OWEN M. FISS

Groups and the Equal Protection Clause

This is an essay about the structure and limitations of the anti-discrimination principle, the principle that controls the interpretation of the Equal Protection Clause. To understand the importance of that principle in constitutional adjudication a distinction must first be drawn between two different modes of interpretation.

Under one mode the constitutional text is taken pure–the primary decisional touchstone is the actual language of the Constitution. The text of the Constitution is viewed as providing an intelligible rule of decision and that text, rather than any gloss, is the primary referent; at most, disagreement may arise as to how much weight should be given to one or two words and what the words mean. This is a plausible –arguable, though far from persuasive–approach to the Free Speech Clause. It is the approach associated with Justice Black.

The second mode of constitutional interpretation deemphasizes the text. Primary reliance is instead placed on a set of principles–which I call *mediating* because they "stand between" the courts and the Constitution–to give meaning and content to an ideal embodied in the text. These principles are offered as a paraphrase of the

The original version of this paper was presented at a conference held in June 1975 at the Institute for Advanced Study, Princeton, N.J., and sponsored by the Alfred P. Sloan Foundation. The paper also benefited from a discussion at the Society for Ethical and Legal Philosophy in October 1975, and from criticism from students in two seminars of mine at the Yale Law School, one in the spring of 1975 and the other (taught with my colleague Bruce Ackerman) in the fall of 1975.

particular textual provision, but in truth the relationship is much more fundamental. They give the provision its only meaning as a guide for decision. So much so, that over time one often loses sight of the artificial status of these principles—they are not "part of" the Constitution, but instead only a judicial gloss, open to revaluation and redefinition in a way that the text of the Constitution is not.

The Equal Protection Clause has generally been viewed in this second way. The words—no state shall "deny to any person within its jurisdiction the equal protection of the laws"—do not state an intelligible rule of decision. In that sense the text has no meaning. The Clause contains the word "equal" and thereby gives constitutional status to the ideal of equality, but that ideal is capable of a wide range of meanings. This ambiguity has created the need for a mediating principle, and the one chosen by courts and commentators is the antidiscrimination principle. When asked what the Equal Protection Clause means, an informed lawyer—even one committed to Justice Black's textual approach to the First Amendment—does not repeat the words of the Clause—a denial of equal protection. Instead, he is likely to respond that the Clause prohibits discrimination.

One purpose of this essay is simply to underscore the fact that the antidiscrimination principle is not the Equal Protection Clause, that it is nothing more than a mediating principle. I want to bring to an end the identification of the Clause with the antidiscrimination principle. But I also have larger ambitions. I want to suggest that the antidiscrimination principle embodies a very limited conception of equality, one that is highly individualistic and confined to assessing the rationality of means. I also want to outline another mediating principle—the group-disadvantaging principle—one that has as good, if not better, claim to represent the ideal of equality, one that takes a fuller account of social reality, and one that more clearly focuses the issues that must be decided in equal protection cases.

I. The Structure of the Antidiscrimination Principle

The construction of the antidiscrimination principle proceeds in three steps. The first is to reduce the ideal of equality to the principle of equal treatment—similar things should be treated similarly. The

second step is to take account of the fact that even the just state must make distinctions, must treat some things differently from others; for example, even the most noncontroversial criminal statute distinguishes between people on the basis of their conduct. Recognition of the inevitability and indeed the justice of some line-drawing makes the central task of equal protection theory one of determining which lines or distinctions are permissible. Not all discriminations can be prohibited; the word "to discriminate," once divested of its emotional connotation, simply means to distinguish or to draw a line. The mediating principle of the Equal Protection Clause therefore must be one that prohibits only "arbitrary"[1] discrimination. The Clause does not itself tell us which distinctions are arbitrary, and as the third step in this process a general method is posited for determining the rationality and thus the permissibility of the lines drawn. The method chosen by the Supreme Court, and the one that generally goes under the rubric of the antidiscrimination principle, has two facets: (a) the identity of the discrimination is determined by the criterion upon which it is based, and (b) the discrimination is arbitrary if the criterion upon which it is based is unrelated to the state purpose.

To illustrate this method of determining whether a discrimination is arbitrary, let us suppose the state wishes to pick the best employees or students for a limited number of openings. That process inevitably involves choices. The state must discriminate. Assume also that the choice is made on the basis of performance on a written test designed to pick the most productive workers or the most brilliant students. The state would then be making an academic discrimination. Presumably it would not be arbitrary since the criterion is related to the state

1. Sometimes the word "invidious" is used interchangeably with "arbitrary" to describe the universe of impermissible discriminations, though with little attention to the special connotations of the word "invidious"—"tending to cause ill will, animosity, or resentment." Professor Karst, in a valuable article, reveals a sensitivity to the difference between the two terms. "Invidious Discrimination: Justice Douglas and the Return of the 'Natural-Law–Due Process Formula'" 16 *U.C.L.A. L. Rev.* 716, 732–734 (1969). He, however, uses the term "invidious discrimination" in a conclusory sense, devoid of descriptive meaning. The term is used "to describe the Court's ultimate conclusion on the question of a violation of equal protection." Ibid., p. 740 fn. 110.

purpose. This would be true even if it turned out that the only applicants selected happened to be white. But suppose the criterion for selection is color: the state grants the position to whites and denies it to blacks, on the basis of their color. That would make the discrimination a racial one and arbitrary because the criterion is not related to the state purpose of selecting the most brilliant students or most productive employees.

In this example, the racial criterion has been deemed arbitrary because it is not related to the state purpose. But Tussman and tenBroek, in their now classic article of the late 1940s,[2] pointed out that un-

2. "The Equal Protection of the Laws," 37 *Calif. L. Rev.* 341 (1949) (hereafter cited as Tussman and tenBroek). Tussman and tenBroek saw three principles, not one, governing the application of the Equal Protection Clause. The one I am describing under the rubric of the antidiscrimination principle was called the "reasonable classification" principle. They also spoke of a principle opposing "discriminatory legislation" and a third guarantee, one of "substantive equal protection." By the latter they meant that certain "rights" (analogous to those that were previously protected by the doctrine of "substantive due process") were to be protected by the Equal Protection Clause; these rights could not be interfered with even though the interference was even-handed. They sought to explain *Shelley v. Kraemer* on the basis of this principle—the enforcement of a racial restrictive covenant was not a form of unequal treatment, but rather an interference with the "right" of a willing seller to sell to a willing buyer. They sought to justify the use of the Equal Protection Clause as "a sanctuary" for these "rights," not because they have any connection to equality, but because, in their words, the Clause "was placed in our Constitution as the culmination of the greatest humanitarian movement in our history." Ibid., p. 364. This doctrine has received little formal recognition by the Court in the past twenty-five years. At most, strands of this doctrine are reflected in the fundamental-right trigger of the strict scrutiny branch of the antidiscrimination principle; in that instance the "right" is used to determine the appropriate degree of fit. On the other hand, what Tussman and tenBroek referred to as the ban on "discriminatory legislation," is completely integrated within what I call the antidiscrimination principle. For Tussman and tenBroek the ban on "discriminatory legislation" was a "criticism of legislative purpose," a "demand for purity of motive." Ibid., pp. 357, 358. Certain legislative purposes, such as the subordination of blacks, were denied to the state altogether, and thus it was irrelevant that the fit might have been perfect between the criterion (or classification) and the (forbidden) purpose. They were not especially clear as to which purposes were forbidden—they spoke in terms of "bias," "prejudice," "hostility," and "antagonism." They also recognized that the word "discrimination" could also be used in the sense that it is being used here (that is, as a term to describe the reasonable classification doctrine), but they failed to integrate the two senses of the word. Ibid., p. 358 fn. 35.

relatedness is not a dichotomous quality. In most cases it is not a question of whether the criterion and end are related or unrelated, but a question of how well they are related. A criterion may be deemed arbitrary even if it is related to the purpose, but only poorly so. Tussman and tenBroek explained that, given the purpose, a criterion could be ill-suited in two different ways: it could be over-inclusive (it picked out more persons than it should) or underinclusive (it excluded persons that it should not). These evils can be described, to use the jargon of contemporary commentators, as ill-fit.[3]

This is the core idea—the foundational concept—of the antidiscrimination principle, one of means-end rationality. But it must also be recognized that the principle contemplates a series of additional inquiries that yield a superstructure. First, the principle requires that the court identify the underlying criterion. This means that a distinction must be drawn between the stated criterion and the *real* criterion. If the challenge is to a statute in all its applications, then the stated criterion may be taken at face value. But if the challenge is to the statute as applied, or to administrative action, then there is no reason why the stated criterion should be treated as the real criterion. The administrator may say he is selecting students on the basis of academic performance, when in fact he is ignoring their test scores, and making his decision on the basis of race.

Second, the court must identify the state's purpose and determine whether it is legitimate. For example, suppose the state's purpose is to subordinate blacks rather than to choose the best students or employees. Then the color black would be well-suited for determining

3. Tussman and tenBroek did not use the term "ill-fit." They did, however, make the points about over- and underinclusiveness (words they actually used and introduced into legal discourse) by the use of diagrams: *P* is the universe of person that *should* be selected given the purpose, and *C* is the universe of person *actually* selected by the criterion. The diagram where *P* is a subset of *C* is used by them to represent overinclusiveness and the diagram where *C* is a subset of *P* is used to represent underinclusiveness. The term "fit" is suggested by these diagrams, and perhaps for that reason is used by the contemporary commentators, such as, Ely, "The Constitutionality of Reverse Racial Discrimination," 41 *U. Chi. L. Rev.* 723, 727 fn. 26 (1974) (hereafter cited as Ely), and is now part of ordinary constitutional parlance.

who should be excluded from the state colleges or jobs, and under a test consisting exclusively of means-end rationality, this use of the racial criterion would be permissible. The Equal Protection Clause would thereby be transformed into a minor protection against state carelessness, permitting intervention only when it was plain that the state did not know how best to achieve its ends.[4] Accordingly, it seemed necessary to go beyond the concept of ill-fit, and the anti-discrimination principle has been modified so as to require that the state purpose against which the criterion is to be measured be legitimate (or permissible).

This account of the inquiry into purpose suggests two steps: first, identifying the state purpose and second, determining whether the purpose is legitimate. But if the court need not take the state's professed purpose as that against which the criterion is to be measured, then the two steps collapse into one. The court fixes the state's purpose by the process of imagination: only legitimate purposes would be imagined, and the judge's mind would scan the universe of legitimate purposes until he identified the legitimate state purpose that was best served by the criterion, the one that left the smallest margins of over- and underinclusiveness. The universe of imaginable purposes would not contain those purposes disavowed by the state, and the disavowal could occur implicitly, for example, it could be implied by the overall statutory framework of the state.[5]

Some have argued that the criterion should be measured against the *stated* purpose. Such a restriction might enhance the invalidating power of the Equal Protection Clause, for it was always assumed—perhaps out of simple fairness—that the process of imagination would yield the best purpose, the one most favorable to the state. This restriction would also reduce judicial maneuverability. And some have further hypothesized that a stated-purpose requirement would invigorate the political process—for "it would encourage the airing and critique of those reasons [justifying the legislative means] in

4. See generally, Note, "Legislative Purpose, Rationality, and Equal Protection," 82 *Yale L. J.* 123 (1972).

5. See, for example, Eisenstadt v. Baird, 405 U.S. 438 (1972).

the state's political process."[6] This hypothesis seems to me to posit a somewhat naive conception of the state political process and what might invigorate it. But more importantly, the restriction is inconsistent with judicial practice in other areas (such as determining whether legislation is authorized by the enumerated powers), and it would be hard to apply. The state rarely identifies its purpose with any degree of precision, and the restriction would be virtually meaningless if, as one proponent of this idea has suggested, "A state court's or attorney general office's description of purpose should be acceptable."[7] For these reasons, the stated-purpose requirement has not taken root, and probably should not be viewed as an important or permanent feature of the antidiscrimination principle.

A third set of auxiliary concepts is responsive to two facts—that the critical inquiry of ill-suitedness, as modulated by Tussman and tenBroek, is one of degree and that some margin of over- and underinclusiveness can always be discovered. Standards must therefore be set for determining how poor the relationship must be between criterion and purpose before it is deemed arbitrary—or to use the jargon of the contemporary commentators, how tight a fit must there be? The doctrines of "suspect classification" and "fundamental right" seek to answer this question. They are essentially standards for determining the requisite degree of fit. In contrast to the more permissive standard called "mere rational relationship" or "minimum scrutiny," which tolerates broad margins of over- and underinclusiveness, these doctrines trigger a strict scrutiny, one that demands a tight fit. If the criterion is "suspect" (the exemplar being race) or the "right" affected "fundamental" (the exemplar being the right to vote)[8] there has to be a very tight fit—any degree of avoidable over-

6. Gunther, "Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection," 86 *Harv. L. Rev.* 1, 47 (1972) (hereafter cited as Gunther). See also Greenawalt, "Judicial Scrutiny of 'Benign' Racial Preference in Law School Admissions," 75 *Colum. L. Rev.* 559, 600 (1975) (hereafter cited as Greenawalt).

7. Gunther, p. 47.

8. See, for example, Carrington v. Rash, 380 U.S. 89 (1965) (denying the right to vote to those serving in the armed forces). As originally conceived, the

inclusiveness or underinclusiveness would be deemed "too much."[9] The use of the term "avoidable" gives the state an out: ill-fit would be accepted if there is no closer-fitting way of satisfying its purpose. Assuming a legitimate purpose and no better alternative, the discriminatory criterion, however "suspect," would seem necessary, and so acceptable under the Equal Protection Clause.

The strict-scrutiny branch of the antidiscrimination principle has necessitated the establishment of methods for determining which criteria are "suspect" and which rights are "fundamental." It has also required, as the fourth feature of the superstructure, the introduction of defensive doctrines—those that allow for the validation of laws and practices that would otherwise seem invalid. There is no analytic reason why these defensive doctrines cannot be applied in the minimum-scrutiny context (and on some occasions they have been); the point is simply that there is less need for them there. In the minimum-scrutiny context, the impulse toward validation could easily be accommodated in the judgment that the margins of over- and underinclusiveness are not "excessive." In the strict-scrutiny

---

"fundamental right" did not have to be of constitutional stature. The constitutional status of the right to vote is shrouded in controversy, but strict scrutiny seems also to have been applied to laws restricting rights clearly of nonconstitutional stature, such as the right to procreate. Skinner v. Oklahoma, 316 U.S. 535 (1952). In 1973, a majority of the Supreme Court said the right had to be of constitutional stature—though in order to pay respects to precedent—they acknowledged that the constitutional right may be an *implicit* one. San Antonio Ind. School District v. Rodriguez, 411 U.S. 1 (1973). It might also be noted that both triggers of strict scrutiny—suspect classification and fundamental right—might be present at the same time (for example, a state denying some ethnic group the right to procreate). The operative significance of this double trigger is not clear. One trigger alone may be sufficient to result in the invalidation of the law, and in that instance the second one would be superfluous—merely frosting on the cake.

9. In these cases it is often difficult to describe the discrimination as "arbitrary" within the ordinary meaning of that word. But one loses sight of the need to make that judgment. The concept of arbitrariness enters only in the establishment of the general method, as a foundational concept, and other factors account for the additional tiers. For an awareness of how the suspect-classification branch of the antidiscrimination principle causes this departure from the rationality test, see Justice Harlan's dissent in Harper v. Virginia Bd. of Elections, 383 U.S. 663, 682 n. 3 (1966).

context, on the other hand, that method of avoidance is not available since, by definition, any margin of over- or underinclusiveness is "excessive."

One defensive doctrine permits the state to take one step at a time. It is a defense aimed at excusing underinclusiveness. For example, assume there is a literacy requirement for voting at time-1. Later, at time-2, the state decides to pass a law establishing that the completion of the sixth grade in a school where subjects are taught primarily in English or Spanish (including schools in Puerto Rico or Mexico) is sufficient proof of literacy. An individual who completed the sixth grade in France (or Poland) complains that the change wrought by the new law is underinclusive–given the purpose of the state, there is no reason why he should not be included. The state might defend on the ground that it is simply taking one step at a time, and the defense has been allowed.[10]

Such a defense obviously has the capacity for completely undermining the complaint of underinclusiveness–each instance of underinclusiveness might be explained as an instance in which the legislature chose to take one step at a time. Tussman and tenBroek, fully aware of this risk, sought to limit the defense by drawing a distinction between the reasons that explained why the state took only one step at a time. If the reach of the law was confined because of *administrative* considerations (for example, additional complications would be introduced) as opposed to *political* considerations (for example, the sponsors could not muster enough votes for extending the law to others), then the one-step-at-a-time defense was allowable. The Supreme Court, on the other hand, has on at least one occasion sought to use the concept of "reform" as the limiting one: the state is allowed to take one step at a time on when the law is a "reform" measure–a law that improves (rather than worsens) the status quo.

