Unpacking the Court*

Lewis A. Kornhauser† and Lawrence G. Sager††

I. INTRODUCTION

Traditional theories of adjudication are curiously incomplete. Their focus is on the process by which a single judge decides or ought to decide cases. Judging is treated as though it were a solitary act, occasioning reference to other judges only to the extent that the artifacts of their past efforts provide authoritative or instructive precedent for the present decisions of the lone judge. The real world of adjudication, though, differs dramatically. Only at the trial level do judges normally decide cases alone. Intermediate courts of appeal and courts of last resort are organized so that judges almost always sit and act together with colleagues on adjudicatory panels. And the size of these panels almost always varies directly with the level of authority they enjoy within their legal system, indicating our widely-shared, naive belief that this multiplicity of judges is in some sense a route to "better" decisions.[1] Appellate adjudication, the common, almost exclusive focus of theories of adjudication, is thus essentially a group process, yet extant theories neither explain the group nature of the process nor take it into account.[2]

1.   Some version of this belief inspired Congress earlier in this century to create three-judge federal district courts—now largely abandoned—to adjudicate cases of particular sensitivity, most notably those involving injunctive disruption of state laws. See 28 U.S.C. § 2281, Act of June 25, 1948, ch. 646, 62 Stat. 968, *repealed by* Act of Aug. 12, 1976, 90 Stat. 1119. This statute was originally enacted in a narrower form as Judicial Code § 266, Act of June 18, 1910, ch. 309, § 17, 36 Stat. 557 (three-judge court required for interlocutory injunction against enforcement of state law). The reach of three-judge panels was expanded by 28 U.S.C. § 380, Act of Feb. 13, 1925, ch. 229, § 1, 43 Stat. 938 (three-judge court required for permanent injunction against enforcement of state law) prior to its repeal.

2.   Considerations of practical politics dominate discussions of the group aspect of judicial decision-making. The Supreme Court is the typical object of such discussions, which are usually concerned with the pragmatic aspects of the melding of the views of the nine justices into patterns of results. Broader theories of the group dynamic of Supreme Court decisionmaking may invoke notions of political stability or representation. But such theories make no effort to tie these concerns to the basic jurisprudential question of how judges do and ought to decide the cases before them.

In this Article, we offer some observations about the attributes of multi-judge courts and their relationship to basic themes in the theory of adjudication. While this first venture falls short of a comprehensive theory, it involves considerably more than minor tinkering with conventional understandings of adjudication. Our effort to incorporate the fact of group decisionmaking into analysis of the judging process has led us to generate a fundamental distinction between "preference aggregation" and "judgment aggregation" in processes of group decisionmaking, to reconsider the traditional taxonomy of schools of jurisprudence in terms of this distinction, to reflect on the idea of representation, and to refine and distinguish the concepts of "consistency" and "coherence" as they apply to judicial decisionmaking.

From the midst of these analytical turnings, three propositions emerge with some force. First, given a reasonable understanding of what the job of judging is and under reasonable assumptions about how well individual judges are likely to do it, enlarging the number of judges who sit on a court can be expected to improve the court's performance. Second, multi-judge courts are quite capable of behaving consistently. If each judge on a court acts consistently from case to case, so too will the court that they constitute. Third, the same simple relationship between individual and group judicial performance does not hold for coherence. The coherence of a body of law generated by a court depends not only on the ability and commitment of each of its constituent judges to behave coherently, but also upon the nature and congruence of these judges' understandings of the criteria for coherence.

We of course hope to leave the reader persuaded that ours is a useful analytical lens through which to view the group aspect of the adjudicatory process, and to secure confirmation of our sightings through that lens. But, more importantly, we hope to convince the reader that the fact of group decisionmaking demands the attention of any serious and complete theory of adjudication.

II. THREE MODELS OF GROUP DECISIONMAKING

Groups can adopt a wide variety of mechanisms for generating decisions on questions of concern to them. The criteria of decision can vary enormously, from chance, to collective preference, to deific will, to moral, aesthetic, scientific, or legal judgment. The designated arbiter of decision can vary, from king or priest, to general electorate, to professional community, to legislature, court, or jury. So, too, the decisionmaking process of the designated arbiter can vary, from oracular ritual to a variety of deliberative and voting procedures. All of these mechanisms of choice involve decisions made on *behalf* of groups. Most of them also involve deci-

sions made *by* groups, in the sense that the actual decisionmaking process involves and responds to expressions of choice by more than a single individual. Group decisionmaking in this second, active sense is our concern here.

Our understanding and evaluation of any group decisionmaking process, like appellate adjudication, depends necessarily upon the purposes and capacities that we impute to that process. These can vary in significant, often ignored, respects. Thus, discussions of group decisionmaking shift uneasily and often unconsciously among at least three distinct models of group processes, which we will label preference aggregation, judgment aggregation, and representation.[3]

A. *Preference Aggregation and Judgment Aggregation*

The most significant analytic divide separates preference aggregation from judgment aggregation. It originates in the difference between expressing a preference and rendering a judgment.[4]

---

3. In Section III, *infra*, we discuss in detail the relationship between preference and judgment aggregation and traditional jurisprudential positions. Here, we offer the eager reader some sense of the connection between our analysis and traditional jurisprudence.

Most jurisprudential theories of adjudication consider adjudication, at least implicitly, to be the rendering of a judgment rather than the expression of a judicial preference. While Ronald Dworkin's theory may represent the purest example of a judgment-based theory of adjudication, (consider, for example, his argument in "Hard Cases" and "Can Rights be Controversial?," both in R. DWORKIN, TAKING RIGHTS SERIOUSLY 81-130, 279-90 (1977)), legal positivism also considers the making of judgments central to adjudication. H.L.A. Hart introduces his game of "scorer's discretion" precisely to emphasize that judicial decisions do not reflect the whim, will or preference of the judge, but rather the judge's considered *judgment* of what the law requires. In scorer's discretion, the game's scorer is not bound by fixed rules, but is free to specify the score at her discretion. In contrast, Hart analogizes judicial decisionmaking to a game in which the score is determined under a fixed rule. H.L.A. HART, THE CONCEPT OF LAW 138-44 (1961).

Some strands of realism and inquiries influenced by economics and political science often regard adjudication as an expression of judicial preference. Jerome Frank's account represents the most non-judgmental, realist theory of adjudication. J. FRANK, LAW AND THE MODERN MIND (1970). Some critical legal theorists may view themselves as heirs to this aspect of the realist tradition; one might for instance interpret the practice of "trashing" as an attempt to establish that decisions that are indeterminate as judgments rest on preference instead. For a discussion of critical legal studies and the indeterminacy thesis, *see* Gordon, *Critical Legal Histories*, 36 STAN. L. REV. 57 (1984); Kelman, *Trashing*, 36 STAN. L. REV. 293 (1984). Both Judge Posner and Judge Easterbrook regard judges as acting on preferences. *See* R. POSNER, ECONOMIC ANALYSIS OF LAW 506 (3d ed. 1986) ("One [possible explanation for judicial behavior] that is consistent with the normal assumptions of economic analysis is that judges seek to impose their personal preferences and values on society."); Easterbrook, *Ways of Criticizing the Court*, 95 HARV. L. REV. 802, 815 (1982) ("Each group of Justices . . . has a preference between the two positions taken by the others."). For more extended comments on Easterbrook, see *infra* notes 27, 35, 37.

The view of appellate courts as representative bodies may be most common among journalists and the lay public. This view, however, does emerge in the legal literature. *See, e.g.*, Hazard, *The Supreme Court as a Legislature*, 64 CORNELL L. REV. 1 (1978).

4. Sen introduces a distinction between the aggregation of individual *interests* and individual *judgments*. Sen, *Social Choice Theory: A Reexamination*, 45 ECONOMETRICA 53 (1977), *reprinted in* A. SEN, CHOICE, WELFARE AND MEASUREMENT 158 (1982). Sen apparently considers the difference between interests and judgments to lie in the nature of the information relied upon in making a

1. *Preferences and Judgments*

When an individual expresses a preference, she is advancing a limited and sovereign claim. The claim is limited in the sense that it speaks only to her own values and advantage. The claim is sovereign in the sense that she is the final and authoritative arbiter of her preferences. The limited and sovereign attributes of a preference combine to make it perfectly possible for two individuals to disagree strongly in their preferences without either of them being wrong. Nor does the presence of disagreement provide a reason for either person to change her preference. Thus, the claim "X prefers meat" neither conflicts with the claim "Y prefers fish" nor provides a reason for Y to alter or even reconsider her preference.[5]

In contrast, when an individual renders a judgment, she is advancing a claim that is neither limited nor sovereign. The claim is not limited in the sense that claims of preference are; a judgment advances a "truth," that is, a proposition to which all other right-thinking persons who may confront the issue must adhere. The claim is decidedly not sovereign in the sense that an individual's adherence to the judgment does not itself justify it. Two persons may disagree in their judgments, but when they do, each acknowledges that (at least) one of them is wrong. Moreover, the disagreement should give each pause: that others evaluate the circumstances differently may be a reason to revise or reconsider one's judgment. The claim "Meat is the healthier food" conflicts with the claim "Fish is the healthier food"; if X and Y have these incompatible convictions, then the position of each is a reason for the other to reconsider.[6]

At the core of the distinction between expressing a preference and rendering a judgment lies the proposition that some questions have "right" or "correct" answers. In this context, a "right" or "correct" answer need not be objectively true or depend upon some ultimate view of the real world; it

decision.

5. While disagreement is not a reason for one to change one's preference, it may provoke a change nonetheless. For instance, X might prefer meat to fish because her preferences for food derive from more basic preferences over taste, nutritional content, and appearance of the food. If X were misinformed about the nutritional content of meat and fish, disagreement might lead her to discover that the nutritional content and health effects of meat were sufficiently less desirable than those of fish so that her expressed preference would change. Her underlying preferences, however, would remain constant. Other cases are more complex. X might prefer agreement with Y to disagreement. The disagreement provides X with new information about the situation and again her expressed preference might change. The disagreement here is a cause of the change, but not a reason for it.

6. Knowledge that one person is wrong does not imply that either person will alter her judgment. Judgment lies within rational discourse, however, and seeks a correct answer. Knowledge that another party has reached a different answer gives one reason to review one's own judgment. This reason is most easily understood in a situation in which no dispute over the criteria of truth exists. If X "judges" that 121 x 11 = 1321, while Y "judges" that 121 x 11 = 1331, both have reason to check their calculations. Of course, Y might choose not to check her calculation because she knows X is a bad calculator or because she must make an immediate decision or because the calculation need only be accurate to within ten percent.

may depend only upon intersubjective agreement over criteria for resolving disputes.[7] When we render a judgment, we assert that the question of the moment has a right or correct answer in some significant and general sense and that our answer to the question embodies that rectitude. In contrast, when we express a preference, we assert only that we value or desire a particular thing or outcome over some alternative. In many social contexts, of course, the appropriate judgment as to a social course of action may depend upon the preferences of individuals who constitute the social entity, but this observation uses rather than attacks the distinction between judgments and preferences.

