should be represented for the purposes of

\textsuperscript{106} RAWLS, supra note 3, at 560.

\textsuperscript{107} SANDEL, supra note 55, at 95.

\textsuperscript{108} The political conception does not take a position in the way that statements of logical laws do not, on their face, take a position about the nature of meaning.
political argument. For this reason, nothing in the very statement of
the political conception of the person conflicts with comprehensive
moralties that are not organized around the ideal of autonomy or
around the thought that we are, by our nature, free beings.

B. Overlapping Consensus

Suppose this enterprise of reinterpretation succeeds — that a lib-
eral conception can be formulated as a freestanding political doctrine,
facially independent of any comprehensive moral conceptions. Pro-
viding this formulation would help in securing social unity under con-
ditions of moral pluralism. It would overcome the threshold problem
that I disclosed earlier. But it would not suffice to defeat the objec-
tions to or the associated internal troubles for Rawls's view.

The objection to the original position was not that its very state-
ment reveals it to be part of a liberal philosophy of life but rather that
citizens will be drawn to it — will find it a reasonable device for set-
tling on principles of justice — only if they endorse such a philosophy.
So, too, even if the formulation of a political conception is freed from
objectionable sectarianism, it may still win support only from adher-
ents to a single comprehensive doctrine or a narrow range of such
doctrines. Consider an analogy: logical laws can be formulated in a
freestanding way, independent of controversies in the theory of mean-
ing. Still, certain logical laws — such as the law of excluded middle —
will arguably be found compelling only by people who hold particular
views in the theory of meaning — for example, that we can understand
the meaning of statements whose truth or falsity transcends our recog-
nitional capacities.110

Take the claim that people with different conceptions of the good
have the capacity to choose and revise their conception, as well as a
fundamental interest in circumstances that enable them to revise it
should they wish. This claim is an element of the political conception
of the person, and it is one of the aspects of the person known behind
the veil of ignorance. In presenting a political conception of the per-
son, Rawls shows that endorsing this claim does not consist in believ-
ing that reflectively held convictions are uniquely worthy of our full
allegiance or that we are essentially choosers of ends rather than ser-
vants of God; by formulating the political conception of the person as
a freestanding view, he shows that that conception does not imply any
particular nonpolitical view of the person, for the content of the polit-
ical conception is very different from the content of any such view.
Nevertheless, it may be true that we only have good reason to accept
the political conception and the associated account of justice if we en-
dorse a comprehensive liberal philosophy of life.

109. See supra notes 94-95 and accompanying text.
Therefore, we may be misled when Rawls says that "accepting the political conception does not presuppose accepting any particular comprehensive . . . doctrine; rather, the political conception presents itself as a reasonable conception for the basic structure alone" (p. 175; emphasis added). Even if the conception presents itself as political, accepting it may still presuppose accepting a comprehensive view if a single view provides the only reasons for accepting the political conception.

Here, then, we need the idea of an overlapping consensus: the idea that all people can — for the different reasons provided by their own reasonable comprehensive moral views — think that the same conception of justice is correct and not merely an accommodation required to ensure a stable peace under conditions of moral pluralism. Rawls imagines, for example, an overlapping consensus composed of four views, each of which is reasonable and each of which provides a rationale for political liberalism: one rooted in a Kantian morality of autonomy, another in utilitarianism, and a third in a religious conception that endorses free faith, while the fourth treats political liberalism as one part of a pluralistic ethical view — a part that needs to be adjusted to the other parts, though it is not derived from them.

Consider, for example, the political conception of citizens as free. How might these four views endorse the idea that citizens are free as a shared basis for political argument? One aspect of political liberalism — captured in the veil of ignorance — is that citizens have the capacity to revise their aims and an interest in favorable conditions for such revision should they wish to pursue it, but that for the purposes of an account of justice the determinate aims of citizens are irrelevant. The Kantian view accepts this aspect of political liberalism because the Kantian conceives of the reflective choice of ends as a feature of an autonomous life and holds that the protection of citizens who wish to pursue such choice is required by respect for their dignity as autonomous. The utilitarian might endorse the interest in revising aims as fundamental because true happiness — whether consisting of pleasurable feelings or the satisfaction of rational desires — depends on the possibility for such revision.111 The conception of free faith also endorses this interest because of its connections with the appropriate fulfillment of religious obligations: that such fulfillment must reflect genuine "inward persuasion of the mind."112 In short, each view accepts, for its own reasons, a conception of persons and their basic interests that provides shared ground in political argument.

But an overlapping consensus on a conception of justice cannot be sustained simply by the existence of points of agreement, for points of

---

111. See, e.g., MILL, ON LIBERTY, supra note 14, at 116-17. Similar considerations would support a case for the interest within a view emphasizing self-realization.

112. LOCKE, supra note 92, at 18.
disagreement among reasonable views are bound also to exist. Each view implies that the others are a mixture of truths and falsehoods. Why, then, should citizens who endorse a particular moral view — who believe it to be true — not hold that political power ought to be used to advance the values of that view? Why should they endorse as correct a view of justice that is confined to shared ground and accept that public discussion must provide justification according to that view? Three considerations explain this restraint.

1. It is worth emphasizing again that citizens who hold competing comprehensive views may nevertheless agree that the values incorporated within the political conception are important values and that the norms and principles included in it provide genuine reasons. From within each comprehensive view, the political conception states nothing but the truth, even if not the whole truth. As my example about the interest in favorable conditions for revising aims indicates, adherents to different moral conceptions do not think that the political conception reflects a compromise required to ensure a stable peace. Instead they believe that the conception expresses a correct account of basic political interests.

2. In accepting as correct a conception of justice that does not include the whole truth, by their lights, citizens acknowledge both the reasonableness of at least some of the views that conflict with their own and the unreasonableness of imposing arrangements whose justification depends on aspects of their own view that others reasonably reject.

The Kantian, for example, rejects the utilitarian conception of the good as the satisfaction of rational desires, but he can understand the utilitarian view as an application of theoretical and practical reason, appreciate the considerations that lead to that view, and see how its endorsement is compatible with a willingness to cooperate on terms that others can accept. So the Kantian's endorsement of a political conception that contains only part of the truth — that takes political autonomy rather than moral autonomy as a fundamental value in political argument — is not simply a compromise required by the existence of other views. Instead, the Kantian thinks it would be wrong to impose institutions and policies justified by a political conception that is rejected by others who are themselves fully reasonable.

113. It might be said that holding a moral view is a matter of having pro-attitudes rather than beliefs that are apt to be true or false. For a sketch of the difficulties in sustaining this position, see Paul Horwich, Gibbard's Theory of Norms, 22 Phil. & Pub. Aff. 67 (1993) (book review). But see Michael Smith, Why Expressivists About Value Should Love Minimalism About Truth, 54 Analysis 1 (1994), and the reply by Horwich in Paul Horwich, The Essence of Expressivism, 54 Analysis 19 (1994).

114. Rawls mentions the first two considerations at pp. 127-28 (referring to Cohen, supra note 63).

115. See supra note 69.
3. As the second point suggests, the key to the possibility of overlapping consensus is that a conception of justice articulates values of great importance and that the existence of a shared political conception itself constitutes an important good. I suggested a case for these claims in my earlier remarks about the good of consensus on a conception of justice.\textsuperscript{116} Suppose that case is correct and that the political consensus does articulate important values. Suppose, too, that different, conflicting comprehensive moral conceptions agree on a conception of justice. Then adherents to those moral conceptions will be able to say — each from her own standpoint — that it is normally best to uphold institutions satisfying the conception of justice, even when policies selected by the institutions are inconsistent with her particular moral conception. These conflicting moral views will also agree that it is normally best to conduct public discussion about political fundamentals in terms of the values and principles of the political conception rather than to appeal to a particular comprehensive moral view that others reasonably reject.

Much here rests on “normally.” Views that form an overlapping consensus will rarely, if ever, hold that political values are ultimate. For that reason, there may well be occasions when a comprehensive moral view supports the conclusion that the stakes are too high and that political values must give way. Adherents to such a view may be optimistic and see deep disagreement as an occasion for a high-stakes effort to persuade others to drop their ultimate convictions; more likely, however, they will think that the time for debate has ended. Because political values are not widely regarded as ultimate values, this kind of breakdown is always possible. To that extent the bases of civic unity are fragile: such fragility is the inevitable result of the pluralism of comprehensive moralities.