Another defensive doctrine requires the court to rank the legitimate

10. See Katzenbach v. Morgan, 384 U.S. 641 (1966) (involving American-flag schools). In Williamson v. Lee Optical Co., 348 U.S. 483, 489 (1955), the Court said, "The reform may take one step at a time, addressing itself to the phase of the problem which seems most acute to the legislative mind."

purposes of the state—to make a distinction between ordinary and special state purposes. Hence, the concept of "a compelling state interest": the achievement of such an interest or purpose is so important that it excuses imperfect means. This doctrine seems to have its roots in the *Japanese Relocation Cases*, where the Supreme Court permitted the use of a racial or national-origin criterion (clearly a suspect one) for determining who should be relocated and otherwise confined.[11] The state purpose—self-preservation of the nation in time of war—was deemed to be of sufficient importance to excuse the overinclusiveness (not all Japanese were security risks) and underinclusiveness (those of German origin might be as much of a security risk).

This appeal to a compelling state interest must be carefully delineated. In the *Japanese Relocation Cases*, the concept was used defensively, to excuse what would otherwise be impermissible, and there was no doubt that the evil to be excused was ill-fit—over- and underinclusiveness. It was not part of a general balancing test. A minor debate has broken out recently, however, as to whether the exclusive focus must remain on ill-fit, or whether there can be, in a case where there is admittedly a perfect fit between means and ends, a "weighing of ends"—a balancing of the harm to the individuals subjected to the law and the good to be achieved.[12] Professor Brest poses the issue in his hypothetical: "How should a court treat a school principal's decision, based solely on aesthetics, to have black and white students sit on opposite sides of the stage at the graduation ceremony?"[13]

11. Korematsu v. United States, 323 U.S. 214 (1944) (relocation); See also Hirabayashi v. United States, 320 U.S. 81 (1943) (curfew).

12. Compare, for example, Ely, p. 727 fn. 26 ("I have argued that, rhetoric to the contrary notwithstanding, special scrutiny in the suspect classification context has in fact consisted not in weighing ends but rather in insisting that the classification in issue fit a constitutionally permissible state goal with greater precision than any available alternative.") with Greenawalt, p. 565 fn. 41 ("It is clear, however, that in some suspect classification cases, the Court has weighed ends, even though it has not been explicit about what it is doing.").

13. *Processes of Constitutional Decisionmaking* (Boston, 1975), p. 489. The oddity of the example is important: it is testimony of how far one must go to find a situation in which complaint of ill-fit could not be made, and thus it reveals the potential reach of the antidiscrimination principle, even as a means focused tool. It is always possible to find ill-fit. I might also add that the

There is little doubt in my mind as to how a court would or should decide the case: the practice is a violation of the Equal Protection Clause. But that is not the issue. The issue is whether it is possible to get to that result (or get there as easily as one should) from the antidiscrimination principle–taken as the mediating principle of the Equal Protection Clause. I think not. If the court finds a state purpose that does not allow the slightest degree of over- or under-inclusiveness, as is indeed suggested by Professor Brest's hypothetical, then the statute or practice would be valid under the antidiscrimination principle. This would be true even though the criterion is "suspect," for example, race. It would not be permissible, within the structure of the antidiscrimination principle, to decide the question by "balancing" or "weighing" the harm done by the state practice (for example, blacks are stigmatized) against the (noncompelling) interests to be served by the state practice (for example, aesthetic satisfaction). The antidiscrimination principle–as I understand it, as Tussman and tenBroek designed it, and as the Supreme Court has generally used it–is a theory about ill-fit, not about the balance of advantage.[14]

---

example is meant only to reveal the structure of the antidiscrimination principle and it can fulfill that purpose even if it is an unlikely case.

14. I realize it is difficult to document this assertion (or perhaps any assertion about a so-called mediating principle). The meaning of the word "discrimination" as gleaned from ordinary usage or the dictionaries is hardly decisive. The word merely requires that a distinction be involved in the analysis–perhaps be the trigger of scrutiny. It does not set the limit or terms of scrutiny. Nor is there an authoritative (official) text to which I can point and say, here is the full and definitive statement of the antidiscrimination principle. Even if there were, one might contend (as those who take opposite sides in the issue do) that the rhetoric is not decisive–what is important is not what the courts say, but what they do. But let me say by way of defense that it is important to be clear about what is at stake–I am only trying to construct a prototype for purposes of analysis and exposition. I will ultimately contend that this prototype is too limited, and seek to supplement it. Of course, it might be contended that my conception of the antidiscrimination principle is too narrow and that the supplemental principle could be viewed as only a slight modification of what I call the antidiscrimination principle, for example, that it might be called "antidiscrimination principle 2." The force of that contention depends on the degree of resemblance–whether the supplemental principle is a close relative of the antidiscrimination principle or rather the member of a new

II. The Appeal of the Antidiscrimination Principle

Antidiscrimination has been the predominant interpretation of the Equal Protection Clause. The examples I have given are cast primarily in terms of race, but the principle also controls cases that do not involve race. It is the general interpretation of the Equal Protection Clause; and indeed it is viewed as having preemptive effect—if the state statute or practice passes the means-end test, then it does not violate the Clause. There have been exceptions, but they have been criticized precisely because they were departures from the principle. It was the substantive character of the one-man, one-vote standard of *Reynolds v. Sims*[15] that prompted Justice Harlan's strong dissent. The standard criticism of that decision invokes the antidiscrimination principle, which would have allowed distinctions among voters if, for example, those differences in treatment were related to legitimate interests, such as preserving the integrity of government subdivisions.

Why, it might fairly be asked, has the antidiscrimination principle been given this position of preeminence? The Equal Protection Clause may need some mediating principle, but why this one? An answer couched in terms of text and history does not suffice. The antidiscrimination principle is not compelled or even suggested by the language of the Clause. That language stands in sharp contrast to that of the Fifteenth Amendment, which does speak in terms of discrimination—the right to vote shall not be denied on account of race. Nor is the antidiscrimination interpretation securely rooted in the legislative history of the Clause. The debates preceding the adoption of the Equal Protection Clause, as best I have been able to deter-

---

family altogether. To decide that issue it is important to understand the intellectual roots of what might be deemed the primary version of the antidiscrimination principle, and there is no dispute that the version focusing on means—the one that conceives of the evil as ill-fit and does not weigh ends—is the primary one. At the very most commentators such as Brest or Greenawalt argue that the principle should be extended far enough to embrace a weighing of ends, and concede that it is predominantly expressive of a conception of means-end rationality.

15. 377 U.S. 533, 615–624 (1964). See also Justice Harlan's dissent in Harper v. Virginia Bd. of Elections, 383 U.S. 663, 682 (1966).

mine, do not justify this choice.[16] Nor is there any reason to believe that antidiscrimination was chosen by the Court as *the* interpretation because of some special view of the legislative history. Yet, even with history and text aside, it is important to note that the predominance of the antidiscrimination principle can in large part be traced to considerations that are particularly appealing to a court.

First, the antidiscrimination principle embodies a conception of equality that roughly corresponds to the conception of equality governing the judicial process. When we speak of "equal justice" we have in mind a norm prohibiting the adjudicator from taking into account certain irrelevant characteristics of the litigants—their race, wealth, and so on. That is the message conveyed by the blindfold on the icon of justice. The antidiscrimination principle also invokes the metaphor of blindness—as in "color blindness."[17] The overarching

16. The materials which have been examined, including the debates on the Fourteenth Amendment itself, are not revealing, and obviously not decisive on an intended mediating principle. See, for example, H. Flack, *The Adoption of the Fourteenth Amendment* (Gloucester, Mass., 1908); Frank & Munro, "The Original Understanding of 'Equal Protection of the Laws,'" 50 *Colum. L. Rev.* 131 (1950); J. James, *The Framing of the Fourteenth Amendment* (1956); J. tenBroek, *Equal Under Law* (New York, 1965). In all fairness I should report that tenBroek's research led him to conclude that "equal" was of secondary importance to "protection" in the Fourteenth Amendment. He wrote, "It was because the protection of the laws was denied to some men that the word 'equal' was used. The word 'full' would have done as well." tenBroek, p. 237. I have not examined the history of all the civil rights debates during the Thirty-ninth Congress. As secondary sources on these debates, see C. Fairman, *Oliver Wendell Holmes Devise History of The Supreme Court, Volume VI, Reconstruction and Reunion 1864–1888*, part one (New York, 1971), pp. 1117–1300; Bickel, "The Original Understanding and the Segregation Decision," 69 *Harv. L. Rev.* 1 (1955); and Casper, "Jones v. Mayer: Clio, Bemused and Confused Muse," 1968 *Sup. Ct. Rev.* 89. But the provisions of the Civil Rights Act of 1866 which are still with us, 42 U.S.C. 1981 and 1982, do not speak in terms of discrimination. Instead, they say all persons "shall have the same right . . . as is enjoyed by white citizens. . . ."

17. This metaphor first surfaced in Justice Harlan's dissent in Plessy v. Ferguson, 163 U.S. 537, 559 (1896) ("our Constitution is color-blind"). The metaphor was suggested by the attorney for the blacks, Albion W. Tourgee, and he—clever lawyer that he must have been—understood why this metaphor would be appealing to a judge. Page 19 of Tourgee's brief in *Plessy* reads: "Justice is pictured blind and her daughter, the Law, ought at least to be color-blind."

obligation is to treat similar persons similarly, declaring certain individual characteristics—such as color—irrelevant.

It is natural for the Justices to seize upon the ideal of their craft in setting norms to govern others. Their craft sets limits to their horizons, it influences their choice among the many meanings of equality. This limit on vision may have been reinforced by the fact that some of the early equal protection cases challenged exclusionary conduct occurring in the course of the judicial process.[18] Moreover, the words "protection of the laws" in the Clause may have led the Justices to think primarily of the administration of justice, and the concept of equality that governs judicial activity in general (equal justice). At some point in history the word "equal" shifts its location so as to deemphasize the word "protection"—it becomes understood that the Clause guarantees "the protection of equal laws," rather than just the "equal protection of the laws"; but the implications of the original version still linger.[19]

Second, the antidiscrimination principle seems to further another supposed norm of the craft—value neutrality—that the judges not substitute their preferences for those of the people. The antidiscrimination principle seems to respond to an aspiration for a "mechanical jurisprudence"—to use Roscoe Pound's phrase—by making the predicate of intervention appear technocratic. The antidiscrimination principle seems to ask no more of the judiciary than that it engage in what might at first seem to be the near mathematical task of determining whether there is, in Tussman and tenBroek's terms, "overinclusiveness" or "underinclusiveness," or, in the terms of the contemporary commentators, whether there is the right "fit" between means and ends. The terms used have an attractively quantitative ring. They make the task of judicial judgment appear to involve as little discretion as when a salesman advises a customer whether a pair of shoes fit. Moreover, under the antidiscrimination principle, whatever judgment there is would seem to be one about means, not

18. For example, Strauder v. West Virginia, 100 U.S. 303 (1880); Virginia v. Rives, 100 U.S. 313 (1880) (both involve the exclusion of blacks from the jury).

19. Tussman and tenBroek, p. 342.

ends, thereby insulating judges from the charge of substituting their judgments for that of the legislature. The court could invalidate state action without passing on the merit or importance of the end—a task, it might be argued, that is especially committed to the more representative branches of government.[20]

The belief that the countermajoritarian objection to judicial review can be avoided by a "mechanical jurisprudence" is false. The entitlement of the judiciary to intervene is no less controversial because only the means are being attacked. They too have been chosen by the people. And there is, in any event, nothing mechanical about the antidiscrimination principle. The promise of value neutrality is only an illusion. On the explicit level, the court must determine whether the state end is legitimate, which classifications are suspect, which rights are fundamental, which legitimate state interests are compelling, and whether the occasion is a proper one for invoking the one-step-at-a-time defense. On the implicit level, the preferences of the judge enter the judicial process when he formulates the imaginable state purposes and chooses among them, and also when he decides whether the criterion is sufficiently ill-suited to warrant invalidation—whether the right degree of fit is present. In contrast to the case of shoes, the concept of fit here has no quantitative content. It only sounds quantitative—as do the words "how much" when used to describe the intensity of affection.

A third explanation for the predominance of the antidiscrimination principle may be found in another supposed ideal of the law—objectivity. In this instance the aspiration is for rules with three characteristics: (a) the rules can be stated with some sharpness or certainty; (b) they are not heavily dependent on factual inquiries or judgments of degree; and (c) they are not time-bound. Rules of this sort are thought to be more "manageable"[21] and to conform to some abstract view about the necessary attributes of "legal rules"—a view

20. See Gunther, pp. 21, 23, 28, 43, who echoes the sentiment expressed in Justice Jackson's concurrence in Railway Express Agency v. New York, 336 U.S. 106, 111–13 (1949). Professor Gunther makes the argument as part of his plea to abandon the fundamental-right branch of the strict-scrutiny inquiry, and thus to make the antidiscrimination principle more focused on means.

21. Gunther, p. 24.

likely to be shared by those seeking a "mechanical jurisprudence" and value-neutrality. I once again doubt the validity of the supposed ideal of objectivity,[22] but it cannot be denied that the antidiscrimination principle makes some contribution toward the satisfaction of this ideal, perhaps to a greater degree than toward the ideal of value-neutrality. With the possible exception of the inquiries necessary to identify the true criterion, an inquiry that need be taken only when administrative action is being challenged, the antidiscrimination principle is not especially fact-oriented. More often than not, over- and underinclusiveness is established, not by a presentation of evidence, but rather by the process of imagination–imagining whether, given the state purpose, other persons might be included within the coverage of the statute or whether people who were included might properly have been excluded. Little turns on the actual numbers involved. Moreover, although uncertainty and gradations of degree may be introduced by certain of the critical judgments required by the principle, for example, judgments about which purposes are legitimate and what is the requisite degree of fit, it is entirely possible that these judgments could be made with a high degree of generality. Once the Supreme Court spoke to an issue–for example, once the Court declared a certain criterion (such as race) to be suspect–a flat rule would emerge (no racial discrimination) that could easily be applied by the lower courts. Indeed, when the antidiscrimination principle was adopted by the legislative branch, and made the central regulatory device of the Civil Rights Acts of 1964 and 1968, it was expressed in a form that satisfied the objectivist ideal. The statutes specified the criteria (such as race, sex, religion, and national origin) that could not be the basis of discrimination.[23]

22. Fiss, "The Jurisprudence of Busing," 39 *Law & Contemp. Prob.* 194 (1975).

23. Some statutes make exceptions for certain criteria: under the Civil Rights Act of 1968, religion is a permissible criterion for allocating housing owned by a religious society, and under the Civil Rights Act of 1964 discrimination on the basis of sex is permitted in employment when sex is a "bona fide occupational criterion." Moreover, the courts have permitted the remedial use of a criterion that seemed to be flatly forbidden (for example, color-conscious

Fourth, the appeal of the antidiscrimination principle may derive from the fact that it appears highly individualistic. The method for determining the permissibility of classifications does not rely, so Tussman and tenBroek proclaim, on the concept of a "natural class" (where "natural" refers not to the biological origins of the class, but rather to the fact that it is not formally created by the law in question). They acknowledge that a judgment about the arbitrariness of a classification might conceivably depend on whether it "coincides with" social groupings deemed appropriate; in that instance, the central inquiry in an equal protection case would be "whether, in defining a class, the legislature has carved the universe at a natural joint."[24] Tussman and tenBroek sought to avoid that inquiry, an inquiry they declared to be "fruitless," and did so by making the permissibility of the classification turn exclusively on the relation of means to end. Hence, the antidiscrimination principle would seem individualistic in a negative sense—it is not in any way dependent on a recognition of social classes or groups. Indeed, that is why means-end rationality is such an attractive concept: it avoids the need of making any statement about the basic societal units.

To some degree this appearance is misleading. The foundational concept—means-end rationality—is individualistic. It is not dependent on the recognition of social groups. On the other hand, elements of groupism appear as one moves up the superstructure. For one thing, the recognition and protection of social groups may be required to determine which state purposes are legitimate, or even to rank state purposes to apply the compelling state-interest doctrine. The paradigm of a state purpose that is illegitimate is couched in terms of a group: "The desire to keep blacks in a position of subordination is an illegitimate state purpose." And the standard of illegitimacy is constructed by attributing what might be viewed as a group-oriented

---

employment is often decreed to correct the effects of past discrimination)—a judicial improvisation that has been ratified by Congress in the course of reenactments (Equal Employment Opportunity Act of 1972). From this perspective, despite the striking difference in language, the civil rights statutes have been treated as mini-equal protection clauses.