This distinction and its ramifications may seem too sharply drawn, and we want to anticipate some potential reservations. One objection that suggests itself is that preferences are simply a common subspecies of judgment, distinguished only by their personal subject matter. This claim is inspired by the possibility that a person can be wrong about her preferences. If we take preferences to be comprised solely of the immediate, accessible feelings of desire experienced by an individual, she can never be wrong about their content. But what of claims that, for example, neurosis, ideology, brainwashing, or a simple lack of careful reflection have distorted an individual's observation of her true sense of value? These claims, which are familiar and in appropriate contexts persuasive, hold that the "true" preferences of an individual may not be reflected in her superficial experience of transient wants.[8] But if we can be wrong about our preferences, in what sense are we the final arbiters of our preferences, and how, if at all, do expressions of our preferences differ from judgments about other truths in the world?

The claim that a person is in error about her preferences can only mean that there is something distorting them, and that once it is banished, her authentic sense of self will assert itself, allowing her to recognize the true objects of her desire.[9] Her authentic underlying self remains the ultimate authority as to her true preferences. She cannot be wrong about her preferences in the same way that she can err on the question of whether or not her ankle is broken. A preference is not a condition that happens to be lodged in a particular person, about which thinking persons can disagree.

---

7. Again, Hart's example of the game of scorer's discretion makes this point quite dramatically. *See supra* note 3. For purposes of this argument, then, we need not take a position on various complex issues of skepticism and the relativism of ethics and "truth" itself.

8. Much of the argument of Habermas and the Frankfurt School, for example, rests on the claim that people do not act on their true interests. For a lucid and illuminating discussion of true or real interests, see R. GEUSS, THE IDEA OF A CRITICAL THEORY 45-54 (1981).

9. Again, Geuss explicates these ideas more clearly than the school which has advanced them. *See id.*

In short, the claim that preferences are simply judgments about certain personal matters is not persuasive. A more or less parallel claim, however, pushes in just the opposite direction: that normative "judgments" are, at their roots, matters of preference.[10] On this view, at some point in all normative discourse reason ceases to drive the argument—assuming that it ever did—and recourse must be made to presuppositions of value. These root presuppositions, it could be argued, share the distinguishing quality that we have assigned to preferences: they are the product of the non-rational, sovereign choice of those persons who adopt them.

This view of normative discourse is decidedly odd. Suppose that X holds the view that every human life is intrinsically, deeply, and equally valuable, and Y holds the view that no one need value any life other than her own. It might be true that reason would fail either as a means of convincing the other. But each would hold that the other was wrong, very wrong, and would deny that the other was entitled to his or her view of the matter. When we advance claims of this sort we assert normative propositions binding on others and recognizable by all right-thinking persons. We do not assert the state of our own emotional impulses.

Still, the value-preference view of normative discourse is surprisingly common, and it is not necessary that we assume the burden of vanquishing it simply to explore the perplexities of multi-judge courts. Suppose we assume that there is a place in normative conversation at which it is appropriate for a claimant in effect to say "I have presupposed this value, and my commitment to it is nonrational, beyond debate, and wholly within my prerogatives; were you to adopt a contrary presupposition, it wouldn't be wrong, merely different." Or suppose we assume, more specifically, that, in the realm of legal adjudication, normative presuppositions of value of this sort ultimately drive the arguments and decisions of judges. To be sure, such presuppositions are properly understood to be preferences. But it hardly follows that the preference/judgment distinction collapses or loses its value. Even at its most extreme, this skeptical stance is best explained in terms of and depends upon the preference/judgment dichotomy and is not an attack on the dichotomy at all. If some or all normative discourse is understood to rest upon a foundation of presupposed, arbitrary values, it remains the case that such root normative preferences can support large structures of contingent truths of the form "Given my presupposition A, it follows that B." Such contingent truths, of course, are matters of judgment, despite their preference-based predicates, because they claim the assent of all right-thinking persons, persons who share criteria for evaluating the truth of such contingent claims.

---

10.   J.L. MACKIE, ETHICS 15-49 (1977), advances a careful argument for this position.

2. *Group Aggregation of Preferences and Judgments*

When we assign to groups the task of making certain social choices, the distinction between preference expression and judgment rendition remains. Consider a parent of three children who has promised them a night's treat. The two available options are going to the movies or playing miniature golf, and the parent puts the choice to the children for discussion and vote because he believes that the choice of entertainment should "reflect the preferences of the children." This group decisionmaking process is preference-based in two respects. First, what each child is being asked to express is his or her preference between the entertainment alternatives. Second, the collected preferences are valued intrinsically, that is, precisely because they are held by the children, and not because they are a means to an extrinsic truth. This is a paradigm of preference aggregation.

In contrast, consider the organizers of a horseracing event, who have designated a panel of three stewards to decide controversies that may arise during a race. The stewards are asked to rule on an objection by one jockey that another jockey improperly cut her off. Here, the group decisionmaking process is judgment-based. Each steward is being asked to render her judgment on questions like "Did Jockey A improperly cut off Jockey B?" The organizers value these collected judgments extrinsically, that is, not merely as a reflection of each steward's view, but as a means of determining a truth extrinsic to the stewards themselves. This is a paradigm of judgment aggregation.[11]

The distinction between preference aggregation and judgment aggregation sharpens our focus on the question of multi-judge courts. Group decisionmaking obviously commends itself in the classic democratic context: if we are in the business of structuring governance decisions for a particular group, and think it appropriate to make the preferences of the members of the group decisive over the decisions in question, then it makes perfect sense to set up a group decisionmaking process like a town meeting or a simple voting procedure. But adjudication does not fit this simple and appealing model.

Suppose first that we try to explain adjudication as an exercise in preference aggregation. What plausible theory makes the preferences of the handful of men and women who serve as judges attractive as the determinants of social policy? Set against and dwarfed by this very basic problem

---

11. Some readers will correctly observe that a utilitarian values the preferences of individuals intrinsically, but as a means to an extrinsic truth, namely, the choice of the morally correct course of action. For our purposes, there is nothing particularly confounding about this view. Utilitarianism is simply one reason for preference aggregation. Although it involves the pursuit of an extrinsic truth, it introduces that truth at a different and later stage of analysis, and makes that truth dependent on the intrinsic preferences of individuals.

is the question most pertinent to our inquiry, which is whether and why we ought to be more at ease with a system in which the preferences of three or even nine judges determine questions of importance to large segments of the society.

It will probably come as no surprise to the reader that we in fact strongly lean toward a view of adjudication as an exercise in judgment aggregation;[12] indeed, we understand most plausible schools of jurisprudence to embrace this view. But characterizing adjudication as judgment aggregation merely deepens the problem of multi-judge courts. If we care about the preferences of the members of a particular group, then canvassing that group makes good sense. But, if we seek to discover an extrinsic truth, then consultation of many rather than one is not obviously desirable. If you want the correct answer to the question "Do the children prefer to go to the movies or to play miniature golf?," assembling the moppets and asking them is an obvious way of getting it. But if you want the correct answer to the question "Did Jockey A improperly cut off Jockey B?," why get the judgments of *three* stewards? In what respects, if any, are three stewards' heads better than one? Would nine be still better than three? If legal adjudication is understood as an exercise in judgment aggregation, and what we want from courts are correct answers as in horse racing, these questions have an obvious—and obviously problematic—bearing on the operational virtues of multi-member courts.

Before we explain our sympathy for the view that adjudication is an exercise in judgment aggregation, and before we tussle with the question of whether adding Homers reduces nods, we want to complete our analysis by introducing a third, overlapping model of group decisionmaking: representation.

B. *Representation*

The distinguishing quality of representation is that one group (the active group) is expected to reach a result that emulates the result that would be reached by another group (the reference group) if the latter undertook to decide the matter. Typically, the active group is a small subset of the reference group.

Consider the producers of a costly film, who are anxious to choose between two endings, and arrange a series of screenings as a means of gauging the preferences of their national audience. The audiences in these small screenings are expected to register preferences, and the results are valued solely as a reflection of the preference distribution of the much

---

12. We offer some justification of this judgment below. *See infra* Section III.

larger audience to which the film will eventually play. This is a paradigm of representation.

Instances of representation can be categorized by the type of decision that is being reflected and by the process of reflection. The information desired about the reference group might implicate preference aggregation, where the question concerns the distribution of preferences in the reference group among possible outcomes. Or it might implicate judgment aggregation, where the question concerns the prevailing judgment within the reference group as to which of the possible outcomes is correct.

In either case, the means of getting the desired information from the active group can assume two quite distinct forms, and the questions put to the individual members of the active group will diverge accordingly. The question to the individual might assume the form: "What is your preference (or judgment) between state of affairs A and state of affairs B?" In this case, the active group will be treated as a surrogate for the reference group, with the hope that the resulting preference or judgment information will fairly approximate the pattern of preferences or judgments held by the reference group. Alternatively, the individual question to the representative group might assume the form: "What do you think is the distribution of preferences (or judgments) in the reference group between state of affairs X and state of affairs Y?" If so, the members of the active group will be expected to render judgments as to the pattern of preferences or judgments in the reference group.

Representation cases in which the active group is treated as a surrogate for the reference group, with members of the active group voting their own preferences or judgments, introduce a novel element in the modeling of group decisionmaking processes. Such cases pose the special question of the quality of the the active group's surrogacy, the "fit" of its results with those which would be arrived at by the reference group.[13] In contrast, in those cases in which the members of the active group are expected to render judgments about the anticipated response of the reference group, the active group is called upon to behave like any other judgment aggregation group.

---

13. The Supreme Court is often popularly regarded as a representative institution. *See supra* note 3. Consider Roman Hruska's remark, on Carswell's nomination to the Supreme Court, that mediocre people are entitled to representation too. The representative model, however, has wider application. One might consider the panels of three chosen to decide a case as representing the entire banc of judges on a circuit. One of the statutory provisions concerning the composition of panels is consistent with this representative view. 28 U.S.C. § 46(b) (1982) requires that the majority of the judges of each panel "be judges of that court." This provision increases the likelihood that the panel decision will "fit" the *en banc* one. We are grateful to our colleague Richard Revesz for pointing us to this example and to the example *infra* note 24.

C. *Measures of Performance for Group Decisionmaking*

These models of group decisionmaking embody distinct purposes, and hence suggest distinct measures of performance. In judgment aggregation contexts, the principal measure of performance is accuracy: the tendency of a group decisionmaking process to reach "correct" results. In any judgment aggregation situation, there are, by hypothesis, results that can be understood as either correct or incorrect, because the decisions of individuals are being received and aggregated with the view of approaching some extrinsic truth or of meeting some communally accepted criteria of evaluation. But it bears emphasizing that the notion of correctness, of accuracy, need not rest on any philosophically significant proposition about the existence of particular forms of truth. The rightness or wrongness of possible outcomes in any given judgment aggregation context may be dependent upon culture, context, politics, or the purely artificial constructs of a game or geometry.[14] There may be many or few correct answers, or a single correct answer. Accuracy is the ability to "get it right," however that favored outcome is defined or derived.

The principal measure of performance in preference aggregation is the ability of a particular process to reflect correctly the preferences of the members of the decisionmaking group. We will refer to this measure as authenticity.