Despite this fragility, one can hope that civic breakdown will not occur. More immediately, the existence of cases in which it does occur, together with the fact that we all have more to say than we are prepared to say in politics, does not imply that consensus is impossible or unattractive, or that operating on the shared ground of a political conception of justice is merely a compromise dictated by circumstance.

IV. CONSENSUS?

I said earlier that the idea of consensus is likely to elicit a skeptical response,\textsuperscript{117} and I want now to explore some of the sources of that skepticism. I will consider four objections to the idea of an overlapping consensus.\textsuperscript{118} Because I find the idea of consensus attractive, I

\textsuperscript{116.} See supra notes 67-74 and accompanying text.
\textsuperscript{117.} See supra note 37 and accompanying text.
\textsuperscript{118.} The first objection I will consider overlaps with the fourth objection discussed by Rawls
will present replies to each of the objections. The four objections form a natural sequence, beginning from the thought that it is simply naive to expect consensus in a large-scale political society. The second and third objections present different variants of a common concern: that the case for consensus reveals that it can be achieved only through an objectionable exclusion of views that fall outside the consensus. The fourth objection accepts the possibility of consensus but argues that an overlapping consensus truncates political argument; by effectively taking comprehensive moral views as given, overlapping consensus fore-stalls the deeper agreement that might emerge from a more vigilant political criticism.

A. Hopelessly Naive

Consider the depth and extent of disagreement on any important political issue: from abortion and taxes to health care reform and trade policy. Against this background of disagreement the idea of consensus may strike us as hopelessly naive. This objection gains added force from Rawls's rejection of the possibility of comprehensive moral agreement. If we are prepared to exclude convergence on morality quite generally — to affirm the fact of reasonable pluralism as a "permanent feature of the public culture of democracy" (p. 36) — why should we find agreement on a political conception of justice plausible?

It will not suffice to say that political agreement is more plausible than comprehensive moral agreement because matters of political justice are a proper subset of moral issues, and agreement on a proper subset is more likely than agreement on the wider set itself. Issues about abortion are a subset of the moral, but I think most of us would be nearly as surprised by consensus on the morality of abortion as by consensus about morality in general. Moreover, it is not enough simply to point to the possibility of agreement on a political conception of justice among people who have different comprehensive views. That possibility is established by the coherence of the idea of an overlapping consensus. But the coherence of that idea does not suffice to show that it is any more realistic than agreement on comprehensive moral views, which is also possible.

To answer these doubts, we need a mechanism — a social or political process that might produce convergence on political values but that does not similarly generate consensus on comprehensive moral values. The right place to look for such a mechanism is at the level of shared institutions, as they might plausibly play an educative role with respect to political ideas, but not with respect to comprehensive moral

---

at pp. 158-68, though my reply differs from Rawls's in important details. The other three objections I will discuss differ from those Rawls considers at pp. 145-58.
conceptions. Before explaining this role, however, I need to make two background points.

First, it is worth emphasizing that we are concerned with agreement on conceptions of justice, not with a convergence of interests. Of course, if people are moved principally by interests, then the absence of such convergence may imply that agreement on justice is not a matter of great moment. Still, the immediate issue is convergence on justice — which, after all, seems less hopeless than an absence of conflicts of interest.

Second, the agreement on justice will be limited in various ways; it will not extend to all judgments of policy or even to all fundamentals that might possibly arise. In overlapping consensus, agreement on procedures and basic protections — in Rawls's terms, on constitutional essentials and matters of basic justice — suffices to make the remaining disagreements less important or less immediate.

Even with these two points of clarification, it may still seem unrealistic to expect agreement on matters of basic justice, given persisting differences in moral outlook. But perhaps we can address this concern about realism if we keep in mind the institutional aspect of the acquisition of political ideas and values. Although it is implausible to expect agreement on a conception of justice to result from a convergence of practical reasoning conducted within different, independent moral traditions, it is not so implausible to expect such agreement to emerge from the acquisition of ideas and principles embodied in shared institutions. As I indicated in the earlier discussion of stability, Rawls's views about the development of moral-political understandings are deeply institutional. The acquisition of conceptions of justice proceeds via participation in institutions of various kinds — families, associations, the state. The formation of moral-political ideas and sensibilities also proceeds less by reasoning or explicit instruction — which may be important in the formation of comprehensive moral views — than by mastering ideas and principles that are expressed in and serve to interpret these institutions. The underlying idea — which traces to Rousseauian and Hegelian theories of will formation — is that people living within institutions and a political culture shaped by

119. This distinction is implicit, I believe, in Rawls's remarks on the "wide role" of a political conception "as educator." When a political conception is fully public, citizens "are presented with a way of regarding themselves [as free and equal] that otherwise they would most likely never be able to entertain." P. 71.

120. I do not mean to deny that convergence of independent traditions is a possibility; my point is that an account of political consensus should not depend on it. Bernard Williams has argued that if there were moral consensus it could not be explained by the (perspective-independent) truth of the moral beliefs on which different traditions converged. See BERNARD WILLIAMS, ETHICS AND THE LIMITS OF PHILOSOPHY 132-55 (1985). Rawls's account of the possibility of consensus on a conception of justice does not require that the truth of the conception explains the agreement on it.

121. See pp. 158-68; RAWLS, supra note 3, at 462-79.
certain ideas and principles are likely to come to understand those ideas and principles and to develop some attachment to them.\footnote{122}

Take, for example, an aspect of the political conception of the person — the (political) idea that citizens are equals in possessing to a sufficient extent the capacity for a conception of the good and for a sense of justice. This idea is manifest in various ways in the practices and traditions of interpretation and public discussion associated with citizenship in a democracy: for example, equality before the law, or equal civil and political rights. Moreover, a stable democratic political process, in which individuals and parties seek to win support for their projects from other citizens, puts some pressure on views to endorse the idea of citizens as equals.\footnote{123} We can understand how citizens quite generally might acquire an understanding of one another as moral equals by holding the position of citizen and living in a political culture in which ideas of equality associated with that position play a central role in political discourse.\footnote{124}

The different comprehensive views that accept this political understanding of equality will have different ways of fitting it into their broader conceptions. Some will accept political equality as following from their more fundamental moral or religious convictions; others will accept political equality as an important, nonderivative value. But what keeps the expectation of general agreement from being hopelessly naive is the plausible thought that citizens who grow up within a reasonably stable democracy will find this (self-)conception familiar and attractive: the political ideas "expressed" in common, public institutions and appealed to in the culture to justify those institutions will shape citizens' moral-political education.

Of course, the acquisition of moral ideas does not proceed exclusively through institutions. So citizens will need to find or to make a place within their comprehensive views for the political ideas and self-conceptions they acquire through institutions: to find a way to combine, for example, a conception of human beings as servants of God bound by natural duties with a political conception of citizens as free, equal, and self-governing. Many views — religious, moral, philosophical — have sufficient internal flexibility or openness to make such ac-


\footnote{123. For further discussion, see Cohen, \emph{supra} note 37.}

\footnote{124. Consider in this connection the virtually unanimous popular endorsement of political equality and equality of opportunity indicated in Herbert McClosky & John Zaller, \emph{The American Ethos: Public Attitudes Toward Capitalism and Democracy} 74 & tbl. 3-5, 83 tbl. 3-9 (1984).}
commodations possible. But because political values are a subset of moral values, we have no reason to expect the accommodation of shared political values to produce a more comprehensive agreement that extends to moral values generally; no institutional mechanism in a democratic society imposes pressure to overcome fundamental differences among moral, religious, and philosophical traditions. The pressure of the shared institutions in forging political agreement ends even as considerable disagreement remains.

To be sure, this explanation provides only the barest sketch of a reply to the objection about realism, but it makes an essential point that is commonly overlooked when political philosophy is understood simply as applied moral philosophy. Political ideas are institutionalized in a democratic society in ways that comprehensive moral — or religious or philosophical — ideas are not. More precisely, comprehensive ideas are institutionalized — if at all — in more particular social associations that are not shared: different churches, for example, advance different comprehensive views. So citizens acquire conflicting comprehensive views through such associations. Political ideas, by contrast, are acquired in part through shared associations. So an account of how consensus might emerge on a political conception of justice among citizens living in a political society can draw upon resources unavailable to an account of a more comprehensive moral consensus. Of course no political mechanism can guarantee agreement: the development of an overlapping consensus requires, as I mentioned, that separate traditions are each able to accommodate the political values within their view, and nothing guarantees that they are able to do so. But we are not looking for a guarantee; we only need a mechanism that might plausibly produce convergence of political values even under conditions of moral pluralism.