24. Tussman and tenBroek, p. 346.

purpose to the Equal Protection Clause—to protect blacks from hostile state action.[25] Admittedly the paradigm of a "compelling state interest" is not often expressed in group terms. That doctrine emerged in the *Japanese Relocation Cases* of the 1940s and there the state interest deemed "compelling" was "self-preservation of a nation at a time of war." But, as Justice Brennan recently perceived, the concept of a compelling state interest might be stretched to embrace the protection of certain groups. In the context of a statute that embodied a classification favoring women, he wrote: "I agree that, in providing a special benefit for a needy segment of society long the victim of purposeful discrimination and neglect, the statute serves the compelling state interest of achieving equality for such groups."[26]

The suspect-classification doctrine also affords some recognition to the role or importance of social groups or natural classes. This is apparent from the original and classic statement of the doctrine by Justice Black: "It should be noted, to begin with, that all legal restrictions which curtail the civil rights of a single racial group are immediately suspect."[27] Tussman and tenBroek, intent on keeping groups out of their account of the Equal Protection Clause, quoted this passage of Justice Black but were then careful to add that "suspect classification" should not be thought of as coextensive with a "single racial group." The obvious next question is whether there are any "suspect classifications" that do not identify a natural class or social group. To this, they simply replied, "[A]n attempt at an exhaustive listing of suspect classifications would be pointless. It suffices to say that this is of necessity a rather loose category."[28] In the last twenty-

25. I think this might—to be somewhat cynical—explain why Tussman and tenBroek tried to talk in terms of two different principles—that of "reasonable classification" and that of "discriminatory legislation"—even though they perceived the interconnections between the principles and the fact that they might be embraced with one principle—what I have called the antidiscrimination principle.

26. Kahn v. Shevin, 416 U.S. 351, 358–59 (1974) (dissenting opinion, joined by Justice Marshall). I suspect that group recognition might also enter through deciding when to honor the one-step-at-a-time defense. See *Katzenbach v. Morgan*, referred to fn. 10 above, another opinion written by Justice Brennan.

27. Korematsu v. United States, 323 U.S. 214, 216 (1944).

28. Tussman and tenBroek, p. 356.

five years, the category has been kept "loose"; but the important fact to note is that almost all of the serious candidates for the status of suspect classification are those that coincide with what might be conceived of as natural classes–for example, blacks, Chicanos, women, and maybe the poor. Moreover, although Tussman and ten-Broek did not even try to explain why certain classifications were suspect, it is not at all clear to me that an adequate explanation can be given that does not recognize the role and importance of social groups.

Some might explain the suspectness of race, to use the exemplar of a suspect classification, in terms of the special history of the Equal Protection Clause.[29] But that explanation does not altogether avoid the reference to groups, for it may be contended that the Clause was not intended to ban the racial classification but rather to protect blacks–as a group–from hostile state action. And in any event, that explanation might be too confining. It anchors the category of suspect classification in historical fact, without room for the kind of generality expected of constitutional doctrines, a generality that might be sufficient to embrace new situations (for example, the demand of women that sex be treated as a suspect classification).[30] Others might seek to explain the suspectness of race solely on the grounds of immutability.[31] That would avoid the reference to groups, but would be an inadequate explanation for it would also render suspect such classifications as "height," "good hearing," "good eyesight," or "intelligence"–a result the antidiscrimination theorist would no doubt deny. A final explanation for the suspectness of race that might avoid the reference to groups, and that indeed does have a connection with the foundational concept of means-end rationality, asserts that race is "generally . . . irrelevant to any legitimate public

29. In his dissent in Harper v. Virginia Bd. of Elections, 383 U.S. 663, 682, fn. 3 (1966), Justice Harlan argued that "insofar as that clause may embody a particular value in addition to rationality, the historical origins of the Civil War Amendments might attribute to racial equality this special status."

30. See Getman, The Emerging Constitutional Principle of Sexual Equality, 1972 *Sup. Ct. Rev.* 157.

31. See, for example, Frontiero v. Richardson, 411 U.S. 677, 686 (1973).

purpose."[32] This would make the individualism of the antidiscrimination strategy pure. But I fail to understand how a claim about general practices (race is generally unrelated) can yield a special standard about the degree of fit (or relatedness) to be required—and that is precisely the function of the suspect-classification doctrine—to trigger strict scrutiny, making any avoidable over- or underinclusiveness impermissible.

Once again, what we are left with is an illusion—that the antidiscrimination principle need not depend on the recognition of "natural classes." This illusion of individualism can be maintained only by ignoring or failing to justify some of the key elements of the antidiscrimination strategy—elements that might be deemed part of the superstructure, but are nonetheless esssential for they have made us willing to live with that strategy.

Wholly apart from this question of whether all the elements of the antidiscrimination principle—the superstructure as well as the foundation—are explicable on individualistic premises, it should be noted that the antidiscrimination principle furthers the ideal of individualism more subtly by making classification the focus of the Equal Protection Clause. Classification is the triggering mechanism and the object of inquiry. To be sure, not all classifications are prohibited, only those that are imprecise. Yet the demand for greater and greater precision in classification has the inevitable effect of disqualifying one classification after another, and that demand is consistent with, and indeed furthers the ideal of treating people as "individuals"—recognizing each

32. "Developments in the Law–Equal Protection," 82 *Harv. L. Rev.* 1065, 1108 (1969). In response to this explanation, Professor Ely argues, "The fact that a characteristic is irrelevant in almost all legal contexts (as most characteristics are) need not imply that there is anything wrong in seizing upon it in the rare context where it does make a difference.[39]" But this response seems to confuse the suspect-classification doctrine with the forbidden-classification doctrine—an absolute ban on all racial classification. This is seen most clearly by noting that Ely's fn. 39 refers to that portion of the Tussman and tenBroek article that eschews the forbidden-classification doctrine (for the reason he articulates) not to the portion of their article that deals with the suspect classification doctrine. It should be noted that Professor Ely's explanation for the suspectness of race, the we-they theory discussed below, is a theory cast in group terms—it seeks to explain legislative motivation in terms of group membership.

person's unique position in time and space, his unique combination of talent, ability and character, and his particular conduct. The pervasiveness of this ideal in society cannot be denied, nor is it likely that judges would be insensitive to it.

The tie between individualism and the antidiscrimination principle may also stem from the fact that it yields a highly individualized conception of rights. Under the antidiscrimination principle, the constitutional flaw inheres in the structure of the statute or the conduct of the administrator, not in its impact on any group or class. Any individual who happens to be burdened by a statute or practice, or any individual excluded from the benefits, can complain of the wrong. True, other persons--namely all those within the legal classification--can make a similar complaint; and in a sense the individual is making the claim as part of a group or class (the legal class). But, the individual's entitlement to relief is not dependent on the interests or desires of others similarly subject to or excluded from the statute or practice.

Such an individualized conception of rights coincides with one strong view of what we mean by a "constitutional right"--the vindication turns on the judgment of the tribunal, not upon the views or action of third parties. Institutional considerations also make the individualized conception of rights appealing. The Equal Protection Clause is primarily enforced through litigation, and it is especially difficult to fit the vindication of group rights into the mold of the law suit. There is no way of making certain that the plaintiff is the appropriate representative of the group, and even more, there is no mechanism for resolving intraclass conflicts--differences among the members of a group as to what is in their best interest.[33]

33. Such differences frequently arise and they are not in any way resolved by the class action device. That procedural device (viewed from the plaintiff's side) only legitimizes the concept of the self-appointed representative, and then seeks to erect safeguards--such as notice--to limit, as far as possible the risk of abuse arising from this power of self-appointment. But the factors that tend to legitimize the mechanism of self-appointment also tend to undermine the effectiveness of the safeguards--each individual stands to gain so little. When the stakes are small it does not make sense for an individual to start a law suit, and for the very same reason it does not make sense for the individual to respond to the notice in order to scrutinize the adequacy of his self-appointed

Finally, under the antidiscrimination principle, equal protection rights are not only individualized, but also universalized and this is another source of its appeal. Everyone is protected. There may be a limitation on the laws brought within its sweep—they must contain a discrimination or classification. But once a distinction among persons is made by a state statute or practice, that measure can be tested by anyone who happens to be burdened by it. Even the suspect-classification doctrine can be construed in universalistic terms—any racial classification, whether black, yellow, or white, is suspect. In contrast, a mediating principle that is, for example, built on the concept of social groups might not be so universal in scope, since it is conceivable that some individual adversely affected by the state might not be a member of one of the protected groups.

The universalizing tendency of the antidiscrimination principle no doubt accounts for its popular appeal—no person seems to be given more protection than another. This universalizing tendency also appeals to a court. It relieves the judiciary of the burden of deciding who will receive the protection (in the jurisdictional sense) of a constitutional provision and then explaining why some are left out. It also creates a strategic advantage for the court—it enables the court to use the Equal Protection Clause to fill some of the gaps created by the (temporary?) retirement of substantive due process. The antidiscrimination principle can be used just as comfortably to challenge a statute that draws a distinction between opticians and optometrists or one that draws a distinction between filled milk and margarine as it can be used to challenge a statute that draws a distinction between whites and blacks.[34] The result may be different, but that

representative. And without such a response, there is no reliable way of judging the adequacy of representation. There is an adversarial void—neither the defendant nor the self-appointed representative has an interest in challenging the adequacy. Indeed, from the defendants' perspective the best representative is an inadequate one (at least if adequacy is judged from the perspective of effectiveness).

34. See, for example, Williamson v. Lee Optical Co., 348 U.S. 483 (1955). It is striking that Tussman and tenBroek built their "reasonable classification" principle (which I call the antidiscrimination principle) almost entirely out of the business regulation cases—the traditional province of substantive due process. This can be explained in part by the fact that at least up until 1949,

is not due to any fundamental shift of theory, but only to a difference in the degree of fit required.

III. The Limitations of the Antidiscrimination Principle

The appeal of the antidiscrimination principle may be unfounded. The ideals served by the principle may not have any intrinsic merit, or the connection between those ideals and the principle may be nothing more than an illusion. As we have seen, the antidiscrimination principle may be criticized on this level. But I believe the criticism runs deeper. The antidiscrimination principle has structural limitations that prevent it from adequately resolving or even addressing certain central claims of equality now being advanced. For these claims the antidiscrimination principle either provides no framework of analysis or, even worse, provides the wrong one. Conceivably, the principle might be adjusted by making certain structural modifications; and indeed, on occasion, over the last twenty-five years, that has occurred, though on an ad hoc and incremental basis, and at the expense of severing the principle from its theoretical foundations and widening the gap between the principle and the ideals it is supposed to serve.

The Permissibility of Preferential Treatment

One shortcoming of the antidiscrimination principle relates to the problem of preferential treatment for blacks. This is a difficult issue, but the antidiscrimination principle makes it more difficult than it is: the permissibility of preferential treatment is tied to the permissibility of hostile treatment against blacks. The antidiscrimination principle does not formally acknowledge social groups, such as blacks; nor does it offer any special dispensation for conduct that benefits a disadvantaged group. It only knows criteria or classifications; and the color black is as much a racial criterion as the color white. The regime it introduces is a symmetrical one of "color blindness," making the criterion of color, any color, presumptively impermissible. Reverse

---

this was the principal use of the Equal Protection Clause. Their formulation was bound by the prior practice, just as mine is bound by the intervening twenty-five years' experience–where the Clause was principally used as a means of protecting the racial minority.

discrimination, so the argument is made, is a form of discrimination and is equally arbitrary since it is based on race.

The defense of preferential treatment under the antidiscrimination regime begins with the search for a purpose that the racial criterion (the color black) would fit perfectly. This is not an easy undertaking. To illustrate the difficulty, let us assume that the policy at issue is one preferring blacks for admission to law school.[35] The first impulse is to identify the purpose as one of increasing the number of black lawyers. Surely if that is the purpose, there is a perfect fit between criterion and purpose—no margins of over- or underinclusiveness. But what appears at first to be a purpose seems to be nothing more than a restatement of the practice. Why does the state want to increase the number of black lawyers? The answer to this question yields what may more properly be deemed a purpose.

An answer cast in terms of the self-interest of the class or in terms of the preferences of those in power (for example, they happen to like blacks) would not be adequate. These answers would not yield a legitimate state purpose. But a number of purposes would be served by the preferential treatment that could be deemed permissible. Here are some examples: to elevate the status of a perpetual underclass by giving certain members of the group positions of power and prestige (on the theory that the elevation of the group will enhance the self-image and aspirations of all members of the group); to insulate the minority from future hostile action by strategically placing members of the group in positions of power; to diversify the student body intellectually and culturally and thereby enrich the educational experience of all; or, finally, to atone for past wrongs to the group. The difficulty with each of these purposes is, however, that once the perspective shifts from groups to individuals, as the antidiscrimination principle requires, the margins of under- and/or overinclusiveness become apparent and indeed pronounced. The fit between criterion (black) and purpose is not perfect—perhaps just as imperfect as the fit between the criterion and purpose when

35. DeFunis v. Odegaard, 416 U.S. 312 (1974). See also two symposia, "DeFunis: The Road Not Taken," 60 *Va. L. Rev.* 917 (1974), and "DeFunis Symposium," 75 *Colum. L. Rev.* 483 (1975).

the action is hostile. The overinclusiveness stems from the fact that there are blacks who are not entitled to the preferential treatment if any of these were the state purpose (the common example is the upper class black, who arguably did not suffer past discrimination and/or independently of the preferential admission to the law school, would be a "success"). The underinclusiveness stems from the fact that there are other persons who are not black and who are nevertheless as entitled to preferential treatment as blacks if the state purpose is what I have imagined (for example, Chicanos, Orientals, the poor).

The next move in the defense of preferential treatment under the antidiscrimination principle is to discover ways of tolerating these margins of ill-fit. The defense of the overinclusion is likely to be couched in terms of administrative convenience.[36] True, not all blacks are entitled to the preferential treatment, but it would be exceedingly difficult and costly to try to pick out those not deserving the preferential treatment, and the costs would not be worth the gains. Administrative convenience may also be used to justify the underinclusion, particularly as it relates to the poor. It would be difficult, so the argument runs, to pick out those nonblacks who have the same social or economic status as blacks and thus, under the stipulated purposes, are as deserving of preferential treatment as blacks. Blackness is an easy criterion to work with, and although there may be mistakes, they are small compared to the costs inherent in the use of alternative criteria ("poor" or "low socioeconomic status"). The difficulty with this administrative convenience argument is that it is standard practice to reject such a defense when whites rather than blacks are the preferred race. Why should this defense be accepted in one context and not another?

Professor Ely's we-they analysis (see fn. 3) might be thought to be responsive to this dilemma. He argues that when the dominant group (whites) use the racial criterion for conferring benefits on the minority (preferential treatment for blacks), there is less reason to be suspicious than when they use the racial criterion for conferring benefits on their

36. See Nickel, "Preferential Policies in Hiring and Admissions: A Jurisprudential Approach," 75 *Colum. L. Rev.* 534, 550–53 (1975).

own class (preferential treatment for whites). When benefits are conferred by one class upon its own members, the risk is high that the arguments about administrative convenience are a sham: it is like voting oneself a pay raise. The risk is high that you would believe any argument in favor of the decision, irrespective of the merits. But when a sacrifice is involved, so the argument continues, as when members of one group (we) confer a benefit on another group (they) at its own (our) expense, then there is less reason to be suspicious of the arguments used to defend that action.

I have some difficulty with the psychological model upon which the we-they analysis rests. It is incomplete. The only motivational factor reflected is self-interest (elaborated in terms of group membership). But, of course, as seems particularly true in the case of preferential admission of blacks to law school, there may be other motivational factors that make arguments about administrative convenience especially suspect. The body making the decision may tend to overvalue arguments about administrative convenience out of a feeling of guilt or fear (for example, of disruption in the university) and there may be little to check those impulses since the costs of the preferential policy are primarily borne by "others" (not the professors or administrators who decide upon the admission policy but by some of the rejected nonblack applicants–not the superstars, but rather by those who are at the end of the meritocratic queue).[37]

The principal difficulty with this we-they analysis is not, however, the incompleteness of the psychological model; for it is conceivable this could be corrected. Rather the principal difficulty stems from what this model (or perhaps any model focusing on the psychology of the discriminators) yields. As Professor Ely acknowledges, the we-they analysis can only provide grounds for (asymmetrical) suspicion, and yet that does not seem sufficient.[38] For even if suspicion turns

37. Burt, "Helping Suspect Groups Disappear" (unpublished manuscript, 1975); Greenawalt, pp. 573–74.

38. I think Ely went wrong in reading the suspect-classification doctrine too weakly. When the Supreme Court and commentators, such as Tussman and tenBroek, spoke of suspect classification, they were trying to express a substantive, not just an evidentiary judgment–that in the generality of cases certain kinds of classification will be invalid. The suspect-classification doc-

out to be wrong, and the argument of administrative convenience is determined to have merit, to be sincere and well-founded, the argument would still be rejected as a justification for ill-fit when preferential treatment is being conferred on whites.