"Fit" is the principal measure of performance in a representative decisionmaking process. By fit, we mean the tendency of the active decisionmaking group to arrive at results that would have been reached by the process' reference group.

The representation model of group decisionmaking and the concomitant notion of fit suggest a fourth measure of performance by which group processes may be judged—reliability. By reliability, we mean the absence of bad surprises: outcomes that differ dramatically from the result that would have been reached by the reference group. Thus, reliability measures a certain kind of stability of performance. This notion is suggested by representation because an important quality of representation, not captured in our formal rendition of the model, is the sense that the goal of reaching decisions that sit relatively comfortably with the reference group is as or more important than the aspirational goal of perfect fit. The polit-

---

14. Hart's example of scoring a game, *see supra* note 3, again provides a revealing example. Consider the pick-up basketball games that are played virtually continuously (during good weather) on the courts on Sixth Avenue between Third and Fourth Streets in Manhattan. During the course of a particular game, the rules of basketball define the "right" or "true" score at each moment of time. Clearly, any "truth" here is constructed, i.e., internal to the social practices of a particular culture and to the use of various English words. Yet, ask a spectator the score and she will render her judgment about it, not express her preference.

ical centrism implied by this goal may extend to and limit the purity of preference and judgment aggregation group decisionmaking enterprises as well; concern for political acceptability may outweigh (or at least modify) our concerns for the revelation of true preferences or the making of correct judgments.

A fifth quality of any decisionmaking process could be treated as a measure of its performance: appearance. By appearance we mean the tendency of a decisionmaking process to inspire belief among those affected by the resulting decisions that those decisions are proper. Appearance is not a freestanding quality; in group decisionmaking contexts, the question would be one of apparent accuracy, fit, reliability, or authenticity. One should not, however, mistake controversy over the appropriate substantive goal for the decisionmaking process with problems of appearance. Someone who believes that adjudication should aggregate preferences will disapprove of decisions based on and appearing as judgment aggregation.

These models, of course, are clean and simple, while the world is neither. Many, perhaps most, decisionmaking groups include among their functions all three enterprises identified here. We may think that Congress should act as a representative group when making appropriations, as a judgment-aggregating body when considering the constitutionality of its acts, and as a preference-aggregating body when setting its work schedule or establishing certain procedural rules. The point of isolating these enterprises in separate models is not to suggest otherwise; rather, it is to facilitate a better understanding of each enterprise in turn.

The world does not neatly conform to our models in another, more problematic respect. Members (or observers) of any particular decisionmaking body may disagree radically over how to understand, in terms of these models, a particular decisionmaking exercise. Such disagreement, for example, occurs frequently in discussions of adjudication, and of whether legislators should exercise their own judgment in resolving controversial issues or should "simply" represent the interests of their constituents. Resolution of any disagreement depends upon what purpose we ascribe to the decisionmaking body. Before any satisfactory evaluation of a particular group decisionmaking process can be undertaken, a prior understanding of the nature of that group's proper purpose or purposes must be in place. Such an understanding of courts and the adjudicatory process is the subject of the next section.

III. THE FUNCTIONS OF ADJUDICATION

It is common to understand theories of adjudication as lying along a spectrum defined by the reach and force of legal authority, by the extent to which one understands judges as bound or unbound by materials resi-

dent in the legal system. At the free, alegal end of the spectrum, realism views judges as charting their own individual courses unencumbered by law in all respects but form, and apparently sees no vice in this practical necessity. Positivism, the middle position, views judges as bound by legal rules as far as they reach, but free to go their own ways in the not infrequent cases in which the extant rules leave the outcome in doubt. At the far, fettered end of the spectrum, legalism views the legal system as containing resources that, appropriately understood, dictate an outcome in most if not all cases, and views judges as bound to mine the system for its guidance—guidance which is, in part, to be found in notions of consistency and coherence more subtle and complex than the fiat to follow the system's announced rules.[15]

This spectrum, ranging from realism and other alegalist positions at one end to legalism at the other, informs but does not parallel our distinctions among the three models of group decisionmaking. The legalist position commits one to a judgment aggregation view of adjudication. If the job of judges in all or virtually all cases is to mine the legal system for "right" answers, then, in setting up courts, we seek a structure that improves their tendency to produce the right results—the quality of accuracy. This easy start suggests that as we move away from legalism we move away from the idea that there are right and wrong answers to legal questions, and hence away from the view that judicial decisionmaking is an exercise in judgment aggregation. On this account, positivist courts would be in the business of judgment aggregation only up to the point where legal rules leave off, and realist courts would never be so engaged. But this tidy rendition overlooks an important distinction between two alegalistic positions, which might be called wise person alegalism and nihilistic alegalism.

Suppose a tribunal of three Wise Persons adjudicates the disputes of an appropriately small society. They do so after full discussion among themselves, and they write opinions containing their judgments and rationales. In the end, each Wise Person solemnly votes and reasons along the lines dictated by her understanding of what result is "best" for the society.

---

15. Some realists, such as Jerome Frank, represent the free end of the spectrum, while Ronald Dworkin exemplifies the legalism end. *Compare* J. FRANK, LAW AND THE MODERN MIND (1930) *with* R. DWORKIN, LAW'S EMPIRE (1986). H.L.A. Hart lies in the middle. H. HART, THE CONCEPT OF LAW (1961). Other important themes in American jurisprudence can be mapped onto this spectrum. The sharp divergence in perspective between Justice Story in Swift v. Tyson, 41 U.S. 1 (1842), and Justice Brandeis in Erie R.R. v. Tompkins, 304 U.S. 64 (1938), for example, has its roots in Story's legalism and Brandeis' streak of realism. Contemporary debate about the legitimate scope of constitutional adjudication is often moved by open or tacit controversy about the extent to which judges are understood to derive or to invent constitutional norms. In discourse of this sort, as one moves towards the realist end of the spectrum, the emphasis shifts from the importance and content of law towards the importance and integrity of judges.

How would such a process differ from the legalist view of adjudication? Missing, at a minimum, would be two related elements of the legalist view. First, the resources upon which each Wise Person can draw are not limited to the set of legal materials; however that set is defined in a particular version of legalism, it will be somewhat more limited than the set of materials available to each Wise Person. Second, each Wise Person is free to go her own way, despite prior decisions of the tribunal.[16] Rationality requires Wise Persons generally to be consistent, but this requirement obliges each Wise Person to attend to her own prior decisions, not the corporate decisions of the tribunal of which she is a part.

One alegalist view of adjudication regards judges as very much like Wise Persons. Judges on this account decide cases, or at least ought to decide cases, on the basis of their individual views of all those matters that bear on the question of what outcome and what articulated rationale are best for their society.

For our purposes, what is interesting about this Wise Person version of alegal adjudication is that judges are understood to be making judgments about what is the best outcome for their society, not merely expressing their preference as to which outcome would best suit them personally. In contrast, the other view of the unfettered judge, nihilism, holds that we ought not seek or expect "correct" judgments from judges, either because there is no meaningful sense in which judicial decisions can be qualified as right or wrong, or because judges are in any event incapable of arriving at correct outcomes through the direct application of their judgmental faculties.[17]

---

16. This license obtains at least on the occasion of the tribunal's first decision, and at least if we imagine that it has begun doing business in a freshly minted society of some sort. These heroic provisos are necessary because the alegal freedom of each Wise Person is subject to powerful erosive forces. Once the tribunal of Wise Persons has begun to accumulate a body of decisions, claims of equal treatment and reliance will gather increasing ethical force in the analysis of each Wise Person, while paradigms of thought will shape their analyses in less explicit but highly significant ways. It may be that the concept of such a tribunal undermines itself in a way that undermines the plausibility of the alegalist vision.

17. Realists did not articulate their theory of adjudication in our terms, and it is thus difficult to identify individuals who might hold (or have held) the nihilist or Wise Person version. We suspect, however, that most realists are of the Wise Person persuasion. Karl Llewellyn, for example, in K. LLEWELLYN, THE COMMON LAW TRADITION (1960), cannot be interpreted as denying either the competence of judges to reach correct answers or the existence of better or worse solutions to legal dilemmas. Thus, he would not appear to be a nihilistic alegalist. Arguably, he is not an alegalist at all because his analysis of adjudication generally assumes that judges restrict their attention largely to "legal materials" and that they adhere to some amorphous set of legal practices, which constrain and shape their judgment. The indeterminacy thesis of critical legal studies is sometimes offered as a form of nihilistic alegalism. Often, however, this claim is made in the context of a particular set of justificatory arguments. Thus, for example, Heller argues that law and economics is incapable of providing answers to every legal question. Heller, *Is the Charitable Exemption from Property Taxation an Easy Case?*, in ESSAYS ON THE LAW AND ECONOMICS OF LOCAL GOVERNMENTS 183 (D. Rubinfeld ed. 1979).

While either Wise Person or nihilist alegalism is formally available to the positivist as a means of filling the void beyond the reach of established rules, the nihilist approach seems poorly suited to the task. A nihilist positivist would have to believe (1) that judges were capable of determining and following the rules of a complex legal system, but incapable of determining the correct outcome of any controversy beyond the pale of established rules, and (2) that despite this judicial incapacity, once a court had more or less arbitrarily set a new rule in place, that rule would and should enjoy abiding respect in future cases.

Our discussion suggests that it is quite possible for realists and positivists to share with legalists the general view that courts are engaged in the enterprise of judgment, that it is meaningful to regard judicial outcomes as correct or incorrect, and that it is of course desirable that judicial institutions be structured to maximize correct results. For the legalist and for the Wise Person positivist and realist, accuracy—the tendency to reach correct judgments—emerges as a core virtue of the adjudicatory process even while they differ about the nature of the process.

Assuming courts remain more or less as we know them, what, for the nihilist, would count as virtue for the process of adjudication? One possibility is to seek from judges simply an authentic expression and tally of their personal preferences in each dispute. But it is hard to imagine why, without some added ingredient of value, the personal preferences of the men and women who happen to be judges should be an even marginally attractive source of social choice; careless bad guesses or outright lies on the part of judges about their preferences seem as worthy a source of choice. The same problem arises from the proposal that what we want from judges are their honest efforts at judgment, even though we—as nihilists—know those efforts to be doomed.

If the preferences of judges or their futile attempts at judgment have value, it must be in their representative capacity as surrogates for society as a whole. On this view, there is a correct outcome, namely that which would be reached by the polity in a plebiscite; but judges lack the capacity to determine that outcome. If, contrary to this nihilist-representative view, judges enjoyed the capacity to determine the polity-favored outcome, we would be restored to judgment aggregation, with a particular, majoritarian understanding of what constituted a correct outcome. In the nihilist version, judges can get it right, but only by inadvertantly mirroring the society of which they are a part. The adjudicatory virtue for a nihilist who held this view would be that of fit between the courts and the rest of society.

Also available to the nihilist as an adjudicatory virtue is appearance. Bereft of hope for decisions that are substantively correct, the nihilist

could still seek courts so structured that their decisions impress a naive population with their worth.