Finally, given the institutional explanation, it is not surprising that the political consensus is itself limited, being principally a matter of agreement on basic political values — such as fairness, equality of citizens, and liberty, for example — rather than an agreement on a definite conception of justice. For no definite conception — no specific interpretation and balancing of the basic political values — is institutionally expressed in the way that the basic values themselves are. Of course there may be an optimal way to articulate and combine those values, and then the underlying agreement may recommend a specific conception. But that is a matter for further argument — for political philosophy. It is not a conclusion that is manifest from the values themselves or from their institutional articulation.

125. See pp. 159-61.
126. Recall the contrast I drew earlier between Hegel and Rawls, supra section I.B.
127. The claim that there is such an optimal way provides the basis of Rawls’s argument for justice as fairness. See p. 9.
B. Unattractively Explained

Let us suppose that this explanation of the difference between the expectations of political and comprehensive moral consensus can be sustained. Then, a second objection seems natural: that the institutional explanation limits the attractiveness of the consensus it explains. An attractive explanation would see political consensus as emerging from a convergence of argument within conflicting moral and religious traditions, or perhaps from unconstrained practical discourse among adherents of separate traditions.128 In either case, political consensus would reflect the operation of reason, driving separate moral positions to common political conclusions.

By contrast, the explanation I have just sketched traces the emergence and reproduction of political consensus to shared background institutions. Through these institutions, citizens acquire moral-political ideas—including ideas of person and society. Moreover, the role of the institutions is crucial, because the content of a political conception for a democratic society does not rely only on practical reason; rather, it draws also on “political conceptions of society and person”—in particular, the idea of citizens “regarded as free and equal in virtue of their possessing the two moral powers to the requisite degree.”129 Thus, the political conception of justice expresses an ideal of political deliberation and justification in a democratic society, not a more generic conception of justification through reason.130 So it is especially implausible to think that the political conception might arise simply from the work of practical reason within and among traditions. Precisely this implausibility, however, may make an overlapping consensus seem less a result of free reflection than a product of the institutional constraints under which political argument proceeds.

This criticism rests on an exaggerated distinction between institutional constraint and free reflection. Recall the background assumptions: the deliberative liberties are in place—and have a fair value—and the society features a range of comprehensive views, which provide intellectual and practical elaborations of different moral, religious, and philosophical traditions. Suppose now that as a consequence of democratic institutions and the position of equal citizen within these institutions, the members of such a society acquire a shared understanding of the equality of moral persons. Suppose, too,

129. P. 109; see also pp. xx, 107-10.
130. See also Joshua Cohen, Deliberation and Democratic Legitimacy, in The Good Polity: NORMATIVE ANALYSIS OF THE STATE 17 (Alan Hamlin & Philip Pettit eds., 1989). I emphasize there that a conception of reasons suited to the ideal of deliberative democracy reflects an ideal of free deliberation among equals. Id. at 22-23.
that citizens adjust their comprehensive views — if they have them — to accommodate this shared understanding. For example, they adjust their conceptions of flourishing and true happiness to the many directions in which citizens develop and pursue their native abilities; they adjust their conception of the conditions required for salvation to accord with the circumstances of a political society that includes citizens of different faiths; and they adjust their views of the “nature” and “proper conduct” of men and women to take account of the equality of men and women as moral persons. Under these conditions, we face strong pressure to regard the acquisition of shared ideas and the adjustment of comprehensive views as a matter of learning rather than mere inculcation via institutional constraint: how, we may ask, does the inculcation work, given a background of deliberative liberties with a fair value? Why are the shared ideas that emerge resilient in the face of challenge?

Of course, we can only presume learning. Someone may be able to show how the agreement reflects power, limited information, confusion and weakness born of moral cacophony, or a deep disparity between the apparent logic of institutions and their real operation. But the presumption is significant and imposes a serious burden on those who would treat the agreement merely as a product of inculcation and constraint.

Consider again the political conception of the person: in particular, the idea of the equality of citizens as rooted in their possession of a capacity both for a sense of justice and for a conception of the good. Assume that people brought up in a just, democratic society find this conception compelling, and that this is so whether their comprehensive views are secular — perfectionist, utilitarian, Kantian — or religious.131 Suppose further that considerations within their own comprehensive views support the conception of citizens as moral equals. But suppose also that citizens reflect on the fact that their traditions would likely have evolved differently under different institutional conditions; had their traditions not been subjected to these particular institutions, the traditions would not now provide the resources to support the political conception. If, for example, these same citizens had been raised in a more hierarchical society, their conceptions of flourishing, salvation, and gender might not be so egalitarian. How, they might ask, could the fact that a conception of justice is rooted in the political conception of the person give any special weight to the conception of justice, given the historically contingent attractiveness of the conception of the person?

The problem with this objection is that it neglects the content of

131. To be sure, important historical strands of these views have rejected the political conception of equality. But we have already rejected the idea that the political conception must emerge from the separate elaboration of competing traditions.
the institutional conditions under which the political conception of the
person emerges. Recall that we are assuming that the deliberative lib-
erties of citizens are secure and that citizens have a fair chance to exer-
cise those liberties. Though the political conception of the person does
not arise through reasoning that proceeds outside an institutional set-
ting, it must successfully withstand pressures arising from the institu-
tionalization of deliberation itself, from freedom of expression and
association, and from a fair distribution of resources. The attrac-
tions of the political conception of the person, then, are assumed to
survive criticisms that might be directed against it. If they do survive,
then how could the mere fact that people would find other views at-
tractive under different circumstances provide a reason for rejecting
the views that they do hold? The fact that citizens' views are in part
institutionally explained should not lead us to think that an allegiance
to them is merely a product of political circumstance rather than free
reflection, given the specificity of the institutions and their role in pro-
tecting public deliberation.

C. Objectionably Exclusionary

The third objection begins from the observation that the difficulty
of achieving consensus depends on the range of positions among which
agreement is sought. As this range narrows, the likelihood of agree-
ment increases. But at the same time, concern intensifies that this nar-
rowing requires arbitrary and exclusionary restrictions on the set of
relevant alternatives. Such restrictions would of course diminish the
interest of the agreement.

Let us bring this observation a little closer to the ground: Rawls
tells us that an adequate conception of justice must be able to win the
support of "reasonable citizens who affirm reasonable comprehensive
doctrines." Other views likely exist and ought not to be suppressed:
"That there are doctrines that reject one or more democratic freedoms
is itself a permanent fact of life..." (p. 64 n.19). But the fact that
certain doctrines do not accept the political conception of justice as
the correct account — the fact that they do not compose part of the
overlapping consensus — raises no troubles, Rawls claims, for the jus-
tification of the political conception. If a political conception is re-
jected by unreasonable comprehensive views, the legitimacy of the
exercise of power through institutions justified by that conception is
not undermined. Reasonable comprehensive doctrines "are the doc-
trines that reasonable citizens affirm and that political liberalism must
address" (p. 36).

The difficulty should now be clear: although confining the range of
relevant conceptions to reasonable views increases the likelihood of

132. On institutionalizing deliberation, see Cohen, supra note 130, at 26-32.
133. P. 36; see Cohen, supra note 63, at 281-85.
agreement, it also prompts concern that the label *unreasonable* will be used to exclude views arbitrarily — simply to ensure agreement or to silence dissent. We may state the objection as follows: If *unreasonable* simply amounts in the end to an abstract abbreviation for "disagrees with the dominant political conception of justice," then of course all reasonable views will support the political conception. But then the idea that an adequate conception must win the support of reasonable citizens who affirm reasonable doctrines will be of uncertain interest. If, however, *reasonable* is defined independently from acceptance of the political conception — say, in terms of a willingness to entertain and respond to objections — then reasonable citizens will likely affirm reasonable views that reject the political conception.

To respond, I should first note that even if acceptance of a particular political conception of justice in part constituted "reasonableness," the idea of an overlapping consensus would still be of interest. Given the fact of reasonable pluralism, a political conception that could be supported on the basis of premises provided by a variety of conflicting comprehensive moral conceptions would still be desirable. Because such conceptions would be reasonable in part because of their support for the political conception, we could not construe support from competing reasonable conceptions as providing an entirely independent check on the acceptability of the conception of justice. Still, this constitutive interpretation of *reasonable* would permit us to make a case for the thesis that consensus on a political conception of justice is compatible with moral pluralism — that it does not require agreement on a comprehensive conception of the good.