For example, imagine it is the 1940s, the state electorate is predominantly white and the state legislature directs the law school to adopt a preferential admission policy in favor of whites.[39] Assume also that this policy is justified on the ground that whites are better prepared academically (given the dual school system) and that the state wishes to have the most brilliant persons as members of the bar. Color is used because of administrative convenience. Under the we-they analysis, there is reason to be suspicious about this explanation of the use of race; the policy in effect serves the class interest of whites and it is likely that the white legislators (or administrators) will undervalue the costs to the blacks of an imprecise fit and mistakenly believe only the "negative myths" about blacks. But it would seem to me that even if that suspicion were refuted, even if it were (somehow) demonstrated that the argument about administrative convenience was sincere and well-founded and in some sense accurate, the result would still be unacceptable under the Equal Protection Clause. Those who are committed to the antidiscrimination principle can reach this result only by insisting that the arguments of administrative convenience are no defense to ill-fit, even assuming that they are sincere and well-founded. It is that move, above all, that plays into the hands of those who wish to attack preferential treatment in favor of blacks: why, they will ask, should administrative convenience be

---

trine was simply a way of avoiding a flat no-exception per se rule—a doctrine of forbidden classifications (no racial classifications at all). When it rejected the forbidden-classification doctrine, the Court may have been looking ahead to the problem of preferential treatment and for that reason rejected the forbidden-classification doctrine. But the discussion of this issue in Tussman and ten-Broek, and the fact that the doctrine originates in the *Japanese Relocation Cases*, leads me to other explanations: a forbidden-classification doctrine would be hard to reconcile with the generality of the language of the Equal Protection Clause; and it would have tied the hands of the legislators and administrators too much, precluding the use of the classification in even extraordinary instances—such as those involving the relocation of the Japanese.

39. See Sweatt v. Painter, 339 U.S. 629 (1950).

allowed to justify ill-fit when the state action is beneficial to the minority but not when it is hostile to the minority?

The we-they analysis cannot resolve this dilemma because it is only an evidentiary, and not a substantive, approach. There are, however, two other strategies employed within the context of the antidiscrimination principle to justify preferential treatment, and these do seem more substantive. One such strategy shifts the definition of the harm —classifications should not be judged in terms of the means-end relationship (fit), but rather in terms of whether they stigmatize.[40] An exclusionary classification aimed at blacks stigmatizes them in a way that preferential treatment does not stigmatize whites; administrative convenience cannot justify or offset the stigmatizing harm caused by the exclusionary policy, and the nonstigmatizing preferential policy does not call for a justification (or if it calls for one, it need be only a weak one, for which administrative convenience will suffice). I am willing to assume that the preferential policy does not stigmatize the rejected applicants,[41] and yet this strategy still seems unsatisfactory. It moves beyond the structure of the antidiscrimination principle in that (a) the evil becomes stigma rather than ill-fit, and (b) it contemplates a weighing of ends—a judgment as to whether the state interest is of sufficient importance to offset the harm. A connection with the original antidiscrimination principle remains: the trigger remains the same—classification. But that seems to be a trivial connection. The tie with the foundational concept—means-end rationality—is severed. And once this step is taken, it is hard to confine the modification. It is difficult to explain why classification should be the only trigger for the Equal Protection Clause and, even if it is, why that Clause should be concerned only with stigmatic harm (or why stigmatic harm alone is capable of overriding a defense of administrative convenience). The rejected white applicant may not be stigma-

40. See Brest, "Palmer v. Thompson: An Approach to the Problem of Unconstitutional Legislative Motive," 1971 *Sup. Ct. Rev.* 95, 116 fn. 109; and Brest, *Processes of Constitutional Decisionmaking* (Boston, 1975), p. 481. See also Greenawalt, p. 566.

41. Fiss, "School Desegregation: The Uncertain Path of the Law," 4 *Philosophy & Public Affairs* 3, 13 (1974). Some have argued that a preferential admissions policy stigmatizes the blacks admitted under it. See Burt, fn. 37 above.

tized, but he is being harmed in other ways; and there is nothing in the theory underlying the antidiscrimination principle that suggests why nonstigmatic harm should be given a subordinate (or weaker) status. And if it is not given that subordinate status, then we are back to the same dilemma of symmetry: why reject administrative convenience in one context and not in the other?

The final move in defense of preferential treatment under the antidiscrimination principle is to invoke one or both of the standard defenses. The charge of underinclusiveness might be defended on the one-step-at-a-time theory—this time the law school helps blacks, next time Chicanos, and so on. This would be considered a "reform" measure, while the exclusionary policy would only be a "regression." The defense of compelling state interest may also be deployed: the purposes served by the preferential admission program are especially desirable or important, while those served by the exclusionary one (having the bar consist of the academically superior) are ordinary.[42] The upshot of both defenses is that the ill-fit (the over- and underinclusiveness) is excused in one case (the preference for blacks) but not in the other (the exclusion of blacks).

The problem with both these defensive moves is that they are devoid of any theoretical foundations. In the ultimate analysis, they are resolution by fiat; for the antidiscrimination principle does not supply any basis or standards for determining what is "reform" and what is "regression," what is an "ordinary" state purpose, and what is a "special" one. These distinctions can only be made if the court has some notion of what is "good" or "desirable," only if the court identifies certain substantive ends as those to be favored under the Equal Protection Clause. As an intellectual feat this may be possible but not within the confines of the antidiscrimination principle. That principle

42. Greenawalt, pp. 574–79, distinguishes between a "compelling interest" and a "substantial interest." He argues, "Because benign racial classifications are less 'suspect,' however, a 'substantial' public interest should be enough to support them." Karst and Horowitz, on the other hand, contend that racial classifications "must be tested against the exacting standard of the 'compelling state interest' formula." "Affirmative Action and Equal Protection," 60 *Va. L. Rev.* 955, 965 (1974).

disclaims any reliance on substantive ends.[43] Indeed that is thought to be a primary source of its appeal.

In my judgment, the preferential and exclusionary policies should be viewed quite differently under the Equal Protection Clause. Indeed, it would be one of the strangest and cruelest ironies to interpret that Clause in such a way that linked—in some tight, inextricable fashion—the judgments about the preferential and exclusionary policies. This dilemma can only be avoided if the applicable mediating principle of the Clause is clearly and explicitly asymmetrical, one that talks about substantive ends, and not fit, and one that recognizes the existence and importance of groups, not just individuals. Only then will it be possible to believe that when we reject the claim against preferential treatment for blacks we are not at the same time undermining the constitutional basis for protecting them. Of course, even if the antidiscrimination principle were not the predominant interpretation of the Clause, it might still be possible to formulate a claim against preferential treatment. The element of individual unfairness to the rejected applicants inherent in preferential treatment could be considered a cost in evaluating the state action in the same way as a loss of liberty or a dignitary harm might be. The failure of the state to include other disadvantaged groups, such as the Chicanos, might also become significant. But the impenetrable barrier posed by the seemingly symmetrical antidiscrimination principle would be gone. The stakes would not be so high.

Nondiscriminatory State Action

The antidiscrimination principle has created several gaps in the coverage of the Equal Protection Clause. The principle purports to be universalistic in terms of the persons protected, and yet it turns out to be far from universalistic in terms of the state practices proscribed. The gaps in coverage arise from the fact that not all objectionable state conduct is discriminatory. Discrimination involves a choice among persons and, as I said, an antidiscrimination principle operates by prohibiting government from making that choice arbitrarily. But there are

43. It might also be noted that this disclaimer is also inconsistent with the legitimate purpose limitation.

government enactments or practices where no choice is made among persons and of these it does not make sense to ask whether there is "arbitrary" discrimination. I am not complaining of the fact that the antidiscrimination principle leaves standing state conduct that should be invalidated; but rather that it provides no frame of reference for assessing certain types of state conduct and for that reason is incomplete.

This gap in part accounts for the difficulty the Supreme Court has had with some of the classic state action cases. One such case is *Shelley v. Kraemer*.[44] The Court there invalidated a state policy of enforcing racially restrictive covenants, and although that result seems right, on an analytic level *Shelley v. Kraemer* is generally deemed to be an extraordinarily difficult case–the Finnegans Wake of constitutional law.[45] The difficult question was not, in my judgment, whether the state judges who enforced the restrictive covenant were acting as representatives of the "state." True, that issue was discussed by the Court, but it hardly seemed of any moment. Rather the troublesome question arose in trying to determine whether the state's action was the kind of "action" prohibited by the Equal Protection Clause. The Clause was viewed as prohibiting (racial) discrimination, and only that. The state asserted that its policy was not in any way discriminatory–restrictive covenants would be enforced against blacks and whites alike.

The basis of the Court's rejection of this defense remains a mystery to me to this day. Only a couple of sentences in the opinion purport to be responsive. In one the Court mentions the factual assertions of plaintiffs that, by and large, these racially restrictive covenants are used against blacks, rather than whites. The Court seemed willing to assume the truthfulness of this assertion as a factual matter, but it was hesitant to conclude much from it. That seemed a sound instinct, provided the Court was confined to the antidiscrimination principle and wanted to invalidate the policy, rather than its application; as long as the state stands ready to, and in fact would, enforce

44. 334 U.S. 1 (1948).

45. Kurland, "Foreward: Equal in Origin and Equal in Title to the Legislative and Executive Branches of the Government," 78 *Harv. L. Rev.* 143, 148 (1964).

a racially restrictive covenant against whites this state policy cannot itself be deemed a form of racial discrimination. The other response is the well-known passage of *Shelley v. Kraemer* declaring that the Equal Protection Clause protects individual rights.[46] But I fail to see why this is responsive to the state's defense—there is no discrimination by the state. The more appropriate response to the state would be to reject its premise as to the kind of state action prohibited by the Clause. Why, I would ask, must the action of the state be discriminatory before it is deemed a violation of the Equal Protection Clause?

Recently, the Court was faced with another state action case that presented a similar problem and, given the faulty frame of reference, the result was not so fortuitous. In *Moose Lodge*,[47] Pennsylvania granted a liquor license to a private club that discriminated on the basis of race. The club refused to serve blacks. The liquor license did not confer monopoly power but it was of great benefit to the club and, even more importantly, it had the effect of limiting the places available in the locality for blacks to purchase liquor. Only a limited number of liquor licenses were made available in each area, and a license was required before liquor could be sold. Once again, the defense of the state was that it did not discriminate on the basis of race: the state did not exclude blacks from the Lodge, nor did it grant the license because the Lodge was discriminating on racial

46. The text reads: "The rights created by the first section of the Fourteenth Amendment are, by its terms, guaranteed to the individual. The rights established are personal rights." 334 U.S., at 22. A clever lawyer might have asserted that the discrimination was not between whites and blacks, but rather between two classes of sellers—those who sell land burdened with a restrictive covenant and those who sell unencumbered land. But if that were the challenged distinction, we have moved beyond the realm of suspect classifications and thus might have to operate under a minimum scrutiny inquiry. The Court did not seem willing to operate at that level; for them it was a racial case—a wrong to blacks. It is interesting to note that Tussman and tenBroek did not see *Shelley v. Kraemer* as resting on the "reasonable classification" (or antidiscrimination) principle. They did not view the case as a racial one, but rather as a matter of "substantive due process"—interference with the liberty to sell—though recognizing, given the bad taste left by that doctrine, that it might have to be called "substantive equal protection." Tussman and tenBroek, at 362; see fn. 2 above.

47. Moose Lodge No. 107 v. Irvis, 407 U.S. 163 (1972).

grounds. The state simply regarded the admission practice of the club as an irrelevance. The Court thought it had to link the state with the discriminatory refusal to serve. According to Justice Rhenquist, the black was claiming that "the refusal to serve him was 'state action' for the purposes of the Equal Protection Clause."[48] Of course, under this formulation the black could not win, and relief was in fact denied.[49]

Both the result of *Moose Lodge* and the mode of analysis seem wrong. The Equal Protection Clause does not govern the behavior of private clubs, but it does govern the conduct of the state. The state was not discriminating, racially or otherwise, but it was engaging in conduct–the conferral of liquor licenses (without regard to admission policies)–and that action could be evaluated in terms of the Equal Protection Clause. Moose Lodge may be free to discriminate, but that does not make it correct for the state to confer a scarce franchise on the club, thereby foreclosing opportunities to blacks. It was the premise of the Court that the only kind of action denied the state is discriminatory action that prevented it from focusing on this foreclosure of opportunities to blacks.

In the state action cases such as *Shelley v. Kraemer* and *Moose Lodge* there were clearly acts of discrimination–the restrictive covenant itself or the refusal of the club to serve blacks. Those discriminatory acts were not performed by the state, and the Court saw the question as whether the discriminatory acts could be *imputed* to the state.[50] In another group of cases, also put beyond the Equal

48. 407 U.S., at 165.

49. With a nod toward *Shelley v. Kraemer*, the Court prohibited the state from directing the club to comply with the national Supreme Lodge rules, which embodied the racial policy. The local was allowed to make up its own mind.

50. I think this imputation occurs in the Proposition 14 case, Reitman v. Mulkey, 387 U.S. 369 (1967). The Court thought it had to impute the discrimination (the refusal to sell) to the state, and it tried to create the linkage by saying that Proposition 14 had the effect of "encouraging" the discrimination. This conceptualization probably prevented the Court from relying on the rationale linked to Charles Black's name, "Foreword: 'State Action,' Equal Protection and California's Proposition 14," 81 *Harv. L. Rev.* 69 (1967) (an additional obstacle was created for blacks in their effort to obtain protective legislation).

Protection Clause by the antidiscrimination interpretation, there was no clear act of discrimination, racial or otherwise, that could easily be imputed to the state. The acts in question clearly belonged to the state, but they were not discriminatory. The antidiscrimination principle left the Court even more at sea. Here I have in mind the on-off decisions of government–the decision whether or not to have a public facility, such as a swimming pool[51] or a public housing project.[52]

The principal challenge to these decisions focused attention on the basis for the decision–why did the town close the swimming pools? Why did the town refuse to build a public housing project? These questions *seem* similar to the inquiry required by an antidiscrimination principle, but that appearance is misleading. The kind of decision most amenable to an antidiscrimination analysis is one choosing among persons, a state decision, for example, about who shall be admitted to the swimming pool or the public housing project. With decisions of that sort, the why-question asks for an identification of the *criterion* of selection: why were these individuals and not others allowed into the swimming pool or the housing project? Answer: because of their race. Once identified, the criterion of selection could be judged under the antidiscrimination principle in terms of its relatedness to the state purpose. With an on-off decision, such as a decision whether or not to provide a public service, the why-question asks, not for the criterion, but for the *motive* or *purpose* itself: why did the town close the swimming pool? Answer: in order to save money or to prevent integration. Why did the town refuse to build a public housing project? Answer: in order to save money or in order to limit the number of poor persons in the community. Accordingly, in this context there is no criterion of selection that can be evaluated for its relatedness.[53] Guided by the antidiscrimination principle, and

51. Palmer v. Thompson, 403 U.S. 217 (1971).

52. James v. Valtierra, 402 U.S. 137 (1971).

53. In some situations, it may be able to reformulate the on-off decision as a choice among various public activities: why did the town decide to close the swimming pool rather than stop the buses. That why-question does yield a criterion of selection, but a different type of one–not an individual trait (such as race or performance on a test), but rather communal goals (to promote full employment). And it is the former type of criterion of selection that permits a court

that alone, one hardly knows where to begin in analyzing these governmental decisions, and I think that accounts for the difficulty the Court has experienced with them. One might conclude that these decisions are not invalid, or that relief should not be provided because it is impossible to fashion an appropriate judicial remedy. But the difficulty with the antidiscrimination interpretation is that it puts these on-off decisions beyond purview of the Equal Protection Clause.

The Problem of Facially Innocent Criteria

The classic state action cases and those cases involving on-off decisions reveal the inability of the antidiscrimination principle to deal adequately with state conduct that does not discriminate among persons. Another problem area arises from state conduct that does in fact discriminate among persons, but not on the basis of a suspect criterion. The discrimination is based on a criterion that seems innocent on its face and yet nonetheless has the effect of disadvantaging blacks (or other minorities). For example, when the state purports to choose employees or college students on the basis of performance on standardized tests, and it turns out that the only persons admitted or hired are white.