Appearance, of course, could be a virtue for the legalist or Wise Person alegalist (either positivist or realist) as well. It is perfectly consistent to desire both that the judicial system deliver correct judgments and that it have the appearance of doing so. Less clear is whether it is consistent to perceive the job of judges and the virtue of courts as that of arriving at correct judgments, and yet see the fit between courts and the rest of society as a virtue as well.

The identification of accuracy, authenticity (as a predicate for fit), reliability, fit, and appearance as potential, positive measures of judicial performance brings us to the nub of our inquiry: How, if at all, can these qualities of adjudication be improved by increasing the number of judges who make up a court? In what follows, we restrict our detailed attention to the feature of accuracy and the question of why three heads might be better than one as a means of "getting it right," or why nine might be better than three. We adopt this focus because we, in common with most philosophers of law, consider most adjudication to be judgment-based, and because it is the relationship between accuracy and tribunal size that seems most perplexing. Close consideration of the other possible measures of judicial performance as they relate to multi-judge courts we leave for another day.[18]

---

18. We are unable to resist some brief comments on how increasing the number of judges might improve the quality of both preference and judgment adjudication as measured by authenticity, fit and reliability.

A preference-based theory of adjudication raises many perplexing problems, including what it might mean for a group to have a preference, and under what circumstances a collective choice rule exists. Authenticity, the principal quality measure of preference-based adjudication, may be affected in two ways by the size of the judicial panel. Smaller panels are likely to have more homogeneous preferences and hence admit a collective choice rule. On the other hand, increasing panel size from one to three or three to nine may improve the quality of deliberation. The panel may be better able to identify the range of possible decisions and their attributes. Similarly, more broadly based deliberation may clarify each judge's preferences over the set of alternative decisions.

The analysis of fit again requires distinguishing preference- from judgment-based adjudication. In preference-based adjudication, fit as a criterion reduces to accuracy if we assume that the size of the group does not affect the quality of deliberation. Judgment-based adjudication presents an additional problem. The best fit for which we may hope would be that the judgmental competence of the active group equal that of the reference group. If everyone has identical judgmental competence, fit and accuracy again are equivalent. If, however, people differ in their judgmental competence, accuracy implies that we should select judges with the highest judgmental competence, while fit requires us to select judges randomly.

Reliability is the property of avoiding outrageous outlying results. Hence, the analysis of reliability depends on the interpretation of dispersion, or unreliability, one offers. First, we might care about the reproducibility of the result, i.e., how likely it is that a different active group drawn from the same reference group would reach the same answer as the first active group. Reproducibility requires that adjudication be neutral with respect to the set of judges before whom a case appears. Unless everyone has perfect judgmental competence, perfect reproducibility is unattainable. It can be shown, however, that the probability of a second group's reproducing the result of the first increases with the number of judges.

In judgment-based theories of adjudication, notions of consistency and coherence cloud the general question of the accuracy of decisions and are especially problematic features of any effort to adjust that theory to take account of the phenomenon of multi-judge courts. Accordingly, we follow our general discussion of accuracy with a close look at consistency and coherence in the multi-judge context.

IV. MULTI-MEMBER COURTS AND ACCURACY

The inquiry into the relationship between court size and accuracy requires consideration of two different facets of the multi-judge adjudicatory enterprise: (1) the basic fact of aggregation—that the judgment of more than one judge will figure in the resulting decision, and (2) the process of deliberation. We will consider these in turn.

A. *Aggregating Judgments*

A highly simplified model of adjudication indicates how increasing the number of judges can improve judicial accuracy. Suppose the following were true of the adjudicatory process: (1) adjudicated controversies have only two possible outcomes; (2) all judges are equally likely to choose the correct outcome; (3) each judge is more likely to choose the correct outcome than the incorrect one;[19] (4) simple majority rule is the voting procedure; and (5) each judge's decision is independent of the other judges' decisions. (This last condition formalizes the non-deliberational aspect of the procedure. If deliberation matters, we would expect that one judge's decision would affect other judges.)

Given these circumstances, we may consider each judge's decision as the

---

The second interpretation of dispersion, and hence reliability, raises questions beyond the adjudicatory model underlying the previous discussion. This alternative model seeks to formalize the conception of adjudication as selecting a position from the set of alternatives arranged along a spectrum from "liberal" to "conservative" with a "reliable" mechanism selecting a "centrist" position. More precisely, reliability as stability assumes that courts confront more than two possible outcomes about any two of which we say, "outcome A is within (distance) d of outcome B." If outcome A* is the outcome that the reference group would have reached, then a decision mechanism X is more reliable than a decision mechanism Y, if mechanism X is more likely to reach a decision close to A* than mechanism Y. Mathematically, this criterion is, for every d, the probability that the outcome under X is within d of A* is greater than the probability that the outcome under Y is within d of A*. If we consider majority rule procedures in which judges are selected randomly from the reference group and the preferances of each potential judge are single-peaked, then judicial panels will select the preferred point of the median judge. In this case, the claim that increasing the number of judges (N) on a panel increases the reliability of the court is equivalent to the claim that as N increases the distribution of the median order statistic gets "tighter" around the true median. We have not succeeded in proving this conjecture.

19. Of course, if each judge were more likely than not to get the answer wrong, we would not seek her judgment about the outcome. As discussed *infra* text accompanying note 20, when we allow different members of the panel to have different degrees of judgmental competence, our analysis requires only that the mean competence exceed ½.

draw of a single marble from a bag with marbles of two colors (white for a correct decision, blue for an incorrect decision), mixed in proportion to the likelihood of any judge's choosing the correct outcome. Adding judges simply adds draws (with replacement); as long as the proportion of white marbles in the bag exceeds ½, the more draws there are, the more likely it becomes that more than half of the marbles drawn will be colored white or "correct." The fact that there are more judges on a panel thus implies that the panel is more accurate, i.e., more likely to reach the correct decision.[20]

The positive relation between improved accuracy and greater number of judges persists when we relax various of our initial assumptions. For example, the analysis changes little when we adopt the more reasonable assumption that the probability that a given judge will reach a correct decision varies among judges. If we assume that judges generally get it right—that individual likelihoods of correct decision are drawn from a normal distribution with mean $p > ½$ and variance $p(1-p)/N$—then as N, the number of judges, increases, the probability of a correct decision increases as well.[21] Moreover, for small groups, simple majority rule makes it more likely that the group will produce a correct decision than that the most competent member of the group will.[22]

Thus far we have assumed that the panel of judges may choose between only two possible outcomes. Most appellate decisions conform to this requirement in the sense that there will be only two possible mandates (reversed/affirmed or judgment for appellant/appellee).[23] We may, however, relax this assumption (believing instead that judges choose from many

---

20. This theorem was first published in Condorcet's 1785 *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, in CONDORCET: SELECTED WRITINGS 33 (K. Baker ed. & trans. 1976). For proofs of this result, see Grofman, Owen & Feld, *Thirteen Theorems in Search of the Truth*, 15 THEORY & DECISION 261 (1983) [hereinafter Grofman]; Nitzan & Paroush, *A General Theorem and Eight Corollaries in Search of Correct Decision*, 17 THEORY & DECISION 211 (1984).

We assume that the panel always has an odd number, $N = 2n + 1$, of judges. Each time n increases by 1, N increases by 2. The proofs of the theorems stated in the text thus proceed by induction on n.

21. In terms of the example above, we now have N judges with N bags. Although we do not know the proportion of white marbles to blue marbles in each bag, we do know that if we dumped them all together we would have the same proportion as before.

22. Grofman, *supra* note 20, and Nitzan & Paroush, *supra* note 20, offer proofs of this result. Corollary 1 to Theorem V in Grofman, *supra* note 20, at 268, states: "If judgmental competence is normally distributed with mean $\bar{p}$ and variance given by $[\bar{p}(1-\bar{p})/N]$, then we obtain results essentially identical" to those when $p = \bar{p}$ for all panel members—the equal competence assumption.

The result related to Theorem XI is "[i]f judgmental competence is normally distributed with mean $\bar{p}$ and variance given by $[\bar{p}(1-\bar{p})]/N$, then it is more probable that the majority choice in small groups ($N < 35$) will be correct than the judgment of the most competent member of the group . . . ." *Id.* at 272.

23. We return to the problem of multiple outcomes below. *See infra* text accompanying notes 41-44.

outcomes rather than between two judgments) without changing our conclusion. It will still be true for a fixed number N of judges that as the probability of each judge's reaching a correct decision increases, the probability of the panel's arriving at a correct decision also increases. Moreover, if each judge is more likely than not to reach a correct decision, then, as N increases, the probability of the panel's reaching a correct decision also increases.

Another assumption under which we have been operating so far is that the judges will decide cases by simple majority vote. Simple majority vote specifies that the number of votes necessary to reach a given outcome is the smallest integer that is greater than one half of the number of voters, $N/2$. The most obvious way to vary this voting procedure is to increase or decrease the number of votes required to reach a given outcome. By doing so, a panel of N judges would have $N + 1$ voting procedures among which to choose, requiring any integer from 0 through N as the requisite number of votes.

Without further modification of the voting procedure, the options thus introduced have odd properties. If one adopts a sub-majority rule, the court could reach two or more outcomes simultaneously, and these could be starkly contradictory in nature. For all super-majority rules, the court could fail to reach any outcome at all. Where there are only two possible outcomes, simple majority rule emerges as uniquely decisive; it alone always identifies a single, correct outcome. But when we allow more than two possible outcomes, this property of majority rule disappears.

In practice, sub- and super-majority rules are made workable by the addition of favored or default states. Where sub-majority rules are at play, typically one outcome is favored, in the sense that it will be adopted if it receives k votes, whether or not some other outcome also receives k. The Supreme Court's "rule of four" in certiorari grants is an example of a favored state. Where super-majority rules are at play, one outcome is the default state, in the sense that if no other outcome receives k votes the default state will prevail. The Article V amendment procedure in the Constitution is an example of the use of a default state; absent the required super-majorities, a proposed amendment fails.

Favored and default states make sub- and super-majority rules decisive. They also make these rules nearly equivalent, in the sense that the rules "Four votes are required to grant certiorari" and "Six votes are required to defeat certiorari" on a nine-judge court are distinguished only by the impact of abstentions.[24]

---

24. In fact, the Supreme Court rules are not as clear as the text suggests. While the positive votes of four judges are required to grant certiorari, the positive votes of five judges are sufficient to lead to a decision that certiorari was improvidently granted. The Court does not report the votes of individual

The choice of voting procedure from among the various sub-majority, simple majority, and super-majority options determines the type of error that is favored in a judgment aggregation enterprise. Consider the actual context of appellate review, where the central question of outcome is whether the lower court is to be affirmed or reversed. Typically, appellate courts in our legal system operate on the basis of simple majority vote; if the number of judges is odd, this procedure is decisive on the binary question "affirm/reverse." There is no need for either a default or a favored state. Either possible error—affirming an incorrect decision or reversing a correct decision—is equally possible. By adopting a different rule, "The decision below is affirmed unless four of the five judges on this court vote to reverse," the court would favor errors involving the affirmance of wrong decisions and disfavor errors involving the reversal of right decisions. The degree and direction of such error deflection could be adjusted easily by adopting various sub- or super-majority procedures and designating the appropriate favored or default states.[25] Thus, explicit consideration of the effect of multi-member courts and the selection of voting rules can be crucial to understanding why a particular rule of law emerges.