Although *reasonable person* is a normative notion, the constitutive interpretation of *reasonable* is not right. Instead, persons count as reasonable only if they are concerned to live on terms that are acceptable to others who share that same concern (pp. 48-54). In addition, they must acknowledge the "burdens of judgment": the conditions that cause disagreement among persons who affirm the importance of cooperating on terms that others can accept — that is, among persons who are reasonable in the first sense (pp. 54-56). Thus, reasonableness is defined abstractly and not — as with constitutive interpretation — in terms of the acceptance of a particular political conception. It more or less directly follows from these two features of reasonableness, however, that reasonable citizens will endorse certain basic liberties (pp. 58-61): how else could they show that they wish to live according to principles that they can justify to others, given disagreements with others that reflect the burdens of judgment?

But doesn't this characterization of *reasonable* show that the restriction of the overlapping consensus to reasonable views endorsed by reasonable citizens is arbitrarily exclusionary? Perhaps the arbitrariness is not as transparent as the constitutive interpretation suggests. Still, the restriction may seem to provide license to define away dis-
senting views as unreasonable and to exclude them from public discussion, while celebrating public consensus among the reasonable. Three points suggest otherwise.

First, we need to distinguish between tolerating a view and ensuring that it forms part of the overlapping consensus. It is no crime to be unreasonable — to favor institutions and policies that cannot be justified to others — or to express an unreasonable view, nor does the endorsement of such a view have any bearing on basic rights. The basis for such rights as expression and association is independent of the content of one’s views. Insofar as unreasonable views are “excluded,” then, that exclusion is of a special kind.

Second, it is a mistake to suppose that, as a general matter, dissenting views turn out unreasonable according to the account provided earlier. Consider, for example, dissident movements on the left in the recent history of this country. Why would anyone think that anti-intervention movements, or movements for civil rights, racial equality, women’s equality, economic justice, and gay and lesbian rights, are or were unreasonable? All these movements appeal, as a general matter, to political values in the democratic tradition. They struggle against the injustice of circumstances in which life chances are fixed by race, class, gender, or sexual orientation. Critics of these movements may disagree with the ways they have articulated democratic values, but we expect reasonable people to disagree.

As an example of a view that is at least in part unreasonable, Rawls mentions — plausibly, I think — the position that would deny to a woman “a duly qualified right to decide whether or not to end her pregnancy during the first trimester.”

134. Rawls says: “That there are views that reject one or more of the democratic freedoms is itself a permanent fact of life, or seems so. This gives us the practical task of containing them — like war and disease — so that they do not overturn political justice.” P. 64 n.19. This remark does not imply that we may do whatever we judge appropriate for containing objectionable views, any more than we can fight a disease by simply quarantining people who are sick. On tolerating the intolerant, see Rawls, supra note 3, at 216-21; on the right of subversive advocacy, see pp. 340-56.

135. Consider, to take just one example, proposals to regulate pornography in order to ensure sexual equality. See, e.g., Catharine A. MacKinnon, Only Words (1993) (reviewed in this issue — Ed.). These proposals appeal to political values. They do not reject the value of liberty generally, or freedom of expression in particular. Instead, they offer a particular way to combine freedom of expression and equality. Although I do not agree with these proposals, it is simply wrong to argue that they reject the value of freedom of expression or that the arguments for them rely on a particular comprehensive view. See Joshua Cohen, Freedom of Expression, 22 Phil. & Pub. Aff. 207 (1993); Joshua Cohen, Pornography: Left (Apr. 1994) (unpublished manuscript, on file with author).

136. P. 243 n.32. This right is much weaker than the right upheld in Roe v. Wade, 410 U.S. 113 (1973), which is not confined to the first trimester.
reproduction of political society over time . . . and finally the equality of women as equal citizens” (p. 243 n.32). Second, he claims that any “reasonable balance” of these values will support the “duly qualified right” (p. 243 n.32). To deny the right is either to deny, at the first step, that the equality of women is an important political value, or to claim, at the second step, that one of the other values — say, the due respect for human life — overrides the value of the equality of women, even if we confine our attention to the early stages of pregnancy.

Assume that the case for denying the right accepts the equality of women and is based on the value of due respect for human life. What prevents someone who accepts the three values from rejecting the duly qualified right as inconsistent with the due respect for human life? The problem is that people reasonably disagree about the precise content of the value of “due respect for human life.” Given the complexities of the question of the status of the fetus, the conscientious rejection by many citizens of the claim that due respect for human life requires that we treat the fetus as a human person in the first trimester, the weight of the equality of women as a political value, and the importance of justification to others when such weighty values are at stake, how could it be reasonable to urge the state to endorse and to enforce the view that due respect for human life bars first-trimester abortions? Someone who rejects first-trimester abortions may reply that when it comes to preventing the murder of innocent babies, being right is more important than being reasonable. But that reply concedes the point about reasonableness, which is the only issue I am now addressing.

Coming now to the third point about the exclusionary character of the notion of reasonableness: it is not arbitrary to worry only about ensuring support from the reasonable conceptions endorsed by reasonable citizens and therefore to exclude unreasonable views from an overlapping consensus. Such views do not aim to find terms that can be justified to others, and to that extent they deny the values of self-government and cooperation on terms of mutual respect. Moreover, one of the reasons for seeking common ground among conflicting views in the first place — for rejecting the appeal to the truth of our own view — is that we regard it as unreasonable to impose political power on others in the name of values that they reasonably reject — even if those values are correct. So the rationale for an overlapping consensus commits us to regarding views unconcerned with common ground as unreasonable. To permit those views to shape the content of a conception of justice is to permit the content of justice to be determined by the power of those views to make themselves heard. But no attractive conception can be built around such an accommodation to power.
D. Overly Accommodating

The final line of criticism I wish to explore accepts the ideal of political consensus but urges that an overlapping consensus is too limited. There are several variants of this concern, but I will focus here on one that takes Rawls's idea of public reason as its immediate target (pp. 212-54).

According to the idea of public reason, we should set aside comprehensive conceptions of the good in certain political settings—when discussing constitutional essentials and matters of basic justice (pp. 227-30)—and conduct political argument on the shared ground provided by political values. The criticism I have in mind rejects these limits of public reason because the constraints they impose on political deliberation prevent us from achieving a deeper level of political agreement than the idea of an overlapping consensus promises.137

To be sure, Rawls describes several exceptions to the requirement of respecting the limits of public reason—several cases in which it is permissible to appeal to a wider range of moral values than those within a political conception of justice (pp. 247-52). But limits remain. None of the exceptions mentioned in Liberalism—and none added in a recent essay modifying Liberalism's account of public reason138—would permit citizens, in the normal course of political argument, to bring the comprehensive views of others to the surface for the purpose of criticizing those views and the political implications that flow from them. Nor does the Rawlsian view encourage or require citizens to express their comprehensive conceptions in the course of political debate with a view to opening those conceptions up to the challenge of public discourse. The account of public reason may seem, then, to undervalue the importance of forms of critical discourse that do not respect the distinction between moral and political argument and as a result to truncate politics and practical reason. This tendency might seem objectionable for two reasons.

137. "[Liberalism] forgets the possibility that when politics goes well, we can know a good in common that we cannot know alone." Sandel, supra note 54, at 183. See also the illuminating remarks by Seyla Benhabib on the limits of liberal and discursive models of the public space in Seyla Benhabib, Models of Public Space: Hannah Arendt, the Liberal Tradition and Jürgen Habermas, in Situating the Self 89 (1992). Benhabib explores feminist criticisms of "overly rigid boundaries . . . between matters of justice and those of the good life, public interests versus private needs, privately held values and publicly shared norms." Id. at 111. In the end, however, I am not sure how far her own view differs from Rawls's. Here I will note just one reason. Benhabib uses the term political discourse in a very expansive way. See id. at 104. So her concern to open up public, political discourse to more comprehensive views—both matters of justice and those of the good life—reflects her idea that such discourse "can be realized in the social and cultural spheres as well." Id. Political discourse covers debates in "cultural journals" about sexual and racial stereotyping, for example. Id. As I explain in the text, see infra text accompanying note 140, Rawls uses the terms political and public more narrowly. So he agrees that the limits of public reason do not apply to political discourse, understood in such a capacious way. See pp. 214-15.

First, actual conceptions of the good may reflect traditions of injustice. A consensus that assumes such conceptions without challenging them — putting them behind a veil of ignorance, at the basis of an overlapping consensus, or off the political agenda — is for that reason less compelling as an account of ideal justice. According to the objection, if we wish to link justice and consensus, we need a consensus that emerges from unconstrained discussion, in which we may call on people to articulate their comprehensive conception of the good, which others may then challenge.