As originally conceived–both by Tussman and tenBroek and by the Supreme Court in the important formative period of the 1940s and 1950s–the antidiscrimination principle promised to evolve a small, finite list of suspect criteria, such as race, religion, national origin, wealth, sex. These would be presumptively impermissible. The great bulk of other criteria may ultimately be deemed arbitrary in some particular instances because of ill-fit, but they would be presumptively valid. For these criteria–which I call *facially innocent*–the mere rational-relation test would suffice, and the probability would be very high that the statute or administrative action incorporating or utilizing such criteria would be sustained.

---

to ask the foundational question of the antidiscrimination principle–one of over- and underinclusiveness of persons. The reader should consult Professor Brest's article on *Palmer v. Thompson* (see fn. 40 above), for he views the motivational analysis as a much closer relative of the antidiscrimination analysis than I would.

In some instances the presumption of validity may be dissolved, and the contrary presumption created, through the use of the concept of the *real* criterion. The plaintiffs can charge cheating: while the state says that it is selecting on the basis of an innocent criterion (such as performance on a written test), in truth the selection is being made on the basis of a suspect criterion (race). The substantiation of this charge confronts the plaintiffs with enormous evidentiary burdens. No one can be expected to admit to charges of cheating, and rarely is the result so striking (for example, the twenty-eight-sided voting district of *Gomillion v. Lightfoot*[54] or no blacks on the work force) as to permit only one inference—discrimination on the basis of a suspect criterion. But if the charge could be substantiated (perhaps with an assist from the reallocation of the burdens of proofs when the criterion had almost the same effect as a suspect one), then there would be no problem of using the strict-scrutiny branch of the anti-discrimination principle: the real criterion, as opposed to the stated criterion, is a suspect one, and there the court should insist upon a very tight fit between purpose and criterion. The troublesome cases arise, however, when the charge of cheating cannot be substantiated, where, for example, the court finds that in truth the jobs were allocated or students selected on the basis of academic performance. What then?

One possible response is, of course, to apply the mere rational-relation test and validate the practice: there is certainly some connection between the state's purposes and these criteria. The fit may not be perfect, but perfection is not required. But the courts have balked. They have been troubled by the fact that the practice is particularly injurious to a disadvantaged group and for that reason have scrutinized state conduct with the greatest of care. The judicial inclination is all toward invalidation.[55] This impulse seems correct as

54. 364 U.S. 339 (1960).

55. See, for example, Lau v. Nichols, 414 U.S. 563 (1974); Keyes v. School Dist. No. 1, Denver, Colo., 413 U.S. 189 (1973); Swann v. Charlotte-Mecklenburg Board of Education, 402 U.S. 1 (1971); Griggs v. Duke Power Co., 401 U.S. 424 (1971); Gaston County v. United States, 395 U.S. 285 (1969). Some of these cases, such as *Griggs*, involve civil rights statutes. The language and legislative history of those statutes generally restrain rather than encourage the judicial inclination. The Court may have been responding to the fact that

a matter of substantive justice, and yet it is difficult to reconcile treatment of the facially innocent criteria with the original, modest conception of the antidiscrimination principle.

One response–that which emerged most clearly during the mid-1960s phase of the Warren Court[56]–was to postulate a second trigger for strict scrutiny–impingement of a fundamental right. This, of course, constituted a radical modification of the antidiscrimination principle, for it introduced a ranking of ends or interests, and it meant that the antidiscrimination principle could no longer be justified exclusively in terms of means-end rationality. The pretense of value-neutrality could not be easily maintained. That is why Tussman and tenBroek did not anticipate the doctrine,[57] and why Professor Gunther, who wrote in the early days of the Burger Court, and who appears committed to returning the Equal Protection Clause to its former glory as a means-focused inquiry, disavows the fundamental-right trigger of the strict-scrutiny inquiry.[58] But even if this modification of the antidiscrimination principle were accepted, and ends were ranked and weighed, the problem of justifying the judicial treatment of facially innocent criteria that especially disadvantaged blacks could not be solved. Admittedly in some instances, the criteria in fact invalidated impinged on what might be considered a fundamental interest; the literacy test for voting might be such an instance.[59] On the other hand, the striking fact is that the Supreme Court has afforded strict scrutiny to facially innocent criteria that do not im-

---

legislative revision is easier when a statute is being construed, though my impression is that civil rights statutes soon become minimum guarantees not open to the ordinary legislative processes of revision, and thus have the same degree of permanence as a constitution.

56. See Note, "Developments in the Law–Equal Protection," 82 *Harv. L. Rev.* 1065, 1120, 1127 (1969), which was written at the peak of the fundamental-right development, codified the development and was thereby instrumental in legitimating it.

57. See fn. 2 above.

58. Gunther, p. 24. Professor Gunther retains the suspect classification branch of the traditional equal protection analysis.

59. Gaston County v. United States, 395 U.S. 285 (1969). In that case, however, the Court did not seek to explain its decisions in terms of a fundamental interest, and this is not surprising since the opinion was written by Justice Harlan. See fn. 8 above.

pinge on fundamental interests—for example, test scores as a criterion for jobs[60] and residence as a criterion for assignment to schools.[61] Although "jobs" and "schools" might in some view be deemed fundamental, certainly as important as the right to procreate, one of the first rights deemed "fundamental," this extension of the fundamental-right test would make what might first have seemed to be an exception much greater than the rule.

A second, and seemingly more modest way of rationalizing the judicial treatment of facially innocent criteria, is to introduce the concept of past discrimination. Strict scrutiny should be given, so the argument runs, to state conduct that perpetuates the effects of earlier conduct (it might be state or private) that was based on the use of a suspect classification. Conduct that perpetuates the effects of past (suspect-criterion) discrimination is as presumptively invalid as the present use of suspect criteria. An objective civil service test is presumptively impermissible whenever it perpetuates the past discrimination of the dual school system (the dual school system put the blacks at a competitive disadvantage and the test perpetuates that disadvantage). The use of geographic proximity is an impermissible criterion of school assignment whenever it perpetuates the past discrimination of the dual school system. The racial assignments of that school system led to the present residential segregation and account for the location and size of the school buildings, and both of these factors in turn explain why the use of geographic proximity as a criterion of assignment results in segregated patterns of school attendance today.

A ban on "the perpetuation of past arbitrary discrimination" looks like a close cousin of the ban on "arbitrary discrimination." But this tie can only be maintained at great expense to important institutional

60. Griggs v. Duke Power Co., 401 U.S. 424 (1971). Although that case involved a statutory discrimination claim (Title VII of the Civil Rights Act of 1964), recent cases illustrate use of the same principle outside the statutory context, e.g. Bridgeport Guardians, Inc. v. Bridgeport Civil Serv. Comm'n., 482 F. 2d 1333 (2d Cir. 1973), cert. denied 421 U.S. 991 (1975); Carter v. Gallagher, 452 F. 2d 315, 327 (8th Cir. 1972), cert. denied, 406 U.S. 950 (1972).

61. Keyes v. School Dist. No. 1, Denver, Colo., 413 U.S. 189 (1973). But see Milliken v. Bradley, 418 U.S. 717 (1974).

values—those that cluster around the ideal of objectivity, an ideal the antidiscrimination principle is supposed to serve. A true inquiry into past discrimination necessitates evidentiary judgments that are likely to strain the judicial system—consume scarce resources and yield unsatisfying results. It would require the courts to construct causal connections that span significant periods of time, periods greater than those permitted under any general statute of limitations (a common device used to prevent the judiciary from undertaking inquiries where the evidence is likely to be stale, fragmentary, and generally unreliable). The difficulties of these backward-looking inquiries are compounded because the court must invariably deal with aggregate behavior, not just a single transaction; it must determine the causal explanation for the residential patterns of an entire community, or the skill levels of all the black applicants.

There are techniques for reducing these strains. The court can create presumptions to limit the evidentiary inquiries or dispense with the need for a showing of identity between victim and beneficiary, or between past perpetrator and present cost-bearer. But these techniques have their own costs. The use of presumptions involves the court in fictionalizing and thereby impairing its credibility. And, more importantly, once the connections between victim and beneficiary and between past perpetrator and present cost-bearer are severed, we have ceased talking about the perpetuation of past discrimination in any individualized sense. The past discrimination that we are talking about is of a more global character—for example, that the group were slaves for one century and subject to Jim Crow laws for another. The ethical significance of this global past discrimination cannot be denied; it gives the group an identity and might explain why we are especially concerned with its welfare. But at the same time it should be understood that once we start talking of global past discrimination, the link between the proposed anti-past-discrimination principle and the original antidiscrimination principle becomes highly attenuated. We have embarked on another journey altogether, one that is decidedly not individualistic—and, as a result, one important source of appeal of the antidiscrimination principle is lost.

The third move designed to deal with the problem of facially

innocent criteria—the introduction of the concept of de facto discrimination (or discriminatory effect)—does not focus on the past. Instead it shifts the trigger for strict scrutiny from the *criterion* of selection to the *result* of the selection process, and the result is stated in terms of a *group* rather than an individual. What triggers the strict scrutiny is not the criterion of selection itself, but rather the result —the fact that a minority group has been especially hurt. (This special hurt is sometimes described as a "differential impact.")

This concept of de facto discrimination also involves a basic modification of the antidiscrimination principle. The trigger is no longer classification, but rather group-impact. This modification deeply threatens two goals allegedly served by the antidiscrimination principle—objectivity and individualism.[62] Of course, even with group-impact as the trigger, it is still possible to ask the ultimate question of the antidiscrimination principle—are there (excessive) margins of over- and underinclusiveness? But the modification does reveal the basic poverty of antidiscrimination theory. It makes me acutely aware of the failure to explain (a) why this finite list of criteria deemed suspect should ever trigger the demand for stricter scrutiny and (b) why scrutiny should be confined to determining the fit of the criterion and purpose. The concern with the result reveals to me that what is ultimately at issue is the welfare of certain disadvantaged groups, not just the use of a criterion, and if that is at issue, there is no reason why the judicial intervention on behalf of that group should be limited to an inquiry as to the degree of fit between a criterion and a purpose.

62. And that is why Professor Goodman, upon becoming aware of this fact, believes he has discovered the decisive argument against a theory of de facto discrimination: "These and countless other de facto discriminations would be disallowed by a rule condemning, or requiring special justification for, all state action disproportionately harmful to members of minority groups. The objection to such a rule is not solely one of practicality, but also one of principle. It is the individual, not the group, to whom the equal protection of laws is guaranteed.[102]" See "De Facto School Segregation: A Constitutional and Empirical Analysis," 60 *Calif. L. Rev.* 275, 300–301 (1972). Goodman's fn. 102 does not elaborate, but only refers the reader to the mystifying sentence of *Shelley v. Kraemer* quoted above in fn. 46.

IV. The Group-Disadvantaging Principle

The Shift from Classification to Class: Integrating the Concept of a Disadvantaged Group into the Law

In attempting to formulate another theory of equal protection, I have viewed the Clause primarily, but not exclusively, as a protection for blacks. In part, this perspective stems from the original intent—the fact that the Clause was viewed as a means of safeguarding blacks from hostile state action. The Equal Protection Clause (following the circumlocution of the slave-clauses in the antebellum Constitution)[63] uses the word "person," rather than "blacks." The generality of the word chosen to describe those protected enables other groups to invoke its protection; and I am willing to admit that was also probably intended. But this generality of coverage does not preclude a theory of primary reference—that blacks were the intended primary beneficiaries, that it was a concern for their welfare that prompted the Clause.

It is not only original intent that explains my starting point. It is also the way the courts have used the Clause. The most intense degree of protection has in fact been given to blacks; they have received a degree of protection that no other group has received. They are the wards of the Equal Protection Clause, and any new theory formulated should reflect this practice. I am also willing to speculate that, as a matter of psychological fact, race provides the paradigm for judicial decision. I suspect that in those cases in which a claim of strict

63. At least three clauses in the Constitution refer to slaves without using the word "slave," though the circumlocution in fact is so transparent as to be mystifying. The Slave Trade Clause provides, "The Migration or Importation of such Persons as any of the States now existing shall think proper to admit, shall not be prohibited by the Congress prior to the Year one thousand eight hundred and eight, but a Tax or duty may be imposed on such Importation, not exceeding ten dollars for each Person." (Article I, section 9, clause 1). See also Article I, section 2, clause 3 ("three-fifths of all other Persons") and Article IV, section 2, clause 3 ("Person held to Service or Labour"). The issue of circumlocution was debated in the original convention in connection with the Slave Trade Clause, some arguing that the generality of the word "persons" would have unintended effects—impliedly creating national power over immigration. For the discussion during the Constitutional Convention see Farrand, 2 *Records of the Federal Convention* 415 & fn. 8 (1937).

scrutiny has been or reasonably could have been made, it is commonplace for a judge to reason about an equal protection case by thinking about the meaning of the Clause in the racial context and by comparing the case before him to a comparable one in the racial area. Moreover, the limitations or inadequacies of the antidiscrimination principle surface most sharply when it is used to evaluate state practices affecting blacks.

Starting from this perspective, a distinctively racial one, it strikes me as odd to build a general interpretation of the Equal Protection Clause, as Tussman and tenBroek did, on the rejection of the idea that there are natural classes, that is, groups that have an identity and existence wholly apart from the challenged state statute or practice. There are natural classes, or social groups, in American society and blacks are such a group. Blacks are viewed as a group; they view themselves as a group; their identity is in large part determined by membership in the group; their social status is linked to the status of the group; and much of our action, institutional and personal, is based on these perspectives.

I use the term "group" to refer to a social group, and for me, a social group is more than a collection of individuals, all of whom, to use a polar example, happen to arrive at the same street corner at the same moment. A social group, as I use the term, has two other characteristics. (1) It is an *entity* (though not one that has a physical body). This means that the group has a distinct existence apart from its members, and also that it has an identity. It makes sense to talk about the group (at various points of time) and know that you are talking about the same group. You can talk about the group without reference to the particular individuals who happen to be its members at any one moment. (2) There is also a condition of *interdependence*. The identity and well-being of the members of the group and the identity and well-being of the group are linked. Members of the group identify themselves—explain who they are—by reference to their membership in the group; and their well-being or status is in part determined by the well-being or status of the group. That is why the free blacks of the antebellum period—the Dred Scotts—were not really free, and could never be so long as the institution of Negro

slavery still existed.[64] Similarly, the well-being and status of the group is determined by reference to the well-being and status of the members of the group. The emancipation of one slave–the presence of one Frederick Douglass–may not substantially alter the well-being or status of the group; but if there were enough Frederick Douglasses, or if most blacks had his status, then surely the status of blacks as a social group would be altered. That is why the free black posed such a threat to the institution of slavery. Moreover, the identity and existence of the group as a discrete entity is in part determined by whether individuals identify themselves by membership in the group. If enough individuals cease to identify themselves in terms of their membership in a particular group (as occurs in the process of assimilation), then the very identity and separate existence of the group–as a distinct entity–will come to an end.

I would be the first to admit that working with the concept of a group is problematic, much more so than working with the concept of an individual or criterion.[65] It is "messy." For example, in some in-

64. On the plight of the free blacks, see I. Berlin, *Slaves Without Masters: The Free Negro in the Antebellum South* (New York, 1974); J. Franklin, *From Slavery to Freedom*, 3d ed. (New York, 1967); Dred Scott v. Sandford, 60 U.S. (19 How.) 393 (1857).