B. *Deliberation*

Our analysis thus far has assumed that the generation of individual judgments or preferences in group decisionmaking contexts is hermetic, and that the impact of the group environment is felt only at the aggregation stage. In fact, most group decisionmaking includes the exchange of ideas and arguments under circumstances that are calculated to affect individual views. Where the collective enterprise involves judgments rather than preferences, and where the discursive process is formal and structured—as with multi-member courts—it is often called deliberation. Deliberation obviously can affect the attributes of a group decisionmaking

---

Justices on the granting of the writ; it is therefore possible that four Justices could vote to grant certiorari and then have their wishes and their votes frustrated if the five Justices who voted against granting certiorari vote to dismiss the writ as improvidently granted.

Presumably the collegiality of the Court prevents such an outcome, as it has in a related context. The Court requires five votes to grant a stay but only four to grant certiorari. In death-penalty cases, this difference has great significance because failure to stay results in the execution of the petitioner and the mooting of his case. In Darden v. Wainwright, 106 S. Ct. 21 (1985), Justice Powell voted, at the last minute, to stay an execution since his failure to do so would have frustrated a previous grant of certiorari supported by only four other Justices.

25. This designation of default states may play a particularly important role in attempts to maintain coherence in a legal system. *See infra* text accompanying notes 33-34. Errors in one direction may be perceived to do much greater violence to the fabric of the system than errors in the other direction.

process; more particularly, it can affect the accuracy of judicial decisionmaking.

In the course of deliberation, judges can be expected to advance arguments in support of a particular result in the case before the court. Arguments serve two functions in legal analysis: they are both the reason that a judge reaches (or abandons) a particular result, and the announced rationale for having done so, carrying implications for the outcome of future cases. The process of testing any such argument includes canvassing the alternative arguments, elaborating the consequences of these arguments, and discussing the criteria of validity. Unless one presupposes that judges enter this process either fully aware of all points and counterpoints, or obdurately fixed on both result and rationale without reference to the full range of possible argumentation, it must be the case that some judges emerge from the deliberative process with judgments about the appropriate result or rationale different from their initial ones.

Deliberation could affect the behavior of any given judge in at least three different ways. First, a judge may become more fully aware of the range of rationales for one or another possible result and, in some cases, more fully aware of possible results; in this sense her judgment set will have been expanded by deliberation. Second, deliberation may convince a judge that there are good reasons to change her mind. A more complete understanding of the consequences of adopting a particular rule would be a commonplace reason for a change of this nature; less common perhaps, but more dramatic, would be exposure to an entirely different perspective on a given problem, as where a judge is persuaded to reconceive as a torts problem a case she had originally characterized as a breach of contract. Third, the communal pressures of deliberation may induce in a given judge a conscious or unconscious impulse to conform her judgment to a range of results that her colleagues treat as acceptable.[26]

---

26. The impact of deliberation on adjudication seems at least in part redundant: the adversarial process also serves to enumerate alternatives and to frame the context in which a judge exercises her judgment or asserts her preferences. Closer inspection reveals that deliberation and adversarial exchanges may complement each other. While each party will advocate positions most favorable to it, this advocacy ensures that any biases of the judges are exposed to conflicting arguments and views. Similarly, the judges cannot make a simple neutral choice between the arguments offered by each side. The adversaries may urge rationales that have no common ground.

Our conception of the role of the judge and the function of adjudication also suggests why the process might include both deliberation and adversarial argument. Our conception of the judge as arbiter implies, as noted above, that she should formulate the decision or judgment problem neutrally as between the parties. Our understanding that judicial decisions affect parties not before the court and our conception of just treatment of these unrepresented parties argues that judges should consider the interests of those parties not before the court because the adversarial process will not necessarily produce arguments and options favorable to these persons.

Of course, the adversarial process serves other functions in addition to deliberative ones. It allows the parties most directly affected by the judgment to participate in the process and hence may enhance the perception of fairness and legitimacy of the outcome.

Although we have identified ways in which deliberation might have a significant impact on judges and their judgments, we have not demonstrated that deliberation improves the accuracy of judicial decisions. Deliberation in a reflective community exposes individuals to seriously maintained ideas and arguments that they previously overlooked or treated summarily; these ideas and arguments, of course, may be good or bad. Deliberation also generates communal pressure to eschew ideas and arguments that fall outside the range acceptable to the community; such radical ideas and arguments, again, may be aberrant and wrong, or innovative and correct.

Thus, if we favor deliberation as a means of improving group judgment, it must be because we assume that its judgment-enhancing aspects outweigh its judgment-impairing features. Our reasons may include the view that ignorance of useful ideas or arguments is a greater hazard to the capacity of reflective individuals to form correct judgments than exposure to false ideas or arguments, or the view that the deliberative process will do more good in exposing obvious errors and restraining mad acts than harm in suppressing novelty. Demonstrating the validity of these propositions is complex, partly empirical, and beyond our resources here, but the general assumption favoring deliberation as an aid to correct judgment seems reasonable in light of common experience.

V. CONSISTENCY, COHERENCE, AND GROUP DECISIONMAKING

In many judgment-based theories of adjudication, "getting it right" entails maintaining a particular, substantive relationship between the decisions of individual cases. The required relationship is sometimes described as "consistency," sometimes as "coherence," and is commonly understood to be a necessary condition of correct outcomes. Though the terms "consistency" and "coherence" are often used loosely and interchangeably, we think they entail two related, but distinct concepts. In this section, we explore consistency and coherence in the context of multi-member courts. Specifically, we ask whether a panel of judges, each of whom has a consistent and coherent view of the law, will, when they aggregate their judgments, produce a consistent and coherent pattern of decisions.[27]

---

27. To our knowledge, only one prior work has approached these questions in any detail. *See* Easterbrook, *Ways of Criticizing the Court*, 95 HARV. L. REV. 802 (1982). Judge Easterbrook's discussion differs dramatically from ours on a variety of grounds. First, he does not distinguish between consistency and coherence, a distinction we believe fundamental to the understanding of adjudication. Second, we consider explicitly situations in which the application of criteria for coherence are path-dependent and in which the judicial view of the law evolves over time. Third, and perhaps most fundamental, Easterbrook views the court as engaged in preference aggregation, rather than judgment aggregation. ("Each group of Justices, moreover, has a preference between the two positions taken by the others." *Id.* at 815.) These differences render impossible a point-by-point comparison of the two

We begin by defining the concepts of consistency and coherence in the context of the decisions of a single judge. Of the two concepts, consistency is easier to grasp and imposes a less stringent requirement. For our purposes, consistency is simply the state of non-contradiction, and two legal rules are inconsistent if and only if they are contradictory. Thus, any two of these three rules are inconsistent:

(1) Given circumstances A, B, and C, a landlord is obligated to avoid doing X.
(2) Given circumstances A, B, and C, a landlord is under no obligation with regard to the doing of X.
(3) Given A, B, and C, a landlord is obligated to do X.[28]

This simple, straightforward understanding of consistency grows more complex when applied to a sequence of decisions by a single judge. Suppose we have a small state with one judge, Liza (who is an infallible arbiter of factual disputes). When Liza decides controversies brought before her, she simply recites an elaborate statement of the factual circumstances surrounding the controversy ("the story") and announces the outcome, such as that the landlord must pay her tenant $y, or $0, nothing at all ("the ruling"). In such a world, it is hard to understand how Liza's decisions could ever be inconsistent. "Rules" in this world consist of matched pairs of extensively described controversies and the associated outcomes; they read "If [story] then [ruling]." Because the date of the events, the parties involved, or other details will differ, no two stories will be exactly alike, and thus no two rules can be directly contradictory or inconsistent.

Judges in legal realms familiar to us, however, do not behave like Liza. They write opinions that purport to announce and justify the applicable legal rule or rules.[29] In contrast to Liza's matched stories and rulings,

---

articles. But we do take issue directly with Easterbrook's central claim that multi-member courts are organically prone to inconsistent behavior. *See infra* notes 35–37 and accompanying text.

28. However, two rules are not made inconsistent simply because they might operate to place the same individual under inconsistent obligations:

Given A, B, and C, a landlord is obligated to do X.

Given C, D, and E, a landlord is obligated to avoid doing X.

It is quite possible in a consistent regime of rules for an individual to burden herself with contrary obligations, as when a landlord enters into binding agreements both to do and to refrain from the same act.

29. The criteria for the determination of the ratio decidendi, or holding, of a case have received extensive, if inconclusive, analysis in the jurisprudential literature. *See, e.g.*, R. CROSS, PRECEDENT IN ENGLISH LAW (3d ed. 1977). Broadly speaking, one may identify two approaches. The first, intimated in the sentence to which this footnote is appended, takes seriously the stated rationales of the judges themselves. The second approach, more consistent with the view of adjudication implicit in the discussion of this section as a whole, focuses on the pattern of facts and associated outcomes of the cases. It has as its source the legal realist movement with K. LLEWELLYN, THE COMMON LAW TRADITION (1960), providing perhaps the best example. In this latter approach, the rule announced

these rules isolate and identify the particular circumstances that are rele-
vant to the stipulated legal consequences.[30] Legal rules thus assume the
abstract form "If q then r," where q is the set of those circumstances
sufficient to bring about r, which in turn is the set of all legal conse-
quences comprehended by the rule. When the circumstances of legal con-
troversies relevant to the outcome are isolated and identified in this fash-
ion, the resulting rules are abstract enough to allow one to generate and
predict future outcomes. In such an environment the concept of inconsis-
tency once again becomes meaningful. Where two legal rules attach differ-
ent legal consequences to the same stipulated set of relevant circumstances
the rules are inconsistent.[31]

So understood, consistency is not hard to justify as a virtue. It serves the
goal of treating persons subject to the adjudicatory process fairly, and is
essential to the ability of affected persons to anticipate legal outcomes and
plan their affairs accordingly. It is our desire for consistency that in sig-
nificant part animates and shapes the rule of stare decisis.

---

by a judge may not articulate the true rationale for the decision; rather, the rule of law must be
extracted from a careful examination and comparison of the fact patterns and outcomes of closely
related cases. The realist method, of course, remains dependent upon a judge's actual words, because
the pattern of facts analyzed is extracted from the opinion of the prior judge. Moreover, the an-
nounced rules of judges are generally good evidence of the true rules and, in any event, reflect the
rule-oriented thinking and behavior of judges generally.