Second, constricting the arena of public discussion — limiting its scope to what can now be shared — perhaps excludes constructive possibilities of consensus and community that might emerge from challenging received moral traditions. Opening up the public arena by dropping the limits of public reason allows deeper challenges to existing conceptions of the good, thus permitting a more expansive consensus to emerge, if only as an ideal of reason.

To clarify the point of the objection, it may help to distinguish two conceptions of the aim of critical discourse. On one view, the point is to expose unreflective assumptions, thereby freeing ourselves from illusions and a false sense of coherence and necessity. This first understanding neither expects nor hopes that such a critique will generate a new and deeper consensus in which all previous views are understood as partial versions of the truth. According to an alternative conception, critique serves as an instrument of reasonable consensus. Instead of taking differences as fundamental and given, it invites a more searching public debate about hidden interests, suppressed alternatives, and moral disagreements with an eye to transcending current conflicts.

Here I am concerned only with the second line of thought: with the rejection of the limits of public reason in the name of possibilities of more comprehensive agreement, and a corresponding rejection of overlapping consensus for its relaxed accommodation of de facto conceptions of the good. There are two responses.

First, as a matter of clarification: to affirm the limits of public reason is not to deny the importance of a more comprehensive critical discourse, in which conceptions of the good — even if reasonable — are subject to challenge, unmasking, irony, and ridicule. Protection of freedom of expression always permits such discourse, and in some settings — even political settings — it may be entirely appropriate as a way to clarify views, to change minds, and perhaps to establish deeper mutual understanding.

The question is whether comprehensive critical discourse is appro-

---

139. In the legal academy, Duncan Kennedy is the great exponent of this first form of critique. See, e.g., DUNCAN KENNEDY, SEXY DRESSING ETC. (1993).
appropriate in deliberative settings that are concerned with establishing the basic terms of political cooperation in a democratic society and sanctioning the exercise of power to enforce those terms. The idea of the limits of public reason is that “political values alone are to settle such fundamental questions as: who has the right to vote, or what religions are to be tolerated, or who is to be assured fair equality of opportunity, or to hold property” (p. 214). Whatever the benefits of more comprehensive critical discourse in such settings, there is likely to be a cost. Critical discourse is likely to impede cooperation on terms of mutual respect, particularly when the views at issue are acknowledged — as I am supposing they are — to be both fundamental and reasonable. But “[m]any if not most political questions do not concern those fundamental matters” (p. 214). Accordingly, the case for limits on argument in the conduct of debate about issues such as trade policy is correspondingly weaker.

Second, given reasonable disagreements, the basis for expecting that a more comprehensive critical discourse will lead to a deeper consensus is unclear, which implies that the benefits are also unclear. It appears that “difference” is a fundamental fact, as fundamental as our commonalities. People disagree deeply, and political reason appears insufficient to resolve these differences. Putting aside comprehensive metaphysical theories according to which we all are the manifestations of spirit, or religious views accessible through faith, what reason could there be for denying that there are such rationally irresolvable disagreements? Everything points to the permanence of moral disagreement, and nothing points against it: there is the fact of disagreement and the absence of any apparent tendency to comprehensive convergence; we have no theory of the operations of practical reason that would lead us to expect convergence on comprehensive moralities; and there is no mechanism of the kind I sketched earlier in the case of political values that might produce agreement on comprehensive views.

One might argue that differences are not so deep because adherents of comprehensive moral conceptions believe their conceptions to be true and think they can withstand rational criticism. This observation suggests a fundamental common interest — in the truth, in living according to the best conception, or in living according to a view that can stand up to rational criticism — that lies deeper than any of our substantive disagreements about which conception is in fact true or best and therefore ought to guide conduct.

The availability of such abstract characterizations of common interests that underlie moral disagreements is of considerable impor-

140. The issue is not whether critical discourse ought to be legally permissible. The legal right must be established because of the requirement of equal basic liberties. See p. 337.
141. See supra notes 120-26 and accompanying text.
tance and may help to secure mutual understanding and respect. It may be important for me to view people who believe that the best life is a life that comports with God's prescriptions as having the same abstract, fundamental interest as I do — an interest in knowing what is true and in living the best life — even if I cannot imagine myself believing what they believe or conducting myself as they do. We all know how complex evaluative questions are, and we can understand how people conscientiously aiming at the same target might end up in very different places.

Finding deep commonalities of interest within moral differences is, then, a significant value. Nevertheless, the availability of such common ground gives us no reason for expecting a more substantial convergence on comprehensive moralities. The interests are too abstract to provide a basis for such an expectation. People with conflicting religious convictions might acknowledge one another as sharing an abstract common interest in believing the truth and in conforming their conduct to their understanding of the truth. This point of agreement might, in turn, be important in ensuring mutual respect among people with conflicting religious convictions. It provides minimal leverage, however, in resolving religious disagreement, and thus very little reason for expecting people's religious convictions to converge. Why should comprehensive moralities be any different?

V. DEMOCRATIC TOLERATION AND LIBERAL UNIVERSALISM

Early in this review, I described Political Liberalism as a deep and original book. I want to conclude by returning to the sources of that depth and originality, indicating their continuity with Theory.

There is of course no originality in the thought that people with different views of life can live together in a political society, and there is some evidence — relatively little, unfortunately — that toleration is a practical possibility. But the defense of toleration, when it does not appeal principally to the very great practical advantages of toleration, commonly proceeds in an "exclusivist" way. What I mean is that the defense of the claim that a political society ought to permit different outlooks on life to flourish within it commonly proceeds from the perspective of one of those outlooks.142

John Locke's defense of religious toleration, for example, seems to depend for its force on a Protestant view of salvation.143 Or consider John Stuart Mill's endorsement of individuality in On Liberty, his

---

142. There are some exceptions. See, e.g., BRUCE A. ACKERMAN, SOCIAL JUSTICE IN THE LIBERAL STATE (1980). Ackerman emphasizes the independence of political argument from moral argument and also the many routes to liberal political arguments. See id. at 355-59. But his discussion of "four of the main highways to the liberal state" suggests that his liberalism is a partially comprehensive doctrine. See id. at 359-69. I am indebted to John Rawls for a discussion of this issue.

143. See LOCKE, supra note 92, at 17-20.
powerful defense of a society featuring “different experiments of liv-
ing.”\textsuperscript{144} In the course of that defense, Mill urges that “[i]t may be
better to be a John Knox than of an Alcibiades, but it is better to be a
Pericles than either.”\textsuperscript{145} Presumably Mill thought that at least some
experiments in living would proceed more in the tradition of Alcibiades
and Knox than of Pericles. The non-Periclean experiments should
certainly be tolerated; Mill summarizes their toleration in his “harm
principle.”\textsuperscript{146} His reasons for such toleration, however, reflect the
Periclean perspective; they draw on a conception of human excellence
with roots in the “Greek ideal of self-development.”\textsuperscript{147} Mill reveals
the depth of these roots when he urges that “developed human beings
are of some use to the undeveloped” and that “those who do not desire
liberty, and would not avail themselves of it” may nevertheless be won
to the cause of liberty because they might “in some intelligible manner
[be] rewarded for allowing other people to make use of it without
hindrance.”\textsuperscript{148}

Rawls proposes something different, which I referred to earlier as
“democratic toleration.” By requiring toleration as a condition for ac-
ceptable public justification, he aims to free the defense of diverse ex-
periments of living from the outlook of one such experiment. More
broadly speaking, Rawls wishes to free the democratic ideal of a
shared arena of public deliberation among equal citizens from depen-
dence on the particular ethical outlook of any subset of the public.
Whether he succeeds in this enterprise is another matter, though I find
the case compelling for reasons I have already presented. The point I
wish to stress here is that in advancing a democratic conception of
tolerations, Rawls presents a sustained response to an important line of
criticism of classical liberal ideas of citizen, person, reason, and public.
According to the criticism, the superficial and abstract universalism of
these ideas masks a much deeper parochialism. Rawls’s conception of
an overlapping consensus on a political conception of justice suggests a
way to present those ideals as genuinely shared ground.