65. For a sensitive discussion of all the difficulties of working with the concept of groups, see B. Bittker, *The Case for Black Reparations* (New York, 1973), particularly Chapters 8, 9, and 10. It should be noted that the peculiar remedial context of that discussion–the payment of money–accentuates the difficulties; some of these difficulties may be modulated in the injunctive context, where less turns on individual errors of classification. For an earlier legal literature on groupism, see Reisman, "Democracy and Defamation: Control of Group Libel," 42 *Colum. L. Rev.* 727 (1942); Pekelis, "Full Equality in a Free Society: A Program for Jewish Action," in *Law and Social Action*, ed. M. Konvitz (Ithaca, N.Y., 1950), pp. 187, 218. The comparative literature includes Marc Galanter's work on India, "Equality and 'Protective Discrimination' in India," 16 *Rutgers L. Rev.* 421 (1962) and "The Problem of Group Membership: Some Reflections on the Judicial View of Indian Society," 4 *J. of the Indian L. Institute* 331 (1962). For a discussion of the role of groups in John Rawls' work, *A Theory of Justice*, see R. Nozick, *Anarchy, State, and Utopia* (New York, 1974), p. 190; Van Dyke, "Justice as Fairness: For Groups?" 69 *Am. Pol. Sci. Rev.* 607 (1975). And for the recent sociological and psychological literature see R. Dahrendorf, *Class and Class Conflict in Industrial Society* (Stanford, Calif., 1959); M. Gordon, *Assimilation in American Life* (New York, 1964); *Ethnicity*, ed. N. Glazer and D. Moynihan (Cambridge, Mass., 1975); R. Sennett and J. Cobb, *The Hidden Injuries of Class* (New York, 1972).

stances, it may be exceedingly difficult to determine whether particular individuals are members of the group; or whether a particular collection of persons constitutes a social group. I will also admit that my definition of a social group, and in particular the condition of interdependence, compounds rather than reduces, these classificatory disputes. But these disputes do not demonstrate the illegitimacy of this category of social entity nor deny the validity or importance of the idea. They only blur the edges. Similarly, the present reality of the social groups should not be obscured by a commitment to the ideal of a "classless society" or the individualistic ethic—the ideal of treating people as individuals rather than as members of groups. Even if the Equal Protection Clause is viewed as the means for furthering or achieving these individualistic ideals (and I am not sure why it should be), there is no reason why the Clause—as an instrument for bringing about the "good society"—must be construed as though it is itself governed by that ideal or why it should be assumed that the "good society" had been achieved in 1868, or is so now.

The conception of blacks as a social group is only the first step in constructing a mediating principle. We must also realize they are a very special type of social group. They have two other characteristics as a group that are critical in understanding the function and reach of the Equal Protection Clause. One is that blacks are very badly off, probably our worst-off class (in terms of material well-being second only to the American Indians), and in addition they have occupied the lowest rung for several centuries. In a sense, they are America's perpetual underclass. It is both of these characteristics—the relative position of the group and the duration of the position—that make efforts to improve the status of the group defensible. This redistribution may be rooted in a theory of compensation—blacks as a group were *put* in that position by others and the redistributive measures are *owed* to the group as a form of compensation. The debt would be viewed as owed by society, once again viewed as a collectivity.[66]

66. See generally Bayles, "Reparations to Wronged Groups," 33 *Analysis* 177 (1973); Cowan, "Inverse Discrimination," 33 *Analysis* 10 (1972); Shiner, "Individuals, Groups and Inverse Discrimination," 33 *Analysis* 182 (1973); Taylor, "Reverse Discrimination and Compensatory Justice," 33 *Analysis* 185

But a redistributive strategy need not rest on this idea of compensation, it need not be backward looking (though past discrimination might be relevant for *explaining* the identity and status of blacks as a social group). The redistributive strategy could give expression to an ethical view against caste, one that would make it undesirable for any social group to occupy a position of subordination for any extended period of time.[67] What, it might be asked, is the justification for that vision? I am not certain whether it is appropriate to ask this question, to push the inquiry a step further and search for the justification of that ethic; visions about how society should be structured may be as irreducible as visions about how individuals should be treated–for example, with dignity. But if this second order inquiry is appropriate, a variety of justifications can be offered and they need not incorporate the notion of compensation. Changes in the hierarchical structure of society–the elimination of caste–might be justified as a means of (a) preserving social peace; (b) maintaining the community as a community, that is, as one cohesive whole; or (c) permitting the fullest development of the individual members of the subordinated group who otherwise might look upon the low status of the group as placing a ceiling on their aspirations and achievements.

It is not just the socioeconomic status of blacks as a group that explains their special position in equal protection theory. It is also their political status. The power of blacks in the political arena is

---

(1973); Sher, "Justifying Reverse Discrimination in Employment," 4 *Philosophy & Public Affairs* 159 (1975).

67. The critical temporal issue is one of duration, not whether the subordination has taken place in the past or in the future. The past has only an evidentiary relevance; it enables us to make judgments about how long the group will occupy the position of subordination. If the group has occupied the position of subordination for the last two centuries, certainly it is likely they will occupy that position for a long time in the future unless remedial steps are taken. On the other hand, if we were somehow assured that notwithstanding past history the subordination will end tomorrow, there would be no occasion for redistributive strategy on this group's behalf. Similarly, if we are told that today a period of perpetual subordination is about to begin for another group, we should be as concerned with the status of that group as we are with the blacks.

severely limited. For the last two centuries the political power of this group was circumscribed in most direct fashion–disenfranchisement. The electoral strength of blacks was not equal to their numbers. That has changed following the massive enfranchisement of the Voting Rights Act of 1965, but structural limitations on the political power of blacks still persist.[68] These limitations arise from three different sources, which can act either alternatively or cumulatively and which, in any event, are all interrelated. One source of weakness is their numbers, the fact that they are a numerical minority; the second is their economic status, their position as the perpetual underclass; and the third is that, as a "discrete and insular" minority, they are the object of "prejudice"–that is, the subject of fear, hatred, and distaste that make it particularly difficult for them to form coalitions with others (such as the white poor) and that make it advantageous for the dominant political parties to hurt them–to use them as a scapegoat.[69]

Recently, in some localities, such as large cities, the weakness of the group derived from their number has been eliminated; indeed in certain of these localities blacks may no longer be in the minority. The blacks may have a majority of a city council, or there may even be a black mayor. It would be wrong, however, to generalize from these situations. They are the exception, not the rule, and therefore should not control the formulation of a general theory of the Equal

68. For the time being, I put to one side the problems of dilution (districting that divides the group) or submersion (including the group in a larger universe through the technique of multi-member districts) since I do not have a full sense of either the prevalence of these practices or their constitutionality. See Witcomb v. Chavis, 403 U.S. 124 (1971); White v. Regester, 412 U.S. 755 (1973); Richmond v. United States, 422 U.S. 358 (1975); United Jewish Organizations of Williamsburgh v. Wilson, 510 F.2d 512 (2d Cir. 1975), cert. granted, CCH Sup. Ct. Bull. (Nov. 11, 1975) (No. 75–104, 1975 Term).

69. The quoted words are from Justice Stone's fn. 4 in United States v. Carolene Prods. Co., 304 U.S. 144, 152–53 (1938), which in part reads: "Nor need we enquire . . . whether prejudice against discrete and insular minorities may be a special condition, which tends seriously to curtail the operation of those political processes ordinarily to be relied upon to protect minorities, and which may call for a correspondingly more searching judicial inquiry." For a stimulating and illuminating discussion of the footnote, see Ball, "Judicial Protection of Powerless Minorities," 59 *Iowa L. Rev.* 1059 (1974).

Protection Clause. Moreover, these black-dominated political agencies –the black city council or the black mayor–must be placed in context. One facet of their context is the white domination of those extra-political agencies such as the banks, factories, and police, that severely circumscribe the power of the formal political agencies. Another facet is the persistent white domination of the national political agencies, such as the Congress and presidency, agencies that have become the critical loci of political power in American society.

Hence, despite recent demographic shifts in several large cities, I think it appropriate to view blacks as a group that is relatively powerless in the political arena and in my judgment that political status of the group justifies a special judicial solicitude on their behalf. When the product of a political process is a law that hurts blacks, the usual countermajoritarian objection to judicial invalidation–the objection that denies those "nine men" the right to substitute their view for that of "the people"–has little force. For the judiciary could be viewed as amplifying the voice of the powerless minority; the judiciary is attempting to rectify the injustice of the political process as a method of adjusting competing claims. The need for this rectification turns on whether the law is deemed one that harms blacks–a judgment that is admittedly hard to make when the perspective becomes a group one, for that requires the aggregation of interests and viewpoints, many of which are in conflict. It is important to emphasize, however, that the need for this rectification does not turn on whether the law embodies a classification, racial or otherwise; it is sufficient if the state law simply has the *effect* of hurting blacks. Nor should the rectification, once triggered by a harmful law, be confined to questions of fit–the judicial responsibility is more extensive than simply one of guarding against the risk of imprecise classifications by the political agencies. The relative powerlessness of blacks also requires that the judiciary strictly scrutinize the choice of ends; for it is just as likely that the interests of blacks as a group will not be adequately taken into account in choosing ends or goals. Maximizing goals such as reducing transportation costs (a goal that might account for the neighborhood-school plan) or having the most brilliant law students (a goal that might account for requiring a 650

on the LSAT) are constitutionally permissible goals in the sense that there is no substantive constitutional provision (or implied purpose lying behind some provision) that deny them to the state. On the other hand, these maximizing goals are obviously not in any sense constitutionally compelled goals and there is a chance—a most substantial one—that they would not be chosen as *the* goals (without any modification) if the interests of the blacks as a group were adequately taken into account—if the goal-choosers paid sufficient attention to the special needs, desires, and views of this powerless group.

The injustice of the political process must be corrected, and perhaps as a last resort, that task falls to the judiciary. But this claim does not yield any basis for specifying what the corrected process would look like, or what the court should say when it amplifies the voice of the powerless minority. A just political process would be one in which blacks would have "more" of a voice than they in fact do, but not necessarily one in which they would "win." In a sense there is a remedial lacuna; a pure process claim cannot determine substantive outcomes. (At the very most, it could yield those substantive outcomes that would tend to enhance the position of this group in the political process—such as favoring an increase in the numbers of black lawyers given the pivotal role lawyers play in the political process or favoring electoral districting that enhances the power of blacks as a group.) But this processual theory focusing on the relative powerlessness of blacks in the political arena need not stand alone. The substantive standards can be supplied by the other critical characteristics of this social group—perpetual subordination. The political status of the group justifies the institutional allocations—our willingness to allow those "nine men" to substitute their judgment (about ends as well as means) for that of "the people." The socioeconomic position of the group supplies an additional reason for the judicial activism and also determines the content of the intervention—improvement of the status of that group.

I would therefore argue that blacks should be viewed as having three characteristics that are relevant in the formulation of equal protection theory: (a) they are a social group; (b) the group has been in a position of perpetual subordination; and (c) the political power

of the group is severely circumscribed. Blacks are what might be called a specially disadvantaged group, and I would view the Equal Protection Clause as a protection for such groups. Blacks are the prototype of the protected group, but they are not the only group entitled to protection. There are other social groups, even as I have used the term, and if these groups have the same characteristics as blacks –perpetual subordination and circumscribed political power–they should be considered specially disadvantaged and receive the same degree of protection. What the Equal Protection Clause protects is specially disadvantaged groups, not just blacks. A concern for equal treatment and the word "person" appearing in the Clause permit and probably require this generality of coverage.

Some of these specially disadvantaged groups can be defined in terms of characteristics that do not have biological roots and that are not immutable; the Clause might protect certain language groups and aliens. Moreover, in passing upon a claim to be considered a specially disadvantaged group, the court may treat one of the characteristics entitling blacks to that status as a sufficient but not a necessary condition; indeed the court may even develop variable standards of protection[70]–it may tolerate disadvantaging practices that would not be tolerated if the group was a "pure" specially disadvantaged group. Jews or women might be entitled to less protection than American Indians, though nonetheless entitled to some protection. Finally, these judicial judgments may be time-bound. Through the process of assimilation the group may cease to exist, or even if the group continues to retain its identity, its socioeconomic and political positions may so improve so as to bring to an end its status as specially disadvantaged.[71]

70. Compare Justice Marshall's variable approach in San Antonio Ind. School Dist. v. Rodriguez, 411 U.S. 1, 70 (1973).

71. Talcott Parsons commented on the changing status of Chinese and Japanese in the United States in "Some Theoretical Considerations on the Nature and Trends of Change of Ethnicity," in *Ethnicity*, ed. N. Glazer and D. Moynihan (Cambridge, Mass., 1975), pp. 73–74. On the claim that the socioeconomic status of blacks has changed, see Wattenberg & Scammon, "Black Progress and Liberal Rhetoric," 55 *Commentary* 35 (1973) and "Letters, An Exchange on Black Progress: Ben J. Wattenberg and Richard M. Scammon and Critics," 56 *Commentary* 20 (1973).

All this means that the courts will have some leeway in identifying the groups protected by the Equal Protection Clause. I think, however, it would be a mistake to use this flexibility to extend the protection to what might be considered artificial classes, those created by a classification or criterion embodied in a state practice or statute, for example, those classes created by tax categories (those having incomes between $27,000 and $30,000, or between $8,000 and $10,000) or licensing statutes (the manufacturers of filled milk).[72] By definition those classes do not have an independent social identity and existence, or if they do, the condition of interdependence is lacking. It is difficult, if not impossible, to make an assessment of their socioeconomic status or of their political power (other than that they have just lost a legislative battle). And, if this is true, neither redistribution nor stringent judicial intervention on their behalf can be justified. It is not that such arguments are unpersuasive, but that they are almost unintelligible. Thus, in only one sense should the group-disadvantaging strategy be viewed as conducive to "more equality": it will get more for fewer. It will get more for the specially disadvantaged groups but will not provide any protection for artificial classes, those solely created by statute or a state practice. Of course, this loss may be more formal than real. Artificial classes constitute part of the universe that the antidiscrimination principle *purports* to protect, but in truth almost never does protect given the permissibility of the minimum-scrutiny inquiry.

72. What about those with income under $4,000? To some extent Justice Powell addresses this issue in his opinion in San Antonio Ind. School District v. Rodriguez, 411 U.S. 1 (1973) and he there employs a concept of social group similar to the one articulated in this essay. See also Michelman, "Forward: On Protecting the Poor Through the Fourteenth Amendment," 83 *Harv. L. Rev.* 7 (1969), suggesting the inappropriateness of using an antidiscrimination theory in this context, and attempting to shift the mode of analysis (which may take the courts beyond the Equal Protection Clause) to "just wants" and "minimum needs." For further development of this theme, see also Michelman, "In Pursuit of Constitutional Welfare Rights: One view of Rawls' Theory of Justice," 121 *U. Pa. L. Rev.* 962 (1973). On the increasing saliency of primordial rather than economic (or "class") categories, see Bell, "Ethnicity and Social Change," in *Ethnicity*, ed. N. Glazer and D. Moynihan (Cambridge, Mass., 1975), p. 141.

The Nature of the Prohibited Action

The Concept of a Group-Disadvantaging Practice. Some state laws or practices may just be a mistake—they make all groups and all persons worse off, and equally so. These do not seem to be the concern of a constitutional provision cast in terms of equality. Equality is a relativistic idea. The concern should be with those laws or practices that particularly hurt a disadvantaged group. Such laws might enhance the welfare of society (or the better-off classes), or leave it the same; what is critical, however, is that the state law or practice aggravates (or perpetuates?) the subordinate position of a specially disadvantaged group. This is what the Equal Protection Clause prohibits.

Implicit in this formulation of the prohibition of the Equal Protection Clause is a view that certain state practices may be harmful to the members of a specially disadvantaged group, and yet not impair or threaten or aggravate the status or position of the group. For example, it is conceivable that a sales tax is harmful to blacks, and yet, due to several factors—such as the diffuseness of the impact[73] and the nature of the deprivation—it need not be viewed as a practice that aggravates the subordinate status of blacks as a group. What is needed in order to bring a state practice within the equal protection ban is a theory of status-harm, one that shows how the challenged practice has this effect on the status of the group.

It is from this perspective—one of a proscription against status-harm—that discriminatory state action should be viewed. Such action is one form of state conduct that impairs the status of a specially disadvantaged group. The Equal Protection Clause prohibits the state, for example, from using race as the criterion of admission to a swimming pool or a public housing project because that practice tends to aggravate the subordinate position of blacks by excluding them from a state facility. The same is true for the dual school system—the practice of assigning students to schools (or other public facilities) on the basis of race in order to segregate them. Once again,

73. Goodman, fn. 62 above, p. 306.

this state action is prohibited by the Equal Protection Clause because it aggravates the subordinate position of blacks, not because the classification is "unrelated" or only "poorly related" to a (permissible) state purpose.