30. Most aspects of a litigated event are generally considered irrelevant to the determination of the
legal consequences. For example, the dates of birth of the parties, the color of the defendant's hair, her
educational background, and, often, many of the social relations between the parties have no bearing
on the formulation of the legal rule. Formally, the typical rule introduces some quantification over
these "irrelevant" aspects of a story. That is, to state fully the legal rule that party A who fails to
deliver to party B a contracted-for widget on a specified date is legally liable for damages, one must
append a variety of suppressed clauses that identify the scope of various terms in the statement of the
rule. For example, the usual statement of the legal rule omits, among others, the clauses "regardless of
the name of party A, regardless of the name of party B, regardless of the hair color of A and of B,
. . . ." Logic allows us to formulate these rules by stating that the legal rule applies to every situation
that meets certain conditions. Thus, the contract example noted above might have a structure some-
thing like "for every person A and for every person B such that A made an offer at time t_1 and B
accepted such offer at time t_2 (after t_1) for a given consideration, if A failed to meet the terms of the
offer, then . . . ." Of course, to state completely a simple rule in this form presents grave difficulties.
In the example above, we would have to define the terms "offer," "acceptance," and "consideration"
as well as spell out the salient aspects of A's actual offer and the types of conduct that conformed to
the terms. We might represent this complex formulation of a simple rule with the following form:
"For every date t (or series of dates t_1, t_2, . . . t_n with $t_i < t_j$ when $i < j$), and for every pair of
people x and y with the properties $P(x,y)$, $Q(x,y)$, and $S(x,y)$ then r." Of course, the rule might be
more complex. It may involve relations among more than two parties and more than two properties.

31. A legal rule assumes the rough form "If [pertinent circumstances] then [legal consequence]."
In the example above, see supra text accompanying note 28, the inconsistent legal rules have identical
circumstantial parts and differ only in the legal consequence. Specifically, the legal consequence of one
belongs to the set of legal consequences defined by the negation of the other. Inconsistency may arise,
however, even when the circumstantial parts of the legal rule differ. Such inconsistency can arise
because two legal rules give different scope to some variables. See supra note 28. For example, the
abstractly phrased rule f^*, which provides that "For each a, $f(a,b^*)$ implies q" is inconsistent with the
rule f, which provides that "For each a and for each b, $f(a,b)$ implies $not\text{-}q$." Assuming that b^* is a b,
the antecedent (or [circumstance]) of the latter rule implies that for every a, $f(a,b^*)$ is true and thus
the two rules are inconsistent. That is, whenever $b=b^*$, the two rules would require q and $not\text{-}q$.

Nor is consistency particularly hard to achieve, at least in a one-judge world. But even in such a world inconsistency could occur by mistake or design. Concurrent inconsistency—the simultaneous existence of two contradictory rules—presumably can only be the product of error. But consecutive inconsistency—the recognition first of one rule and later of a contradictory successor can well be intentional, as when a rule is abandoned because it is thought to be fundamentally wrong, because, for example, it detracts from the coherence of the main corpus of extant rules.[32]

This last suggestion brings us to the concept of coherence. Coherence is a quality of conceptual unity. Coherence does not require that a system's premises be correct, but it does demand that they form or reflect a unitary vision of that portion of the world modeled by the system. A perfectly coherent legal system would comprise normative elements derivable from a relatively limited number of non-contradictory premises that are reasonably general in form and that join in a recognizable conception of social policy.

As a virtue of systematic thought and discourse generally, coherence stands between the purely formal quality of consistency and the substantive (typically normative) quality of soundness. A system is consistent if all of its elements are derivable from a set of non-contradictory premises. It is sound if, in addition, those premises are correct. Coherence lies between these bounds: A coherent system must be consistent but it need not be sound. Coherence thus brings some order and structure to what might otherwise be a jumble of consistent propositions. This important substantive restraint on a system of thought or discourse is imprecise and bound up with the intellectual culture of which it is a part.

Like consistency, coherence is not hard to justify as an adjudicatory virtue. It serves the same goals of treating persons subject to the adjudicatory process fairly and enabling such persons to anticipate legal outcomes and plan their affairs accordingly.

An example may clarify our discussion of the concept of coherence. Judge Skelly Wright's recognition of an implied warranty of habitability in *Javins v. First Nat'l Realty Corp.*,[33] which wrought a major change in landlord-tenant law, is grounded in a judicial appeal to coherence. Under the pre-*Javins* common law, two perfectly consistent but very different rules were well-established: One governed landlord-tenant relationships;

---

32. Consecutive inconsistency can only arise in a single judge world if the judge, at the time she starts, has an incomplete vision of what the appropriate body of legal rules should be. If she knew at the outset what the "right body of legal rules" was, the sequence of cases would simply provide the order of announcement of the rules. In a multi-judge world, consecutive inconsistency might arise even among judges each of whom had a complete vision of the body of law, because their visions might differ.

33. 428 F.2d 1071 (D.C. Cir.), *cert. denied*, 400 U.S. 925 (1970).

the other governed most of the remaining universe of bilateral promissory relationships. Rather than ground his decision in extrinsic policy claims, Judge Wright reasoned from the pull of material already resident in the legal system. Recognition of the implied warranty of habitability was understood as eliminating a major source of incoherence in the law of bilateral promissory relationships.

Javins illustrates the tension that can exist between consistency and coherence. Consistency is a necessary condition of coherence, and hence the perfectly coherent legal system would also be perfectly consistent. But in an imperfectly coherent system, coherence and consecutive consistency can conflict.[34] Thus, in *Javins*, Skelly Wright could be consistent by following well-established precedent in landlord-tenant law, or coherent by bringing landlord-tenant law into line with promissory relationships generally. He could not be both.

The tension between consistency and coherence has interesting implications for an adjudicatory system. Suppose Liza is the only judge in a common law state, and that over the course of her judicial career she has decided hundreds of cases. She may decide case 51, involving a complicated and arcane issue of tort law, and then not encounter the same issue again until case 451. By the time case 451 presents itself, she may realize that a long line of her decisions in the contracts area has important implications for the 51/451 issue, implications directly contrary to her decision in 51. Coherence and consistency would conflict here, as in *Javins*, and we can well imagine Liza preferring coherence to consistency. Of course, if from the outset Liza is preternaturally able to anticipate all her future cases or in some other way able to behave with perfect coherence throughout her career, the problem will never arise. But even our idealized Liza can encounter the problem if we take a step from our one-judge world to a world a little more real, and posit a legal system with a number of judges, each of whom sits alone and each of whom has equal authority.

---

34. The claim of the text rests on the following understanding of coherence. Consider two idealized bodies of law, each consistent and minimally coherent. The first body of law corresponds in general to the pre-*Javins* state of the law, which treats landlord-tenant relationships differently than other bilateral promissory relations. The second body of law corresponds to the post-*Javins* world, and treats all bilateral promissory relations in a unitary way. This post-*Javins* legal corpus is idealized in the sense that it does not contain those cases decided prior to *Javins* that were inconsistent with the unitary view of bilateral promises. The claim is that the post-*Javins* corpus better satisfies the criteria of coherence than the pre-*Javins* corpus. Although each body of law may be derivable from some smaller set of principles, the claim of greater coherence need not be a claim that the post-*Javins* set of generative principles is logically consistent while the pre-*Javins* set is not. In an idealized body of law, the generative set is by hypothesis consistent. But even such idealized sets will differ in the degree to which they satisfy non-logical criteria embodying, for example, our norms of simplicity, unity, and elegance.

The decision in *Javins*, of course, could not yield a perfectly consistent body of law. It introduces a landlord-tenant rule consecutively inconsistent with the prior landlord-tenant rule. It does so to advance the goal of greater coherence.

Now, if Liza decides case 51 and returns to the same issue in case 451, she has to contend with an intervening body of decisions including the decisions of Judges A, B, C, D, etc. Even if Liza's colleagues have behaved consistently with the rule of case 51, they may have decided other cases that raise problems of coherence, and when Liza returns to the issue in 451, she may well confront a clash between coherence and consistency.

Since the commands of consistency and coherence may well point in opposite directions on occasion, it follows that we cannot insist or expect that judges always behave consistently and coherently. It also follows that adjudication that seeks to be faithful to the requirements of consistency and coherence will display the quality of path-dependence; that is, adjudicated outcomes may vary depending on when they are decided. If an issue comes up as number n it may be decided one way by a conscientious and perfectly competent judge, but if it comes up as number $n+m$ it may be decided differently by the same judge. Path-dependence is a kind of irrationality closely connected with inconsistency, and carries with it a parallel sense of unfairness. Like the trade-off between coherence and consistency, however, it is endemic to systems of adjudicated outcomes, whether participating tribunals are staffed by a single judge or by groups of judges, except in the nonexistent world of a preternaturally perfect, single judge (who lives forever!).

When we turn to multi-member courts, the analysis of a panel's ability to achieve consistent and coherent decisions becomes more complex. Some believe that multi-member courts cannot achieve consistency at all. The argument against the consistency of multi-member courts rests on the possibility that judges on a multi-member court may disagree on the rules applicable to the decisions of specific cases, and that this disagreement will lead to a reeling inconsistency, with the court lurching from rule to contradictory rule.[35]

---

35. Easterbrook, *supra* note 27, at 814–817, presents the argument for inconsistency. He relies on Arrow's Theorem, and in particular on the problem of "cycling." The problem of cycling is introduced if we imagine a three judge court, with the judges obliged to choose among three rationales, A, B, and C. Each judge ranks these rationales in order of preference (where A > B denotes that rationale A is favored over rationale B):

Judge	Rule Ranking
R	A > C > B
S	B > A > C
T	C > B > A

Easterbrook apparently conceives of a court under such a circumstance conducting a series of futile pairwise votes, pursuant to which these discomforting results would emerge:

Rule B would beat Rule A (Judges S & T vs. Judge R)
Rule C would beat Rule B (Judges R & T vs. Judge S)
Rule A would beat Rule C (Judges R & S vs. Judge T)
Rule B would beat Rule A . . .

and so on endlessly, unless the court chose a wholly arbitrary place to stop. This is the problem of

This argument is false. Assuming that each judge on a multi-member court decides cases consistently, then the court as a whole will decide cases consistently. Consider the case most likely to lead to inconsistency, one in which each judge on the panel is firmly committed to a different rule. Suppose that, on a three-judge court, Judge A holds that commercial speech is unprotected by the First Amendment, and that regulations of other speech are constitutional only if they are content-neutral and supported by legitimate governmental concerns; Judge B holds that all speech is absolutely protected; and Judge C holds that the non-commercial speech of natural persons is absolutely protected, commercial speech is subject to reasonable regulation, and that the speech of juridical entities is unprotected. Suppose further that three state laws are challenged on First Amendment grounds; these laws respectively ban: (1) "false or misleading" advertisements, (2) leaflets in any public place (because of litter), and (3) corporate participation in legislative referendum campaigns. We would get the following voting pattern from our judges:

LAW:	False Advertising	Public Places	Referenda
JUDGE:			
A	OK	OK	BAD
B	BAD	BAD	BAD
C	OK	BAD	OK
PANEL:	OK	BAD	BAD

Should any of the three cases recur, the court will decide it as it decided that case the first time. Specifically, it will approve of laws banning false advertising and disapprove of laws banning leafletting in public spaces and corporate participation in referendum campaigns. Thus, though this court lacks a majority rationale, it has a consistent rule.[36] Its rule, while

cycling, which is a critical element in Arrow's Theorem and forms the heart of Easterbrook's claim. But as our analysis of a radically fractured court will demonstrate, *see infra* text accompanying note 36, a perfectly stable voting pattern emerges from such a court. If the individual judges decide cases consistently, like cases will be decided alike, and no problem of consistency presents itself.