To be sure, liberal political thought has always been self-con-
sciously universalistic, speaking in the name of all human beings, and
urging the protection of the rights and interests of all, regardless of
race, class, sex, religion, or any other of the particularisms that distin-
guish and divide us. But critics of liberalism have vigilantly revealed
the hidden (and not-so-hidden) exclusions — of, for example, class,

\begin{itemize}
\item \textsuperscript{144} MILL, ON LIBERTY, supra note 14, at 115.
\item \textsuperscript{145} Id. at 120.
\item \textsuperscript{146} Id. at 72-73, 114, 132, 149-50.
\item \textsuperscript{147} Id. at 120. For interesting suggestions about the connections of this feature of Mill’s
view with his affection for colonialism, see Bhikhu Parekh, Superior People: The Narrowness of
\item \textsuperscript{148} MILL, ON LIBERTY, supra note 14, at 122. For example, the undeveloped “might possi-
bly learn something” from the developed. Id.
\end{itemize}
More Democratic Liberalism

race, and gender — that compromise liberalism’s defining promise: its capacity to say “all” without quite meaning it. Some critics have argued that its promise is essentially compromised. For them, liberal universalism is unavoidably exclusive; its fundamental categories, such as citizen, person, public, and rights, cannot be extended to include all people without losing their definition. These critics argue, for example, that the idea of a public sphere takes shape from its opposition to a private sphere and that the distinction between public and private stands in the way of the equality of women; or that the abstractions that define liberal universalism require that we neglect the more concrete differences — such as class and natural endowment — that shape actual lives.

Liberals, of course, deny that the project of liberal universalism is hopelessly compromised and that abstraction is the enemy of equality and inclusion. But denial is one thing; it is quite another to make a constructive case that liberalism can deliver more fully on the universalistic promise of its classical proponents and to abandon key elements of liberalism to ensure that delivery.

Consider in this light Rawls’s project in *Theory and Liberalism*. *Theory* took seriously the egalitarian critique of liberalism: the charge that the defense of liberty is a defense of the privileges of people with the wealth or status needed to make effective use of their liberty. In response, Rawls moved the idea of the social contract to a higher order of abstraction, presenting it as an agreement among free and equal persons, not among property owners, or among men, or among individuals with definite conceptions of their own advantage. Through this abstract reinterpretation of the social contract, Rawls made a compelling case for the view that the best version of liberalism is more egalitarian and inclusive than had traditionally been thought. In short, Rawls gave us a more genuinely universalistic liberalism, committed to “democratic equality” and less susceptible to charges of...


152. On the Lockean contract as an agreement among property owners, see Cohen, supra note 149.

153. Democratic equality is Rawls’s term for the conception of fair distribution that includes the difference principle. See Rawls, supra note 3, at 75-83; Cohen, supra note 38, at 727-31.
class exclusion.

Although *Liberalism* is not so concerned with the class question, it, too, aims at a more genuinely universalistic liberalism. Generalizing and deepening the ideal of toleration — by carrying it to a higher order of abstraction — Rawls offers a democratic liberalism less susceptible to charges of moral parochialism, sectarianism, and elitism and more suited to "the historical and social circumstances of a democratic society" (p. 154). By "apply[ing] the principles of toleration to philosophy itself," political liberalism leaves it "to citizens themselves to settle the questions of religion, philosophy, and morals in accordance with views they freely affirm" (p. 154).

Here we come to the heart of Rawls's work and the basis of his permanent contribution to political philosophy: he offers us a new version of democratic liberalism, marked by a commitment to liberalism's universalistic promise and a willingness to pursue that commitment by transforming those aspects of liberal thought that are condemned by its own high aspirations.

Consider the common ground of *Theory* and *Liberalism* from a different angle. In his Gettysburg Address, Lincoln said that the United States was "conceived in Liberty, and dedicated to the proposition that all men are created equal," and he wondered whether a political society with such abstract devotions could "long endure." ¹⁵⁴ Perhaps such a society would be unable to make good on the promise of liberty and equality; perhaps dedication to an idea and a proposition would provide too thin a basis for stable social unity.

*Theory* and *Liberalism* are the product of a life's engagement with these concerns. *Theory* gives us an account of what the promise of liberty and equality demands and a measure of how far we are from keeping that promise. *Liberalism* offers hope and a warning: the hope that we can achieve social unity in a democracy through shared commitment to abstract principles, and the warning that any political bonds thicker than these¹⁵⁵ would, by excluding some citizens, represent yet another failure to endure.


¹⁵⁵ For an example of thicker bonds, see the quotation from Pat Buchanan that begins this article. See supra text accompany note 1.
Stephen Holmes has recently published an engaging and stimulating, though finally unsatisfying, book. At a time when modern liberalism is being assailed seemingly from all sides — by fundamentalist Christians, conservative libertarians, critical race and feminist legal scholars, and communitarian political scholars — Holmes endeavors in The Anatomy of Antiliberalism to defend the faith from attack by a discrete and somewhat nonobvious group of theorists. The book purports to weave the works of thinkers as diverse as Joseph de Maistre and Roberto Unger into a coherent tradition of “antiliberalism” and, in so doing, to correct the oft-repeated errors of both historiography and interpretation that run through this tradition.

That he is only partly successful in these aims reflects more on his taxonomic choices than his substantive analysis. The book is at the same time over- and underinclusive. First, it is not at all clear that Holmes has, in fact, identified a tradition of antiliberalism that is more substantial than the extremely broad definition of antiliberal he constructs to encompass the variety of views he highlights; hence the theory is overinclusive. Second, it is underinclusive in that even if the protofascism of Carl Schmitt and the communitarianism of Michael Sandel can be considered part of a unitary tradition — without stripping such a tradition of normative weight — Holmes has neglected to address adequately the salience that the communitarian critique, especially in legal contexts, has for liberalism.

Because he is writing about an opposition theory, Holmes begins with a thumbnail sketch of the liberal tradition. Somewhat disappointingly, he defines liberalism as “a political theory and program that flourished from the middle of the seventeenth to the middle of the nineteenth century” and “continues to be a living tradition today” (p. 3). More specifically, he identifies the core practices of liberalism as religious toleration, freedom of expression, constraint on state action against the individual, broad franchise, constitutional government, and commitment to private property and freedom of contract (pp. 3-4). Four broad core norms in turn support these practices: personal security, impartiality, individual liberty, and democracy (p. 4).

Holmes then offers an overview of the tenets of non-Marxist an-
tiliberalism.³ Warning his readers that antiliberalism is "always a sensibility as well as an argument" (p. 5), Holmes sets forth the common attitudes he identifies in the antiliberal mind. Antiliberals decry atomization and the alienation implicit in rational self-interest (p. 6). They distrust science and technology and the Enlightenment usurpation of religious morality by secular humanism (p. 6). They are hostile to the culture of modernity and tend to conflate the theory of liberalism with the practice of liberal states (p. 6). Moreover, they are apocalyptic: society, at whatever time they write, is in a "crisis" that it can overcome only by eradicating the "virus" of liberalism (p. 7). Additionally, one might add that antiliberals systematically decontextualize liberal theory, thus positing as descriptive claims what are clearly normative aspirations.⁴

Holmes structures the book simply. Part I analyzes a series of representative antiliberal thinkers. Part II refutes the standard historical account of liberalism offered by these theorists and attempts thereby to deepen our understanding of the liberal tradition. Holmes makes an initial distinction in Part I between "hardline" and "softline" antiliberals (p. 88). He devotes the first portion of Part I to the former: Joseph de Maistre, Carl Schmitt, and Leo Strauss. The remaining bulk of Part I addresses the latter: Alasdair MacIntyre, Christopher Lasch, Roberto Unger, and the communitarianism associated with Charles Taylor and Robert Bellah. Apart from the substantive distinction between these two groups, the differentiation tracks historical chronology in an acknowledged way: hardline antiliberals all antedate the Nazi regime.⁵ Holmes employs similar methods in dissecting each of his antiliberals: he points out their internal inconsistencies, their reliance on empirically untrue factual assertions, and their misreading of liberal theory. The method not only serves to keep the reader's attention focused on the similarities among the samples, but it reinforces the narrative structure Holmes imposes on his argument.

He begins his story with Joseph de Maistre, the counterrevolutionary Catholic philosopher of the late eighteenth century. Maistre, as Holmes portrays him, held a foundational view of man's essential

³. Pp 1-2. Holmes is concerned here solely with non-Marxist theorists, noting that the antiliberals he surveys would likely assert that Marxism and liberalism are offshoots of the same tradition.

⁴. Don Herzog devastatingly critiques this methodology — especially as employed by critical legal studies scholars — in Don Herzog, As Many as Six Impossible Things Before Breakfast, 75 CAL. L. REV. 609 (1987).