I acknowledge that in these examples the state action may also be viewed as "arbitrary discrimination." Yet it is important to emphasize that "arbitrary discrimination" is the species, not the genus. Discrimination, arbitrary or otherwise, is only one form—one form among many—of conduct that disadvantages a group. There may be group-disadvantaging conduct that is not discriminatory. This would be true of state conduct that seemed beyond the reach of the Equal Protection Clause under the antidiscrimination principle because it embodied no discrimination, racial or otherwise; for example, the state policies of enforcing all racially restrictive covenants, allocating the scarce supply of liquor licenses without regard to whether the recipients will serve blacks, closing a municipal swimming pool or other public facilities in order to avoid integration, and refusing to build a public housing project in order to limit the number of poor blacks in the community. Similarly, conduct that did discriminate but on the basis of criteria innocent on their face, such as performance on a standardized test for employment or college admission, or geographic proximity for student assignment, could be evaluated from the perspective of whether it had the effect of impairing the status of a specially disadvantaged group. There would be no need to attempt to force them into the "arbitrary discrimination" pigeonhole. The need, instead, would be to formulate a theory that linked the practice and the status of the group. To be sure, such a theory may be highly problematic; for it might well require something more than a statistical showing that by and large the practice hurts blacks more than any other group (the differential impact). To take one central example, in the context of determining whether it is permissible to use performance on a standardized test as an employment criterion, a full theory of status-harm might have to include an assessment of (a) the status (determined on several scales such as income and prestige) of the job itself (professor vs. street cleaner), (b) the public visibility of the position (a judge vs. a chemist), (c) the diffuseness

of the exclusionary impact (whether blacks are the predominant group excluded), and (d) the strength of the reasons justifying the use of the criterion (how accurate a test was it and how significant were the differences).

Admittedly, racially discriminatory conduct need not be viewed from this perspective—as a species of the genus of group-disadvantaging conduct. It could be viewed as the member of another genus, that of unfair treatment: what is wrong, it may be argued, with using race as the criterion for admission to a swimming pool or a public housing project is that it is a form of unfair treatment—an individual is being judged (for the purpose of allocating the scarce resource) on the basis of an irrelevant characteristic. The problem, however, is one of double membership: arbitrary discrimination is a member of the genus of unfair treatment as well as that of group-disadvantaging conduct. Double membership is possible because of an area of overlap of the two genuses (unfair treatment and group-disadvantaging conduct), though to be sure, the genuses are not coextensive, nor is one embraced by the other.

This analytic distinction between the two genuses is important. It preserves the possibility that conduct may be unfair and yet not a group-disadvantaging practice. Preferential treatment in favor of one of the specially disadvantaged groups would be an instance of such conduct. The white applicant who is rejected because of the preference for blacks may have been treated unfairly—may claim that he is being treated unfairly because he is being judged on the basis of an inappropriate criterion (not being black) and because the costs of a social policy are being localized on him.[74] The individual unfairness to the rejected nondisadvantaged applicant is relevant in assessing the justification of the state's refusal to institute the practice of preferential treatment; it might be relevant in fashioning a remedy, or it might even give rise to the violation of another constitutional provision, such as the Due Process Clause. But I do wish to deny that unfair treatment—such as being judged on the basis of an inappropriate criterion—is the domain of the Equal Protection Clause—

74. For elaboration of this point see my article, "School Desegregation: The Uncertain Path of the Law," 4 *Philosophy & Public Affairs* 3, 7–11 (1974).

even though such unfair treatment may be viewed from the individual perspective as a form of unequal treatment. As a protection for specially disadvantaged groups, the Equal Protection Clause should be viewed as a prohibition against group-disadvantaging practices, not unfair treatment.

Even if the claim of individual unfairness is put to one side, and the Equal Protection Clause is viewed as a protection for specially disadvantaged groups, a state policy of preferential treatment for blacks may be nonetheless constitutionally vulnerable. For one thing, it remains to be seen whether this policy in fact improves the position of the disadvantaged group. A preferential law-school admission program for blacks, to take a familiar example, may be justified on the ground that it gives positions of power, prestige, and influence to members of the racial group, positions that they would not otherwise attain in the immediate future, and that the acquisition of those positions will be an advantage to both the individual blacks admitted and, more importantly, to the group. The theory is that an increase in the number of black lawyers will disperse members of the racial group through the higher economic and social strata, raise aspirations of all members of the group, and create a self-generating protective device for the group–providing some members of the group with the power and leverage needed to protect the group from hostile attacks in the future. The status of the group will be improved. This is the theory, but, of course, these assertions of group-benefit are not totally free from doubt; indeed, some might raise the claim of counter-productivity–a regime of preferential treatment casts doubt upon the ability of all members of the group for it gives expression to the belief that the group would not succeed on its own.

If the court truly believed that a state policy–even if called "benign" –impaired the status of blacks then the policy would be invalid. But I doubt whether anyone believes that preferential admissions to law schools for blacks impairs the status of the group; those who argue that the policy is counterproductive today do so for the more limited purpose of casting doubt on the truth of the assertions supporting the policy. They wish to make the factual assertions on behalf of the policy open to controversy. And if that is so, it is important to delineate the

role of the court in such a dispute. The dispute is simply over the constitutional permissibility of the preferential policy, and from that perspective the question for the court is whether there is *some rational basis* for legislators and administrators believing that a preferential policy would benefit the group. The appropriate standard for viewing a policy that appears to the court to benefit a specially disadvantaged group should be a rational-basis standard. The judicial activism authorized by the Equal Protection Clause is, under the group-disadvantaging principle, asymmetrical.

Another possible objection to a preferential admission policy under the group-disadvantaging principle seeks to expand the universe of beneficiaries—the relevant group preferred should not be blacks, but rather the poor. The resolution of this objection might be thought to call for a judicial inquiry into what are the "true" groups in American society today—an inquiry into what group identifications are the most important to the individual, either on a psychological, political, economic, or sociological level, or which ones should be encouraged.[75] However, once again, I believe that the judicial inquiry should be a much more modest one. The court should ask whether there is any rational basis for the legislator or administrator choosing the group delineation that it did. There would be little doubt that an antipoverty strategy—an admission policy preferring the poor—would be constitutionally permissible. But that is not the issue. This particular objection to preferential treatment for blacks—the one that demands the preferred group be the poor rather than blacks—seeks to make an antipoverty strategy the only constitutionally permissible redistributive strategy. It is this constitutional strait jacket that I find is troubling and without basis in the Equal Protection Clause.

The fact that some individual blacks may identify themselves in terms of their economic position ("poor") does not deny—at least today—the reality of the racial identification—that these individuals also identify themselves as blacks or that blacks are a social group. To acknowledge the multiplicity of group identifications is not to embrace a reductionism that denies the reality of some of the groups. Nor can the reductionism be justified on the ground that the preferential

75. Bell, fn. 72 above.

policy seeks to improve the socioeconomic status of blacks. The focus on blacks (as opposed to persons who happen to have the same economic status) should be viewed as a matter of legislative or administrative perogative–a question of setting priorities. The plight of the poor may be bad, but, so the legislator or administrator should be allowed to say, not as bad as that of the blacks.[76] Such a judgment about the urgency of the situation of blacks may be rooted in two considerations. The first is the caste quality of the blacks' low status–the fact that blacks have occupied the lowest socioeconomic rung in America for at least two centuries and will continue to do so unless redistributive measures are instituted. True, we may have always had and perhaps will always have people called "the poor," but that is to confuse a stratum with the occupant of a stratum. The second consideration is that blacks face disabilities not encountered by the poor (even conceived of as a group). These disabilities manifest themselves in all spheres of life–economic, social, and political–and derive from the fact that the individuals are members of the racial group. These are disabilities that do not saddle persons who are poor. Indeed, in order to elevate themselves, the white poor have incentives to disassociate themselves from the blacks and to accentuate the racial distinction. They have incentives to make blackness the lowest status, for of necessity it is a status into which they cannot fall.[77]

Similarly, there are reasons–good ones, though not necessarily compelling–for the legislators or administrators to treat blacks as a single group without trying to sort out the "rich blacks," without trying to fractionate the group. One reason is administrative convenience–the likely number of rich blacks are so few, and the costs of the mechanisms needed to identify them are so high, that the sorting is not worth the effort. Under the antidiscrimination principle, the central evil was one of loose fit and all arguments of administrative convenience were suspect; but under the group-disadvantaging principle, imprecision is not itself a constitutional vice. Imprecision can be

76. See the penetrating article of Duncan, "Inheritance of Poverty or Inheritance of Race?" in *On Understanding Poverty: Perspectives from the Social Sciences*, ed. Moynihan (New York, 1968).

77. Parsons, fn. 71 above, p. 77.

tolerated when the state law or practice seeks to improve the position of a disadvantaged group and is in fact related to that end. Moreover, wholly apart from considerations of administrative convenience, the decision not to exclude the rich black (even once identified) can be justified. The argument is not that rich blacks as individuals are as "entitled" (in the compensatory sense) to as much of an assist as persons who are poor (though that may entirely be possible, for even though these individuals are rich, they are still black and being black may have been as severe a disadvantage in our society as being poor, if not greater). Rather, the claim is that the preference of the rich blacks may be justified in terms of improving the position of the group. Even if the blacks preferred happen to be rich, a benefit abounds to the group as a whole. Members of that group have obtained these positions of power, prestige, and influence that they otherwise might not have and to that extent the status of that group is improved. On the other hand, it is not clear that preferring a poor person confers a benefit on the poor conceived as a group—the preferred individual merely leaves the group; and even if there were group benefits entailed in a preference for the poor, certainly legislators or administrators are entitled to rank the improvement of blacks as a group as a social goal of first importance, more important than elevating the poor conceived of as a social group.

Finally, the preferential admission program for blacks (or any other single disadvantaged group) may be thought vulnerable because of its impact on other disadvantaged groups, whether they be American Indians, Chicanos, or perhaps even the poor, the poor black, or the poor black women (if these latter categories can be considered discrete disadvantaged social groups). The adverse impact on these groups arises in a two-step fashion and is ultimately traceable to the fact of scarcity: starting with a fixed number of openings, the preference for blacks lessens the number of places available to the members of these other groups. What is given to one group cannot be given to another.

There are several lines of response to this particular objection. One might argue, for example, that the nonpreferred group is not as badly off as the preferred group (for example, Chicanos are not as badly off

as blacks). Another might emphasize the indirect quality of the exclusion (exclusion occurs because blacks are preferred and the places are limited), a factor that has an important bearing on the question whether the preferential treatment gives rise to a status-harm to the nonpreferred disadvantaged group. Preferring blacks may limit the number of places open to other disadvantaged groups, but it is not clear that it impairs their status. Finally, the objection might be answered by invoking the standard defense to underinclusion—one step at a time. Under the antidiscrimination principle, this defense was troubling because it seemed to threaten the claim of underinclusiveness; the defense could be raised to every instance of underinclusiveness. With the group-disadvantaging principle, no particular importance is attached to the claim of underinclusiveness itself, and thus the defense is less threatening; and in any event, it is conceivable that substantive standards could be developed to limit the defense and to thereby maintain as much pressure on the political agencies to remain responsive to the demands of all specially disadvantaged groups.

Having identified these possible lines of defense to this particular objection to preferential treatment for blacks, I do not want to obscure its force. The harm to other disadvantaged groups that arises from the fact of scarcity would be a significant point of constitutional vulnerability for a restricted preferential policy. It invokes the prohibition against group-disadvantaging practices and in responding to this objection, unlike the others, no help can be derived from the asymmetrical quality of judicial activism. The claim is that specially disadvantaged groups are being hurt, and the group-disadvantaging principle requires the court to be particularly attentive to such claims. But perhaps it is correct that a *restricted* preferential policy should be constitutionally vulnerable; and that it should be vulnerable precisely for this reason—because of its impact on other disadvantaged groups rather than because of its impact on the dominant group—the rejected white males. This is more than a purely analytic point, for it indicates what must be done to save the policy constitutionally—extend it to those specially disadvantaged groups that are as entitled to the preferential treatment as blacks.

The Accommodation of Considerations of Total Welfare. The overriding concern is with the status of the specially disadvantaged groups, and any state practice which aggravates their subordinate position would be presumptively invalid. There is, however, an element of restraint, and that arises from the fact that certain group-disadvantaging practices further interests of the polity as a whole, some of which are material (for example, increased productivity) and others nonmaterial (for instance, increased individual liberty). There must be some accommodation of other such competing interests. I doubt, for example, whether the Equal Protection Clause should be construed so stringently as to deny the state the right to insist upon certain minimum levels of proficiency or competence for its employees or students, even if such insistence were to aggravate (or at least perpetuate) the subordinate position of blacks. The problem, of course, is that if too large a role is allowed these considerations of total welfare, the protectionist edge of the Equal Protection Clause would be severely dulled: the state could always defend against an equal protection claim on the ground that, even though a disadvantaged group is especially harmed, total welfare is being maximized.

Two analytic tools must be introduced in order to avoid this dilemma. The first is the concept of a nonallowable interest. An example would be the interest of whites to keep blacks in a subordinate position. That interest should be given no weight in determining whether the group-disadvantaging practice is justified. In that sense it has no normative weight, although account of it might have to be taken out of sheer necessity. It *should* not be taken into consideration, but it might *have to be* because the court has no choice. For example, the resistance to the busing decree may be so intense as to be expressed in open rebellion, and in that case the court might have to proceed more slowly.[78] Resistance is an obstacle to be reckoned with, though there are limits to what can be done.

The concept of a nonallowable interest corresponds roughly to the concept of an illegitimate state purpose under the antidiscrimination

78. See Hart v. Community School Board No. 21, 383 F. Supp. 699 (E.D.N.Y. 1974); Cooper v. Aaron, 358 U.S. 1, 16-17 (1958); Buchanan v. Warley, 245 U.S. 60, 81 (1917).

principle. The difference relates not so much to the content of what is allowed or illegitimate; rather the standard for legitimacy or allowability is directly anchored in the governing principle. The antidiscrimination principle talks in terms of fit, in terms of over- and underinclusiveness, and such a principle can provide no standard for determining the legitimacy of state purposes; the antidiscrimination principle had to go beyond itself, and in that sense is incomplete, not in itself fully intelligible. The group-disadvantaging principle, on the other hand, does itself provide a standard of legitimacy or allowability. Interests should not be allowed when they would effectively give the dominant group a veto power over the elevation of the specially disadvantaged group. Such a veto would be inconsistent with what I perceive to be the very purpose of the Equal Protection Clause and with the notion of a constitutional restraint.

The second analytic tool that must be introduced to prevent the weighing process from degenerating into no-protection is the concept of a compelling benefit. This concept structures the relationship between the harm to the group's status and the benefit to the polity. If some compelling benefit is to be obtained by the practice or statute, either because of the importance of the interest served or the size of the benefit, the practice will be permitted, notwithstanding the status-harm to the specially disadvantaged group. The required benefit is linked to the harm: the more severe the disadvantage, the more compelling the benefit must be. But in linking the two, it is important to emphasize that the harm caused the disadvantaged class by the practice need not be greater than the benefit to the polity. That is why, in thinking of the Equal Protection Clause as a redistributional device, the balancing metaphor seems inapt: relief would be granted even if the harm to the group were less than the benefit to society. The more appropriate metaphor would be one that conceives of the status-harm to the disadvantaged group as the trigger of the remedial effort with the legitimate interests of the polity serving as the restraining force. The restraint starts once the benefit is greater than the harm, but only when it reached a certain quantity and intensity—denoted by the term "compelling"—would the remedial effort be brought to a halt.

In the first instance the concept of compelling benefit will require the court to ask if there are alternative ways available to society for furthering its interests, and whether these alternatives are less harmful to the disadvantaged group. If so, the practice will be invalidated, for the state practice is not, in the ordinary meaning of that word, compelled. If, however, there are no alternatives that are less disadvantaging, then the court must move to a second order inquiry: it must gauge the status-harm to the disadvantaged group and the benefit to society and, in the final analysis, determine whether the benefit is of a compelling quality.

It might be noted that the first step of this inquiry is similar to that required by the strict-scrutiny branch of the antidiscrimination principle: strict scrutiny requires a search for better alternatives. But the difference is four-fold. First, we have an explanation for the strictness of the scrutiny, and one that is contained within the principle —the scrutiny is stringent because the status of an already subordinated group is being threatened. Under the antidiscrimination principle, the strictness of the scrutiny depends on a judgment about which interests are fundamental and which criteria are suspect; and although it might be possible to explain and justify those choices, the explanation would take the court far beyond the foundational concept of the antidiscrimination principle—means-end rationality—and thus transform the antidiscrimination principle so as to make it inconsistent with the ideals it is supposed to serve. Second, the action that triggers the inquiry differs. Under the group-disadvantaging principle, it is harm to a specially disadvantaged group, not the use of a criterion (somehow) deemed "suspect" or impingement on an interest deemed "fundamental." Third, there is a difference in the standard of what is a "better" alternative. Under the group-disadvantaging principle the alternative is judged not in terms of fit (reduction of under- and overinclusiveness) but rather in terms of its status-harm to the disadvantaged group. Fourth, the inquiry goes on to a second level if there is no better alternative. Under the antidiscrimination principle, the inquiry stops once it is determined there is no closer fitting or more precise criterion; the statute or practice is validated. But under the group-disadvantaging principle, even if the judge determines there is

no less harmful way of satisfying the purpose, he must still gauge the harm to the group and the benefit to society. He must determine whether the benefit to society is so important or so great as to make the status-harm to the specially disadvantaged group tolerable—an inquiry that has no meaning under the antidiscrimination principle, and one that is inconsistent with the assurance that the principle exclusively rests on the modest conception of means-end rationality.