36. This table provides a useful visual means of depicting not just consistency but coherence as well. Consistency is a logical relationship between events that depends simply on the stability of the vertical elements in the table. Thus, if in each case involving a ban on false and misleading advertising that vertical line is constant, with Judges A, B, and C maintaining their respective positions, the court will adjudicate the case in the same way, and consistency will result. Coherence—to anticipate our discussion—is a substantive relationship along the horizontal elements of the table. For a given judge, say Judge A, to be coherent she must maintain a particular relationship among her votes in these three sorts of cases; for the court to be coherent, the outcomes in the bottom horizontal line must enjoy that same substantive relationship. Coherence demands that the horizontal results be derivable from some unitary set of principles or embedded in some structured theory. The rule of any given judge, or of the court, is coherent only if it constitutes a formulation of First Amendment doctrine that satisfies this requirement. Moreover, even if each judge decides all three cases coherently, it does not follow that the court's outcomes will be coherent. *See infra* text accompanying notes 39-40.

different from the rule of each judge, derives from the rules followed by Judges A, B, and C in a regular and predictable way. Namely, the court outcome in a given case is the outcome common to a majority of the judges. We can state the rule clearly:

> Regulations of speech are unconstitutional except a) reasonable regulations of all commercial speech and b) content-neutral regulations of the noncommercial speech of juridical entities, if the regulations are supported by legitimate governmental concerns.

If each of our three judges applies her rule consistently, the court will apply its amalgamated rule consistently. The suggestion that multi-member courts cannot act consistently is simply false, even in the context of this worst-case, fractured court.[37]

---

37. Easterbrook's contrary claim, *supra* note 27, relies on social choice theory. *See, e.g.*, K. AR-ROW, SOCIAL CHOICE AND INDIVIDUAL VALUES (2d ed. 1963); A. SEN, COLLECTIVE CHOICE AND SOCIAL WELFARE (1970). In particular, he draws on a theorem first proven by Arrow. This theorem, as applied by Easterbrook to adjudication, imposes various restrictions on the choice problem faced by the panel of judges. Specifically, it requires (1) that they be choosing among at least three alternatives; (2) that each judge have a complete and transitive ordering of the alternatives; and (3) that the aggregation mechanism used by the panel meet five conditions: (a) it must produce a complete and transitive (panel) ranking of all the alternatives; (b) the mechanism must produce such ranking regardless of the pattern of individual rankings of the judges (universal domain); (c) if the ranking of any pair of alternatives is unanimous among the judges, the aggregation mechanism must rank that pair of alternatives the same way (pareto criterion) (termed "unanimity" by Easterbrook, *supra* note 27, at 823); (d) the mechanism cannot identify the panel ranking simply by the ranking of a particular judge (nondictatorship); and (e) in determining the panel ranking of two alternatives A and B, the aggregation mechanism may use only information about the order of A and B in each judge's own ranking (independence of irrelevant alternatives). *Id.* at 823. Given these conditions, Easterbrook argues that differential rankings of alternatives by individual judges on a multi-member panel will result in inconsistent decisions.

While Easterbrook assumes that judges choose among multiple rationales, most of our discussion assumes that judges make a binary choice between judgment for plaintiff and judgment for defendant. As our discussion of the multiple outcome case shows, our disagreement with Easterbrook, however, lies deeper than disagreement over the number of options open to a judge. Rather, we differ from Easterbrook in our demands both on individual judges and on the aggregation mechanism itself to have (or to produce) rankings of all alternatives. In our analysis, each judge must be able to make the set of judgments specified by the aggregation mechanism; she must choose only the outcome she thinks best or correct. Thus each judge need not have a complete, transitive ordering over the alternatives. (This ordering would be an ordering of judgments, such as, for example, "alternative A is more correct than alternative B.") This point is most apparent in our discussion of a choice among three alternatives. *See infra* text accompanying note 41. More importantly, we do not demand that the aggregation mechanism yield a panel ranking of the alternatives. It need only choose one. Easterbrook requires each judge and the panel to have a complete (preference) ordering over all alternatives.

For a number of reasons, we feel that our model of adjudication provides a sounder basis than Easterbrook's for a normative theory of adjudication. First, Easterbrook's model treats the existence of more than two alternatives in a somewhat odd way. While each judge has preferences over all alternatives, she is presented choices pairwise. On Easterbrook's interpretation of alternatives as rationales, such a restriction makes little sense. Judges simply are not asked in case A to choose between rationales I and II, in case B between rationales II and III, and in case C between rationales I and III. All of the rationales are available to the judge in each case; in each instance, she is asked to choose the best outcome. As we show below, the sequence of cases may present her with problems of coherence but not of consistency. Second, the panel has no need to generate a ranking of all possible rationales; it must select only one. Again, as emerges below, although courts might adopt a mechanism that did

So far, we have been considering the case of a court so badly fractured that it cannot produce a majority opinion and thus cannot produce an announced rule to govern future cases and command the allegiance promised by stare decisis. One apparent problem of consistency posed by multi-judge courts involves those cases where there is a majority opinion: in such cases, how can we account for concurring[38] and dissenting opinions? Superficially, such opinions may appear to implicate *concurrent* inconsistency, since the judges who author them necessarily favor rules different than those adopted as law by the majority.

But, of course, a concurring or dissenting opinion does not actually install any rule in the legal system; it is merely a judge's announcement of how she would decide the case were she the only judge or were she able to convince a majority of her colleagues. Accordingly, the problems of consistency involved in authoring such an opinion are in no sense particular to the multi-judge court context.

A single judge, facing a novel question, has no problem of consistency, and neither does a judge on a multi-judge court who disagrees with her colleagues; a majority of them agree on rule A, she would choose rule B, and her opinion merely recites this circumstance and advances her reasons for disagreeing with the majority rule. A single judge who disagrees with an established rule does have a problem of *consecutive inconsistency*. She may regard the rule as intrinsically bad, or as a source of incoherence, as did Skelly Wright in *Javins*; in either event, she can alter the rule only at

produce a complete ranking of alternatives, it seems silly to do so when such a choice has no apparent benefits and plunges one into the antinomies of Arrow's theorem.

Two other remarks are appropriate. First, our analysis, just as Arrow's and Easterbrook's (except for a few comments, *see* Easterbrook, *supra* note 27, at 821-23), ignores strategic behavior on the part of judges. As it stands, therefore, our analysis is more appropriate as the basis for a normative, rather than a positive, theory of adjudication. Most jurisprudential theories of adjudication are, of course, normative. In addition, even for a positive theory, an analysis of sincere adjudication must normally precede one that addresses the additional complexities posed by sophisticated judges acting strategically.

Much of the plausibility of Easterbrook's approach derives from its grounding in preference aggregation. In that context, we desire that the panel ranking "authentically" reflect the rankings of the judges. In judgment aggregation, however, we seek not a reflection of the judges' rankings, but the correct judgment. Consequently, many of the conditions imposed on the aggregation mechanism seem inappropriate in the judgment context. If one judge were infallible, then dictatorship (by the infallible) would be a desirable rather than an undesirable characteristic of an aggregation mechanism. Much the same conclusion might apply if the competence of each judge were known and if one judge were significantly more competent than the others. Similarly, the requirement of independence of irrelevant alternatives seems inapplicable. Each judge makes her judgment based on the relevant criteria for making the judgment, not based on how alternative A compares to alternative B.

Second, none of the above argues that Arrow's theorem or social choice theory in general is "wrong" or "false." We contend, rather, that it does not provide the appropriate model in which to analyze judicial decisionmaking.

38. The problem can only be understood to involve "true" concurring opinions, where judges disagree with the majority, not "two cents" concurring opinions, where judges agree with the majority but add observations of their own.

the price of consecutive inconsistency. Likewise, a judge on a multi-judge court can disagree with a previously announced rule only at the same price. But the problem is in no sense specific to multi-judge courts or to concurring and dissenting opinions. It is merely an artifact of institutional arrangements that respond seriously to the demands of consistency but on occasion neglect those demands in the service of other adjudicatory virtues.

As we have seen, a multi-judge court can behave *consistently*. Even our badly fractured court will be consistent if each of our three judges acts consistently. It does not follow, however, that if our three disagreeing judges act *coherently* the court will also be coherent. There is no reason to infer from the fact that each judge's rule is legitimately derivable from a small set of general, consistent, and unitary premises that the court's amalgamated rule is similarly derivable. It is possible, for example, that the axiom set from which each judge's rule derives rests on a different conception of the appropriate basis for decision than the rules of the other judges.

The extent to which incoherence will plague a multi-member court will depend in part on how each judge conceives her role. The most serious problem will arise if no judge believes that she has a duty to accommodate her judgment in a case before her to prior decisions of the court which she believes were wrongly decided. In this circumstance, each judge would hold a coherent view of what the law should be but would feel no obligation to render coherent the actual series of legal outcomes. Both Judge A and Judge B in the First Amendment example above adhered to their views of the First Amendment even though, by the time the referenda case arose, both Judge A and Judge B would have viewed the prevailing law as incorrect. They strove for the correct outcome in the case before them and ignored any demand for coherence in the system of prevailing legal rules. Although judges do not completely ignore the demands of coherence in the legal system, they may frequently adhere for a time to a view of the law at odds with recent decisions and only later strive for coherence. But our example shows that even a brief (three case) period of obstinacy may lead to an incoherent body of law.[39]

---

39. The death penalty might exemplify the phenomenon described in the text. Both Justice Marshall and Justice Brennan have announced that they will not follow the rule of stare decisis. In Gregg v. Georgia, 428 U.S. 153 (1976), Justice Brennan and Justice Marshall announced their view that the Eighth and Fourteenth Amendments prohibit the death penalty. *Id.* at 227 (Brennan, J., dissenting); *id.* at 231 (Marshall, J., dissenting). A computer search found 442 subsequent Supreme Court cases containing the phrase "adhering to our views that the death penalty is in all circumstances cruel and unusual punishment prohibited by the eighth and fourteenth amendments." (LEXIS, Genfed library, US file, Sept. 9, 1986). From this we may infer not only that they will not seek to decide a new case in a manner consistent with the court's prior decision of an identical case, but also that they feel no obligation to generate a coherent extension of the evolving Supreme Court jurisprudence on the death penalty. Put differently, they renounce any obligation to decide "similar" cases "similarly." To the

Even in this worst case for coherence, some forces reduce the difficulties caused by multi-judge courts. If, as we have hypothesized earlier, judges as a group do in fact "get it right" more often than not, then we can expect judges to tend toward the same or similar judgments. When majorities agree as to both outcome and rule, this rule will become the source of claims of consistency on judges, as enshrined in the rule of stare decisis.