⁵. Holmes notes in the introduction that after 1945 the rhetoric of communitarianism was "radically demilitarized." P. 9. Moreover, he dryly detects a lack of martial seriousness in his postwar antiliberals:

When MacIntyre or Unger suggest en passant that killing enemies or risking one's life in the carnage of battle provides a solution to the spiritual emptiness of commercial society, readers cringe but then rightly dismiss the literal implications of what they say. A prewar antiliberal, such as Schmitt, was obviously in greater earnest.

P. 10.
bloodthirstiness.6 Humans gathered together in society necessarily require an authority that is both final and absolute to prevent them from butchering each other. Maistre finds such authority in both a temporal and a spiritual incarnation: the monarch and the Church. Obedience to the absolute commands of these twin sovereigns provides the cement of community without which humans will descend ineluctably into depravity. Hence Maistre vociferously attacked the Enlightenment for replacing obedience with discussion, and he blamed the Reformation in particular for encouraging revolt against spiritual authority (p. 18). The direct result of this “horrifying project of extinguishing both Christianity and sovereignty in Europe” (p. 18) is the Terror (p. 15).

Maistre provides Holmes an excellent opportunity to introduce the broad strokes of his “anatomy”: the necessity of unassailable political authority, or decisionism (pp. 18-19); the privileging of spiritual over temporal community — of sacralized over secular institutions — and the concomitant belief in the false necessity of the existing order (pp. 21-23); the characterization of the scientific method as degrading to morality (p. 23); the denigration of the individual, as compared to the group (pp. 25-27); and finally the simultaneous attitude of revulsion toward and exaltation of violence and bloodshed.7

From Maistre, Holmes neatly segues into a discussion of Carl Schmitt. Holmes points out that both Maistre and Schmitt experienced firsthand a “world-shaking crisis of authority” (p. 37). Schmitt’s revolution was the defeat of Germany in the First World War and the subsequent political instability of the Weimar Republic. Moreover, Schmitt’s earliest major work, Political Theology,8 contains an admiring chapter on the “decisionism” of Maistre and other “counterrevolutionary philosophers of the state.”9 The context of Schmitt’s work is crucial, of course. Holmes is fairly balanced in his portrayal, neither downplaying the particularly egregious aspects of Schmitt’s involvement with the Nazi Party, nor dismissing him as a mere Nazi theoretician.10

7. Pp. 27-32. Maistre’s response to the Terror both as a horrific result of the abstract humanism of the Enlightenment and as part of a just God’s plan for religious revival nicely captures the tension in this attitude. P. 15.
9. Id. at 53-66.
10. As an example, Holmes notes that Schmitt hosted a conference on “German jurisprudence at war with the Jewish spirit” in 1936 to counter SS accusations of insincere antisemitism. Pp. 38-39. Conversely, Holmes refers to commentators who are sympathetic to Schmitt’s post-war denials of culpability in the Nazi catastrophe — including his biographer Joseph Bendersky — as “apologists.” P. 42. Schmitt’s role in the Third Reich may well have been less that of a true believer than that of a pragmatist, conforming his behavior to the dictates of the regime in order to advance his own ideas. See, e.g., George Schwab, Introduction to SCHMITT, supra note
Holmes rightly separates Schmitt’s ideology from his actions and identifies the core principles of his thinking. The first principle is his famous enemy-friend distinction: that personal enmity is often crucial to politics (p. 40). The power to define this enmity is the exclusive province of the sovereign, and Holmes perceptively notes that Schmitt’s hostility to communism was a product of communism’s universalist posture: by positing a universal “class enemy,” communism shifted the people’s focus from the international to the domestic, thus weakening nationalism and encouraging internal chaos (pp. 41-43). The essential antidote to the militant pluralism that consumed Weimar Germany is for Schmitt the “decisive leader,” or dictator. In contrast to the endless discussion that characterizes parliamentary democracy, Schmitt sees in dictatorship both the practical requirement of decisiveness and the theoretical legitimacy of democratic acclaim. As Holmes notes, a commitment to democracy — defined as the direct expression of popular will, unencumbered by dissent, free speech, or opposition parties — is, at best, “perverse” (p. 49). But political legitimacy for Schmitt is a function of the crowd, of the almost mystical identification — captured so hauntingly by Leni Riefenstahl — between the ruler and the ruled.

Holmes rightly admires much of Schmitt’s thought; it is the admiration one has for a beautiful monster, perhaps. He also devastatingly refutes both the descriptive aspects of Schmitt’s critique and his misreading of liberal theory. As Holmes points out, liberal societies are very capable of binding decisionmaking, are aware of and able to accommodate the demands of heterogeneity, and are quite adept at mustering defensive force and effectively governing far-flung empires (p. 58). Moreover, he notes, the rule of law governing the constitutional state as envisioned by John Locke is not the “sovereignty of abstract, self-applying rules” (p. 59) that Schmitt sees as hobbling the executive power, but rather a mechanism for maintaining the personal accountability of the executive when making those “hard” political choices Schmitt so admires (p. 59).

The last hardline antiliberal Holmes examines is Leo Strauss, and, again, Holmes provides an elegant segue. Strauss first gained widespread notice in 1932 with his review of Schmitt’s Concept of the Political. Strauss also provides a kind of bridge between the hard- and

8, at xi, xiii (suggesting that Schmitt made “shocking” compromises with the Nazi regime in order to supplant Nazi “totalitarianism” with his own “authoritarianism”).


13. See, e.g., p. 57 (“Schmitt’s criticisms of liberalism are often interesting and sometimes persuasive.”).

14. P. 60; see Leo Strauss, Comments on Der Begriff des Politischen by Carl Schmitt, re-
softline schools of antiliberalism. If it is difficult to see the connection between the devotion to the Great Books of this cosmopolitan German-Jewish emigre and the hallucinatory fever of Maistre, it is easier to measure the impact of his critique of the modern West on the current cultural debate. The phenomenal success of Allan Bloom's contemporary warmed-over Straussianism should lay to rest any doubts as to his continuing influence.

Holmes focuses on Strauss's reading of ancient texts as teaching that "inequality is central to the human condition" (p. 70). From this conclusion Strauss constructed his own dualistic portrait of the world: there are "philosophers," those that understand this truth and live accordingly, and "the herd," who must be fed the pablum of religion to spare them the pain of knowing this truth (pp. 64-65). This dualism provides the basis, not only of Strauss's contempt for the liberal ideal, but also of his strategy for dealing with its ascendancy. Not only is liberalism's central tenet— the essential equality of humanity— self-evidently wrong, indeed unnatural (pp. 81-82), but the liberal project itself is ultimately destructive of all social order. The masses, stripped of both their illusions and their capacity for obedience, will demand the satisfaction of their base desires well beyond the ability of a well-ordered society to deliver (p. 64). Moreover, the elite, equally seduced by the siren song of equality, will tend not to the ascertainment of eternal verities in service of their role but instead to the domination of nature, through science, necessary to satisfy the appetite of the polis (p. 72).

Hence the Straussian strategy of esotericism. Because of the dangerousness of the philosopher's view of the world, Straussianism argues that philosophers must keep silent when addressing the herd. Far from advocating a remaking of the established order, Strauss seems oddly content with the world as it is — provided it has room for the lonely philosopher and his disciples. Perhaps it is his essential fatalism — certainly an ancient sensibility — that accounts for this. As Holmes points out, Strauss posits a realizable "good society" made up of sedated masses, gentlemen rulers, and philosophers directing the whole works, but at the same time Strauss argues that the creation of such a society depends on chance, rather than the willed intention of those who could truly appreciate such a society (p. 74).

As Holmes shows, the hardline antiliberals of his survey are just

---


16. Holmes deliciously skewers the pretensions of academics by suggesting that Strauss's view of philosophers as "walking time-bombs" may explain much of his appeal. P. 78. As Holmes notes, Strauss's insistence that intellectuals must obscure their understanding of the world lest they destroy the illusions that sustain ordinary men, is "[for] desk-bound scholars... an extraordinarily flattering idea." P. 78.
that; their attacks are not merely against the descriptive conditions of liberal societies but against their normative underpinnings as well. Like Maistre they refute the rule of law and privilege faith and obedience over reason; like Schmitt they disparage democracy by discussion and exalt the *fuhrerprinzip*; like Strauss, they deny any kind of political equality among persons, subscribing instead to a rigidly hierarchical social order. Holmes is at his polemic best in this part of the book. He notes the logical absurdity of Maistre's simultaneous claims that God's will is supreme and that the Enlightenment threatens to supplant God's will (p. 35). He conclusively shows that Schmitt's pose as a latter-day Hobbesian relies on a fundamental distortion of Hobbes's Erastianism to fit the peculiar contours of Nazi antisemitism (pp. 50-53). Finally, he exposes in Strauss the intellectual subterfuge of contrasting ancient philosophy with modern society, rightly chiding him for refusing to consider the conditions of ancient society as the correct comparison (p. 83). But once he turns his attention to the softliners, those who criticize liberalism but when pressed "reveal a surprising fondness for liberal protections and freedoms" (p. 88), the book, and Holmes's argument, begins to waver in focus.