The State Action Requirement. I identified the state action requirement of the Fourteenth Amendment as one of the persistent sources of difficulty with the antidiscrimination principle. The problem to me, seemed to me to stem not so much from the "state" component, but from the "action" component, or more specifically, from the requirement that the "action" be a form of forbidden "discrimination." This artificial conception of the requisite "action" prevented the courts from tracing the effects of the state law or practice on the status of various disadvantaged groups in society, such as the blacks. Under the group-disadvantaging principle, however, this artificial limitation is eliminated.

In removing this limitation I do not mean to suggest an abandonment of the traditional view that the prohibition of the Equal Protection Clause applies only to the conduct—the laws and practices—of the state. Due account has been taken of this view. It is the state that is prohibited from aggravating the subordinate position of specially disadvantaged groups. There is a state action requirement. The most troublesome question that is likely to arise is whether state "inaction" shall be treated as "action." Should the failure of the state to *initiate* redistributive measures on behalf of specially disadvantaged groups—to counteract the inequalities imposed by private activity—be viewed as a practice that aggravates the position of subordination and thus as a violation of the Equal Protection Clause? As a purely analytic matter, it is possible to answer that question in the affirmative. But, on the other hand, there are three factors that argue for a negative answer, for preserving the distinction between action and inaction.

One is a concern for the text. Professor Black has argued that the words "denial" and "protection" suggest to him an affirmative obliga-

tion, an obligation to throw a life preserver.[79] On the other hand, without the distinction between action and inaction, the words "state" and "laws" in the Clause would become superfluous. The line between individual action and governmental action would be obliterated; the Clause would oblige the state to enact laws counteracting private group-disadvantaging practices and thus, in that sense, private action would be covered by the Clause.

The conception of constitutional prohibitions as restraints on the use of governmental power is a second factor that suggests that the distinction between action and inaction be preserved. This negative conception has historical roots. I suspect that this is probably how the framers viewed their task in drafting the first section of the Fourteenth Amendment (in contrast to those provisions of the Constitution establishing the structure of government). I also think the negative conception would be the one most consistent with the goal—of some undeniable validity—to minimize intervention by the judiciary (the agency primarily entrusted to enforce the Equal Protection Clause). If inaction were viewed as action, the intervention by the judiciary would be enormously increased, if not endless. This degree of intervention might be at odds with our democratic traditions, even tempered by a concern for minority rights, and it might push the judiciary beyond the limits of its competence. There is some awkwardness in having the judiciary act as the primary redistributive agency, in large part because it does not set its own agenda and thus cannot rationally order its priorities: should it be bread rather than housing?[80] A court must depend on the choices set by the litigators and I suspect that the resulting pattern of decision would tend to over emphasize those redistributive measures connected to the criminal process (for example, providing free transcripts on appeal). The criminal defendant always has the incentive to litigate, and society facilitates that proclivity by subsidizing litigation costs.

The problems of fashioning an appropriate remedy also bear on

79. Black, fn. 50 above, p. 73.

80. See Winter, "Poverty, Economic Equality and the Equal Protection Clause," 1972 *Sup. Ct. Rev.* 41, 89.

the action/inaction issue. Of course, even if the state action limitation is taken seriously, and the distinction between action and inaction preserved, acute remedial problems can arise. It is these factors, above all, that might persuade a court not to intervene when, for example, a town decides to close the swimming pool or to close its public schools in order to avoid integration. I have little doubt that such action disadvantages blacks, and yet the difficulty of fashioning an effective decree may be so great as to lead the court to deny injunctive (though perhaps not declaratory) relief. These negative decisions of the state come very close to no decision (the pure state inaction category), but once the line is crossed, and the court moves from action to inaction, these problems of fashioning an effective remedy are compounded.[81] The court would have to imagine a universe of possible measures that might be taken (by the legislature) on behalf of the disadvantaged group, make a choice among the permissible ones, compel its enactment, and then police the police.[82] This is not an impossible task but it is exceedingly treacherous, and perhaps further reason for limiting the group-disadvantaging principle to actions of the state.

V. The Choice of Principle

In many situations it will not make a great deal of difference whether the court operates under the antidiscrimination principle or the group-disadvantaging one. An example of such a situation—which I would like to call "first-order"—would be one in which the state excludes blacks from public institutions. In these first-order situations, which were the focus of judicial attention up to the late 1940s, and the following decade or two, the same result is likely to flow from either

81. It is hard to know on which side of the line to place a decision not to enter a field all together, for example, not to build public housing. See James v. Valtierra, 402 U.S. 137 (1971).

82. In addition to enjoining the legislature and compelling the enactment of an ameliorative measure, a court could act negatively. For example, instead of ordering the state legislature to enact a law prohibiting operators of public accommodation from refusing to serve individuals because of their race, it could overturn to trespass convictions of those who were not served because of their race.

principle. I would still prefer the group-disadvantaging principle on the ground of frankness–it more accurately captures the intellectual process that should go on in the mind of the judge. It is nevertheless hard to believe much turns on the choice of principle.

Today, however, we find ourselves beyond these first-order situations. A new situation arises when the court is confronted with challenges to nondiscriminatory state action (such as conferring liquor licenses without regard to admission practices or closing public facilities) and with challenges to the state use of facially innocent criteria (such as test performance for allocating jobs or college places). With these second-order situations, there is more than frankness to recommend the group-disadvantaging principle. I believe this principle will frame matters in such a way as to expose the real issues and thus be more likely to lead to the correct decision–invalidation of those state practices that aggravate the subordinate position of the specially disadvantaged groups. It is, of course, possible that under the antidiscrimination principle a court willing to stretch and strain could reach the same result as it would under the group-disadvantaging principle; but that seems either to be a fortuity, or to require such a modification of the antidiscrimination principle–as evidenced by the "past discrimination," "de facto discrimination," or "fundamental right" offshoots–as to deprive the principle of any intellectual coherence and transform it into something it was never intended to be. In these situations, the group-disadvantaging principle should be preferred because it has a degree of coherence and completeness that can never be achieved with the antidiscrimination principle.

There is, to be sure, a third-order situation–as exemplified by state preferential treatment of blacks. In this instance more turns on the choice of principle than frankness, increased likelihood of the right result, or formal elegance. With these third-order problems there is a genuine conflict of principles. The antidiscrimination principle, with its individualistic, means-focused, and symmetrical character, would tend toward prohibiting such preferential treatment; the group-disadvantaging principle, on the other hand, would tend toward permitting it (and indeed might even provide the foundation for the fourth-order claim that may lie around the corner–that of requiring

the preferential treatment). I believe this conflict should be resolved in favor of the group-disadvantaging principle but conceivably that this preference need not be one that has preemptive effect (all the way back down the scale). The antidiscrimination principle has coexisted with at least one other equal protection principle, the numerical-equality-of-persons principle of *Reynolds v. Sims*; and, similarly, the group-disadvantaging principle can coexist with other principles.[83] If that be so, then the group-disadvantaging principle can be viewed as a supplemental one: the use of the principle becomes more important as one moves from the first order to the second and then on to the third; and when one arrives at the third order, a hierarchy of the principles must be constructed; and when there is a conflict of principles, the principle of first priority takes precedence. I would argue that in this hierarchy the group-disadvantaging principle should be the one of first priority, should be placed ahead of the antidiscrimination principle. But I would be the first to acknowledge that this choice is not an easy one.

Part of the difficulty of making the choice stems from the fact that the usual material of judicial decisions—legislative history and text—provides no guidance. History indicates that the Clause was extended

83. If the group-disadvantaging principle did have a preemptive effect (all the way down the scale), it would knock-out the antidiscrimination principle as it now exists, that is, with its elaborate superstructure. It might, nevertheless, be possible to find a place in the Constitution—either in the Equal Protection Clause or perhaps more appropriately in the Due Process Clause—for a stripped-down version of the antidiscrimination principle—that is, simply a guarantee of means-end rationality. A law that bore no relation (or very little relation to) an end would be thought "arbitrary." (It might even be possible to append to this basic guarantee of means-end rationality a requirement that the end be constitutionally permissible, though the standard of permissibility would derive from other provisions, such as the First Amendment, rather than the Equal Protection Clause itself—now viewed as a protection for specially disadvantaged groups.) This basic guarantee of means-end rationality would sometimes enable a court to invalidate a state law that, for example, curtailed in a strikingly underinclusive way the speaking privilege of some particular individual or individuals who are not members of a disadvantaged group. The court could reach that result under a means-end test without passing on the general validity of the particular end chosen (for example, to prevent incitement), as it might have to if exclusive reliance were placed on a substantive constitutional provision (such as the First Amendment).

to protect blacks. But it does not tell us whether blacks were to be viewed as a group or as individuals, nor does it say much about the intensity or degree of protection that is to be afforded. Similarly the text does little more than give the ideal of equality constitutional status and circumscribe it with a state action requirement. It is hard to believe that the use of the word "person" was intended to foreclose the recognition of the importance and status of groups. In essence, the text clothes the court with the authority to give specific meaning to the ideal of equality—to choose among the various subgoals contained within the ideal. A judge must become a natural lawyer out of default. The ethical issue is whether the position of perpetual subordination is going to be brought to an end for our disadvantaged groups, and if so, at what speed and at what cost.

The antidiscrimination principle roughly corresponds to the lay concept of equal treatment, and some might argue that the antidiscrimination principle should be given priority (and perhaps preemptive effect) because equal treatment is a more widely accepted goal of personal and social action (or more in accord with traditional American values, such as individualism). But this argument seems wrong, even if the informal Gallup Poll came out as imagined. It is not the job of the oracle to tell people—whether it be persons on the street or critical moralists—what they already believe.

For one thing, the public morality may be only an echo: the concept of equal treatment may be the more widely accepted subgoal of the ideal of equality because it more nearly accords with the concept of equality previously propounded by the Supreme Court and because it is the one embodied in the law. The Equal Protection Clause provides the Court with a textual platform from which it can make pronouncements as to the meaning of equality; it shapes the ideal. There pronouncements are viewed as authoritative, part of the "law," and play an important—though by no means decisive—role in shaping popular morality.[84] Law is a determinant, not just an instrument, of

84. The impact of *Brown v. Board of Education* on popular morality, especially of those growing up in the 1950s and 1960s, is ample testimony of this phenomenon. So is the wait in the spring of 1974 for the Supreme Court's decision in *DeFunis*. More seemed to be at stake than a directive against the University of Washington Law School, or for that matter all other state schools.

equality. Of course, this relationship between law (viewed as pronouncement rather than directive) and popular morality does not deny the existence of the latter; an echo is still a sound. But it does mean that the group-disadvantaging principle may also be widely accepted once it too is propounded to be the chosen strategy of the Supreme Court, once the judiciary says that this principle and the concept of equal status has an important claim to the constitutional ideal of equality.

Moreover, deference to the prevailing popular morality seems particularly inappropriate when what is being construed is a constitutional protection for minorities. Our commitment to democracy might dictate a reference to the people for the adoption and amendment of the Constitution, but once those processes are complete, a second reference to the electorate to elaborate the content of the ideal embodied in the Constitution seems inconsistent with the very idea of a constitutional restraint. This is particularly true of one that was in large part intended to protect a racial minority. All constitutional restraints are to be countermajoritarian, and this one particularly so.

It might be contended that the priority between the two principles should be set, not on the basis of text, history, or a rough sense of popular morality, but rather on the basis of certain institutional values —which strategy would best further the ideals of the craft. As we have seen, the predominance of the antidiscrimination strategy could in part be explained in terms of its supposed institutional advantages, objectivity and value-neutrality. Supposedly, judges will not be called on to make judgments about ends. The lines that emerge will be sharp. The decisions of the courts will not be heavily steeped in factual inquiries. In contrast, under the group-disadvantaging interpretation, the courts must deal with highly speculative entities, social groups. Subtle factual inquiries are required as to the contours and status of the group and the impact of a challenged practice on the group. Invariably value judgments would have to be made as to the costs and benefits of the practice to the disadvantaged group and to the polity.

People looked to the Court for some guidance in the solution of an intractable moral problem, and this guidance was supposed to emerge from the Court's decision on the meaning of the Equal Protection Clause.

I am willing to assume that the group-disadvantaging strategy will strain the resources, the imagination and even the patience of the judiciary. From the perspective of "mechanical jurisprudence" the group-disadvantaging principle offers no advantages. But I doubt whether these institutional considerations ought to be the bases for the choice between principles. For one thing, as we saw, this image of what judicial life will be under the antidiscrimination principle—no value judgments, sharp lines and no factual judgments—is largely illusory. The court must make determinations about whether the purpose served by the classification is "legitimate," which classifications are "suspect," what rights are "fundamental," what purposes are special or ordinary, whether it is permissible to take one step at a time, and whether the "fit" is sufficient. The quantitative ring to the terms "fit," "overinclusion," and "underinclusion" is decidedly an illusion. Moreover, once the antidiscrimination strategy is modified to embrace "the perpetuation of past discriminations" or "de facto discrimination," as I believe it must, the factual inquiries become overwhelming and the value judgments used (for example, in determining which effects of past discrimination are to be eliminated) become commonplace.

In any event, even if it can (somehow) be demonstrated that the antidiscrimination principle is more conducive to the traditional ideals of the craft, it still remains to be seen why these ideals—the ideals of "mechanical jurisprudence"—should be preserved at all[85] or at least at the expense of substantive results deemed just. It is understandable why judges will choose that strategy most in accord with the ideal of their craft but that hardly makes it just, nor, for the self-conscious judge, inevitable. The redistributive aims served by the group-disadvantaging principle—the elevation of at least one group that has spent two centuries in this country in a position of subordination—may simply override these supposed institutional advantages.

Finally, even if these institutional arguments provide a basis for preferring the antidiscrimination principle, it must be remembered that at best they dictate a choice of the judicial strategy. These arguments cannot be used by all. A sharp line should be drawn between the principles governing the judicial process and those that govern

85. This view is elaborated in the article referred to in fn. 22 above.

nonjudicial processes, some legislative, some private. This is one sense in which law and morals should be separated. For the institutional considerations have little relevance for those who do not act as judges, but nonetheless must struggle with the task of giving meaning to the ideal of equality—for example, citizens deciding upon admissions policies to a law school and legislators acting under Section 5 of the Fourteenth Amendment. Nor should these institutional considerations be allowed to operate subsilentio in nonjudicial spheres, a danger created when the citizenry gives excessive deference to the judicial pronouncement as to the meaning of equality—when sight is lost of the reason for the choice of strategy, and the citizenry defers to judicial pronouncement as to the meaning of "equality" simply because it is "The Law," or even worse, "The Constitution."

The roots of such excessive deference are deep. It may reflect the psychological need for an authoritative agency to decide questions of individual morality. The need becomes particularly acute when, as is true here, the ethical questions become more difficult, the arguments on each side more balanced, and when it is not just a conflict between liberty and equality, but in essence a conflict between two important senses of equality—equal treatment and equal status. Moreover, when what is demanded is a nationwide morality, one that could subordinate the morality of a particular region, as has been true in racial matters, the Supreme Court is particularly well positioned to perform the function of such an authoritative decision-making agency. The impact of judicial pronouncement on positive morality may also be traced to the strategic position of lawyers in our society, as managers of important nonlegal institutions and as formulators of opinion. The Court is "their" institution and they are bound to look to it in formulating their conception of equality.[86]

But whatever the reasons for this deference, and regardless of how understandable it might be, what strikes me as important is that we should become increasingly aware of the role of judicial pronounce-

86. Indeed a person committed to the Court as a legal institution—because of constitutional structure or results in particular cases—might be willing or even desirous to have the Court's sphere of influence broadened as a means of buttressing or fortifying its legal position.

ment in translating the ideal of equality, the nature of the choices the courts have made, and the explanation for the choices. We should become aware of the fact that the antidiscrimination principle is not inevitable and, indeed, that its predominance may be traceable to institutional values that have little relevance for individual morality or legislative policy. At best they explain the choice of judicial strategy. This increased consciousness does not necessarily mean that the group-disadvantaging principle must be adopted as a matter of individual morality or legislative policy. Indeed considerations that might have given little weight in determining what should be the correct interpretation of the Equal Protection Clause–such as the consideration of individual fairness–may play a larger role in determining what is right for a citizen or legislator. Old dilemmas may appear in new guises.