Consider now the case in which each judge believes coherence in the system of legal rules is a necessary condition of getting the right answer in the case before the court. We must examine two variations. In the first, all judges share common criteria for coherence. Thus, if they are shown a series of legal rules, they will agree in their judgments of the coherence of each system of legal rules though they may disagree about which system is the correct one. Suppose now that the sequence of decisions made by the court is coherent. Then, if the court decides cases one at a time, the evolving legal system will continue to be coherent. This result follows because each judge believes she must decide a case in the way that maintains the coherence of the legal system. Even if, at some point in the decision-making process, the judges disagree radically on how the legal system should develop, for each particular case they will each offer a coherent extension of the line of decisions. And the line of decisions will be extended for the purposes of that case consistently with the proffered extension of a majority of the judges. Moreover, even the judges who believe that the decision is contrary to their conception of the way the law should develop will view the resulting pattern of decisions as coherent. Thus, when judges strive to maintain the coherence of the legal system and share criteria of coherence and, in addition, the law develops in a case by case fashion, multi-judge courts will yield coherent outcomes, even if the same judges facing the relevant cases in isolation and for the first time would disagree on the appropriate outcomes.

Suppose, then, that the judges do not share the same criteria of coherence so that, while each strives to maintain her own conception of coherence in the legal system, they may disagree over the coherence of the resulting outcomes. Suppose that up to some point the judges agree that the pattern of decisions they have made thus far is coherent. They now confront three cases, for example, the three First Amendment cases discussed above. In the first case on false advertising, each judge decides according to her own view of what is correct, which now includes the consideration

---

extent they occasionally attract a majority for their position, they introduce incoherence to the law. Similarly, Justice Rehnquist has adhered to his views on procedural due process in the employment context, *see* Arnett v. Kennedy, 416 U.S. 134, 152-54 (1974) (plurality opinion), despite the fact that six Justices have repudiated those views. *See* Cleveland Bd. of Educ. v. Loudermill, 105 S. Ct. 1487, 1492 (1985) (noting that six Justices specifically rejected plurality opinion in *Arnett*); *id.* at 1502-04 (Rehnquist, J., dissenting) (adhering to his plurality opinion in *Arnett*).

that the legal system as a whole should be as coherent as possible. The court upholds the law but now Judge B believes the legal system is incoherent. When the second case on public places arises, Judges A and C will decide it in a way that each believes will maintain the coherence of the system as a whole, while Judge B will try to lessen the incoherence generated by the false advertising decision.[40] After the second case is decided in accordance with Judge B's and Judge C's view, Judge A will believe the law incoherent, as may Judge B, though Judge B will regard the law as less incoherent after the decision than before. When the referenda case arises, only Judge C renders a decision that necessarily preserves coherence as she sees it. Judges A and B may be forced to decide in a way that lessens as much as possible the incoherence they perceive. But after the decision of the referenda case where C is in the minority, Judge C may also view the legal system as incoherent. At this point, each judge might then view the legal system as incoherent, even though each has striven to preserve its coherence. Thus, if radically fractured courts disagree not only on the correct answer but also on the criteria for coherence, multi-judge courts may result in incoherent rules even when each judge strives to preserve coherence or lessen incoherence. Since the judges disagree on the criteria for coherence, when a decision runs contrary to a judge's view of the law, it may result in an incoherent continuation of the line of decisions. If the judges disagree broadly on the criteria of coherence, incoherence is likely to arise more frequently and subside less with the passage of time.

Thus far our discussion of consistency and coherence has assumed that every case has only two outcomes. A brief discussion of two related examples suggests that, even if cases have more than two possible outcomes, multi-member courts, when they reach any decision at all, will be consistent. These two examples also reveal the complexity of the concept of coherence.

Suppose first that a given case has three possible outcomes: the defendant might win (W), the plaintiff might get remedy 1 (R1), or he might get remedy 2 (R2). The multiplicity of possible outcomes, of course, does not alter the ability of a single judge to achieve consistent (or coherent) patterns of decisions. Consider then a three judge court with Judge A espous-

---

40. For Judge B, the table may no longer represent a true picture of how she would decide cases. In this variation of the problem, Judge B attempts to preserve coherence; thus her view of the correct decision of the public places and referenda cases depends upon the decision in the false advertising case. We have assumed that considerations of preserving coherence have not led any judge, when faced with an incoherent pattern, to alter her view of the correct decision.

It should be noted that if more than two outcomes are possible, one can more plausibly introduce incoherence into the system because a greater range of outcomes increases the scope for disagreement among judges over which outcomes constitute coherent continuations of the line of cases.

ing a rule that dictates outcome W, Judge B a rule that dictates R1, and Judge C one that dictates R2. Unlike the case in which only two outcomes are possible, the judges here cannot agree on a judgment. The panel decision will depend upon the default rule. For instance, one default rule (DR1) would be that, in the absence of a majority agreement on any outcome, defendant wins. DR1 dictates an outcome of W. Another default rule (DR2) might divide the decision into parts. First, the panel must decide whether plaintiff is entitled to any recovery (i.e., whether he has a right). Only then does the panel determine the remedy to which he is entitled. Under DR2, the outcome will depend on how Judge A votes on the remedy question. For DR2 to be meaningful, Judge A must be able to answer the "counterfactual": Given that plaintiff is entitled to recover, what (in my judgment) ought he to recover in this case? Under either DR1 or DR2, the panel will be consistent.[41]

We have suggested two different judicial practices to resolve the indecision in the multiple outcome case. These two practices do not exhaust the possibilities. For instance, judges might consider their rankings of the three different outcomes and base their decision solely on the pairwise comparisons embedded in this ordinal ranking. Of course, this path leads us to Arrow's paradox, and it is not clear why we should follow such a path. The practice chosen to resolve indecision should advance the purposes of the general practice of adjudication. Selecting a practice that leads to Arrow's paradox undermines rather than advances those purposes. In judgment aggregation, the court seeks to identify the "correct" result, given common criteria of correctness. The procedures embodied in DR1 and DR2 may be more reasonable ways of pursuing this goal or subsidiary goals of adjudication than the use of rankings of outcomes judges believe to be wrong as the basis for selecting the "correct" answer. We might, for example, justify DR2 by arguing that the question of whether defendant is legally responsible is qualitatively different from and more important than the question of appropriate remedy.

Our second example assumes multiple issues rather than multiple remedies.[42] Consider a criminal appeal in which the defendant urges that the conviction be reversed on two grounds: (1) that his confession was coerced and hence should not have been admitted at trial; and (2) that the jury

---

41. In this discussion, as throughout the Article, we suppress analysis of strategic voting and vote trading on the part of the judges. One might imagine that Judge B and Judge C would agree on a judgment that gave some remedy to plaintiff. They might choose a compromise remedy, or one judge might accept the position of the other judge in exchange for a concession in another case. We think that, while strategic voting and vote trading may well occur, they are at odds with a judgment aggregation view of adjudication.

42. The second example closely parallels the first. Each ground suggests the appropriate remedy. The two grounds thus correspond to R1 and R2 while upholding the conviction corresponds to W.

was improperly selected (some discrimination claim). Again consider a three judge panel:

ISSUE:	1	2
JUDGE:		
A	REVERSE	AFFIRM
B	AFFIRM	REVERSE
C	AFFIRM	AFFIRM
PANEL:	AFFIRM	AFFIRM

The panel arrives at a paradoxical outcome if it decides the case by voting on an issue-by-issue basis.[43] For each issue, two judges believe the defendant was treated fairly. But two judges believe that in this case, the defendant was treated unfairly.[44] The extent to which the case's outcome strikes one as unfair depends on how one conceives of the process of adjudication. The result strikes us as unfair because both Judge A and Judge B would, if deciding this case alone, reverse. Both judges view the trial as defective; they simply cannot agree on a rationale, a disagreement which we may not think sufficient to justify the defendant's incarceration.

On the other hand, agreement on the rationale for reversal may promote similar treatment of similar cases and thereby promote coherence. To identify issues requires that we decompose our conception of fair trial procedures into simple components. Within each component we may have a unitary theory of decision that provides coherence to that subset of cases. Our understanding, for example, of the evils of coerced confessions allows us to articulate criteria for identification of them, and to relate these criteria both to the evils of coerced confessions and to more general conceptions of a fair trial. A similar understanding of jury discrimination motivates us to isolate it as an "issue." That Judges A and B disagree on the source of unfairness to the defendant threatens to undermine the conceptual unity within each issue. One may then prefer issue-by-issue decisionmaking because it advances coherence.

VI. CONCLUSION

A complete theory of adjudication must account for a complex process of decisionmaking. Theories of adjudication often begin—and end—with a single judge, Liza, considered as though she has decided and will in the future decide all the cases in the legal system. Complications immediately

---

43. Issue-by-issue voting contradicts our description of how judges decide cases. We began by associating outcomes with fact situations and not with "issues" or subsets of fact situations.

44. Both Judges A and B might view the outcome in this case as introducing incoherence into the legal system.

and necessarily ensue when the theory admits a chain of judges stretching backward and forward in time, with Liza merely one temporal link; the difficulties increase when the theory recognizes the existence of other solitary judges deciding cases concurrently with Liza. Even then, the picture is crucially incomplete: the reality of judicial institutions demands that a theory of adjudication account for the fact that judges sit in panels and decide cases as groups. This last aspect of the adjudicatory enterprise generally has been ignored by theorists; in this Article we have sketched several ways in which attention to the group aspect of judicial decisionmaking both deepens and confounds a theory of adjudication.

Our analysis begins with a division of group decisionmaking into exercises of judgment aggregation, preference aggregation, and representation, and with the identification of the measures of performance—the virtues—that appropriately attach to each. On our reading, judgment aggregation aims at right answers (accuracy), preference aggregation at genuine expressions of personal choice (authenticity), and representation at both effective replication of the decisions that the represented group would reach (fit) and a stable centrism of decisionmaking (reliability). In addition, appearance emerges as a free-floating virtue that can attach to any of the three models of group decisionmaking.

In considering how increasing the number of judges could improve the performance of a court, we concern ourselves primarily with adjudication as an exercise in judgment aggregation and hence consider judicial performance in terms of accuracy. We do so because we understand that prominent, competitive views of the judicial decisionmaking process—legalism, positivism, and legal realism—unite in demanding or at least encouraging the judgment aggregation view of adjudication, and because we share that view. Moreover, the link between court size and accuracy seems particularly interesting, especially when accuracy is understood to embrace the qualities of consistency and coherence.

Our analysis of the relationship between court size and accuracy is not definitive, but it does indicate that adding judges improves accuracy under plausibly optimistic assumptions about the general capacity of judges to reach correct outcomes and about the impact of deliberation on this capacity.

Of particular interest to our accuracy inquiry is whether groups of judges functioning as a single court can act consistently and coherently, or whether their group status undermines these common adjudicatory objectives. We demonstrate that, if individual judges decide cases consistently, then the court which they constitute will likewise reach consistent results. But panels of judges, each judge of which decides cases coherently, will not necessarily reach coherent results. To achieve coherence, a panel must

satisfy two additional conditions: Each judge must offer a decision she believes coherent with the prior decisions of the court (as opposed to her own view in prior cases), and the panel must have a common conception of coherence. The condition that the panel share a conception of coherence is deeply problematic for theories of adjudication that seem to tolerate or even encourage conceptual disparity among judges. Our claim in this regard is an important example of the need to attend to the group aspect of judicial decisionmaking.

In the end, we leave groups of judges where we found them, in the very center of the appellate adjudicatory process. The only surviving mystery is how legal theorists can contrive to overlook them.