The major difficulty centers on whether the inclusion of these softliners stretches the conceptual framework of antiliberalism beyond the point at which it ceases to do any interesting work. The truth is that the hard brand of antiliberalism simply does not pose a credible threat to much of anything. Although one could presumably argue that some events in Eastern Europe and the former Soviet Union signal a resurgence of Schmittian nationalism, those events are almost uniformly viewed as pathological by the liberal democracies; they pose no philosophical challenge to modern liberalism. So in order to raise the stakes high enough to be of interest, Holmes must demonstrate that salient internal critiques in modern states are merely the old antiliberal attacks dressed in fancy new clothes. The difference between a Carl Schmitt and a Christopher Lasch, however, may be more a difference in kind than degree.

This certainly seems to be the case with both Lasch and Alasdair MacIntyre. Softline antiliberalism seems most closely to conform to Holmes's idea of a sensibility rather than an argument. The sensibility here is primarily a generalized yearning for the prelapsarian past. For moralists like MacIntyre and Lasch — whom Holmes describes as communitarian conservatives (p. 141) — the failure of modern liberalism has more to do with its being modern than being liberal. For MacIntyre, the condition of modernity is the triumph of the secular over the sacred and a concomitant loss of moral authority. Holmes attempts to portray MacIntyre's response to this loss as "Schmittian" (p. 94). But unlike Maistre's privileging of religious faith as a means of enforcing obedience to the sovereign, MacIntyre seems to view religion as essential to maintaining a normative consensus (p. 97). This is
a subtle, but fundamental, difference, because it goes to the establishment of criteria, rather than of authority. MacIntyre is interested in discussion; his concern is that the radical secularization of modern life has eliminated any baselines by which to judge the merits of good arguments (p. 100). Moreover, he sees in the demise of religion the unraveling of the communitarian claims that bind people to the search for common good (p. 91). Indeed it is this yearning for community valuation that arguably links him to Schmitt and certainly aligns him with the cultural criticism of Christopher Lasch.

Lasch shares with MacIntyre a profound hostility to science and technology (pp. 97, 122-40). Lasch also envisions a bygone halcyon world — in his case a community of yeoman producers, free of the parasitic consumer class — but for Lasch science is merely an element of the lust for “progress” which he blames for the sorry state of modernity. Holmes nicely characterizes Lasch’s outlook as “anti-Promethean” (p. 128) and rightly takes him to task for failing to balance the benefits — like literacy and sanitation — with the costs of liberal progress — like drug addiction and impending environmental catastrophe (p. 137). What Holmes fails to achieve with his analysis of Lasch and MacIntyre, however, is a real connection between their grumblings about the modern world and the kind of sustained attack on liberal norms launched by Maistre and Schmitt. They seem, from Holmes’s description, less a serious threat to the liberal order than a couple of dyspeptic curmudgeons. Moreover, they do not seem antiliberal in any meaningful sense of the word.

This difficulty is even more pronounced in the chapter on Roberto Unger. Holmes describes Unger as a countercultural radical and distinguishes his brand of antiliberalism as that which assails liberalism, not for being anarchistic, but for not being anarchistic enough (p. 141). For Unger, the sin of liberalism is not that it has loosened the bonds of community but rather that it stamps out the “spontaneity of the soul” (p. 158). Hence Unger’s own self-description as a “superliberal” (p. 160). Admittedly, this is a recent stance for Unger, and Holmes capably narrates the labyrinthine progression in Unger’s work from communitarianism to individualism. He also quickly exposes the vulnerabilities of superliberalism as theory, especially the assumption that context smashing can substitute for moral doctrine (p. 169). It seems odd, however, to argue that the call “to redeem liberalism through more liberalism” (p. 160) is an antiliberal one. Moreover, one gets the sense in this chapter that Holmes is shooting ducks in a pond: if Unger quite simply is not a rigorous thinker, and if his grandiosity and confusion keep him firmly at the margins of academic debate, then why has Holmes devoted an entire chapter to his work?

17. The link, however, is tenuous. For Schmitt the community serves as the means of distinguishing friend from enemy; for MacIntyre, it serves to provide a moral vocabulary.
After a too-brief chapter on the communitarian "trap" (pp. 176-84), in which Holmes takes Michael Sandel, Charles Taylor, Robert Bellah, and others to task for attaching moral significance to terms like social that are inherently descriptive (p. 178) and for, again, systemically confusing criticisms of liberal society with criticism of liberal thought (p. 181), Holmes attempts in Part II to answer the standard descriptive claims about liberalism put forward by its critics.

To the charge that liberalism conceives of individuals as atomized, Holmes answers that antiliberals have misinterpreted Locke's requirement of consent to authority as a claim of presocial rationality and decontextualized social contract theory in general (pp. 193-94). Reliance on rational self-interest is a normative and not a descriptive claim: the point is not that people are necessarily the best judges of their needs but that there is no good reason to assume that their rulers necessarily are (p. 197). Likewise, liberal theory, grounded on consent of the governed, is hostile, not to authority as such, but merely to that authority that is arbitrary and capricious (p. 203). Finally, Holmes argues that liberal rights, because they carry correlative duties, are not inherently alienating (pp. 228-31) and that the elaborate procedural mechanisms that liberal states employ to channel naked preferences into political discourse reveal that liberalism has never been grounded in "moral skepticism" (p. 235).

These claims are all good, of course, but what about those community-oriented arguments? It is here that Holmes's book finally does not satisfy. In John Rawls's most recent book, he addresses the tensions between claims of individualism and claims of community by noting that "[a] well-ordered democratic society is neither a community, nor, more generally, an association."18 This rejection of community for Rawls follows from the observation that "the diversity of reasonable comprehensive religious, philosophical, and moral doctrines found in modern democratic societies is not a mere historical condition that may soon pass away; it is a permanent feature of the public culture of democracy."19

I do not mean to equate Holmes's project with Rawls's, but what must follow from Rawls's assertion is that liberalism must find the means to accommodate these competing doctrines without destroying itself and that the challenges posed by various strands of contemporary legal communitarianism cannot simply be brushed off by demonstrating that such claims are based on a misreading of liberal theory. Liberalism, after all, predates modern democracy. The argument that rights generally understood as private are better conceived of as public and deriving from the political community — implicit in advocacy for

18. JOHN RAWLS, POLITICAL LIBERALISM 40 (1993) (reviewed in this issue by Professor Joshua Cohen — Ed.).
19. Id. at 36.
hate-speech codes, for example — cannot be adequately answered by a narrow appeal to correlative duties (p. 228). Inappropriate preference formation in general poses particularly thorny dilemmas for liberalism. One attempt out of the thicket is the revival of republicanism, focusing on the transformative and deliberative nature of political community as a means of mediating the tensions between self-governance and governance by law — between the needs of the individual and the needs of the community.\(^{20}\) Is modern republicanism thus "antiliberal"? Perhaps even more salient is the literature on preference formation itself.\(^{21}\) Liberalism must develop models thick enough to account for the constitutive role of legal institutions, the presence of heuristic biases, and the impact of risk-assessment on the choices actors in a liberal society make. When those choices imply broad consequences for society as a whole, as they often do in the environmental arena, for example, it does not satisfy to note simply that Locke and Montesquieu were aware of the elementary processes of character formation.

In an earlier work, Holmes has written splendidly of the early period of modern liberalism,\(^{22}\) and *The Anatomy of Antiliberalism* provides the same kind of thoughtful and important analysis of the rise of antiliberal thought. His chapters on the hardline antiliberals sparkle with acerbic wit and cogent criticism. But Holmes’s attempt to extend his critique of these theorists to encompass contemporary softliners both robs the hard antiliberals of their rhetorical power and disserves the intentions of contemporary dissenters from the liberal orthodoxy.

— Jeffrey R. Costello

---

