ARTICLES

THE TROUBLE WITH CLASSIFICATIONS

*Aziz Z. Huq**

The Supreme Court relies increasingly on anticlassification rules to implement the Constitution's various commands of evenhanded state treatment. These rules direct attention to whether an instance of a forbidden classification is present on the face of a challenged law. They contain two necessary steps. First, a court defines a general category of impermissible terms. Second, a court ascertains whether an instance of the category is found in enacted text—so triggering the familiar strict scrutiny analysis. So defined, anticlassification rules now dominate equal protection, free speech, "dormant" Commerce Clause, and even free exercise jurisprudence. The Roberts Court celebrates these doctrinal tests as "commonsense," citing their administrability and mechanical quality as safeguards against problematic judicial discretion.

This Article challenges this account of anticlassification rules as simple and transparent. It draws extensively on conceptual tools from the philosophy of language to elucidate the inherent complications and internal tensions of the doctrine. Defining and drawing bounds around categories such as "race" and "content discrimination," for example, cannot be done without a theory of what philosophers of language call "natural kinds" and "social kinds." Yet when courts identify instances of impermissible categories in legal text, they tend to fluctuate erratically between semantic and communicative theories of meaning. A careful examination of these, and other, hidden premises of anticlassification clarifies apparent doctrinal inconsistencies. Absent a systematic theorizing of such difficulties, anticlassification rules cannot be coherently or consistently applied. Reckoning with these difficulties suggests that the Court's main normative justifications for anticlassification have a narrower reach than commonly appreciated.

* Frank and Bernice J. Greenberg Professor of Law, University of Chicago. This research was supported by the Frank J. Cicero Fund. I am grateful to the editors for their careful work on this article.

INTRODUCTION

In the opening moments of oral argument in *Students for Fair Admissions, Inc. v. President & Fellows of Harvard College* (*SFFA*),[1] the constitutional challenge to private and public universities' use of race in admissions fleetingly seemed to stumble. Three Justices aired what seemed a grave difficulty in petitioner's case: Did Harvard College and the University of North Carolina (UNC), in fact, classify by race? If so, what did that entail? First, Chief Justice Roberts asked if consideration of an "essay about having to confront [racial] discrimination" counted as such.[2] Justice Kagan next pressed on whether applicants' "own racial experiences" could be legitimately considered.[3] Piling on, Justice Barrett questioned whether admissions could account for "racial identity . . . as a source of pride."[4] The gist of these questions was that race might play a role in universities' admissions processes without counting as a problematic racial classification. Deflecting these inquiries, petitioner's counsel, Cameron Norris, offered a simple formula: universities may do these things, but they "just cannot consider race itself."[5] And when the Court finally resolved the case in June 2023, Chief Justice Roberts laid down that rule, unadorned by further explanation: universities that "can and do[] take race into account" employ a constitutionally suspect "racial classification."[6] The difficulties of oral argument, it seemed, had vanished into air—albeit without explanation.

The puzzles raised by Chief Justice Roberts, Justice Kagan, and Justice Barrett deserve closer attention. Across a range of constitutional contexts, the Supreme Court increasingly uses some sort of *anticlassification rule* to implement the Constitution's demands for even-handed treatment. Anticlassification rules isolate the question whether instances of an impermissible category occur in the verbal formulation of a binding enactment. At the same time, they promise to exclude fickle judicial consideration of either intentions or purposes. Anticlassification rules are now prominent in the Equal Protection Clause of the Fourteenth Amendment,[7] as the affirmative action

---

1 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141 (2023).

2 Transcript of Oral Argument at 7, *Students for Fair Admissions*, 143 S. Ct. 2141 (No. 20-1199).

3 *Id.* at 8.

4 *Id.* at 10, 9–10.

5 *Id.* at 11.

6 *Students for Fair Admissions*, 143 S. Ct. at 2155, 2162 (first citing Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 397 F. Supp. 3d 126, 144 (D. Mass. 2019); and then citing Grutter v. Bollinger, 539 U.S. 306, 326 (2003), *abrogated by Students for Fair Admissions*, 143 S. Ct. 2141).

7 U.S. CONST. amend. XIV, § 1.

decision shows. They loom large, too, under the Free Speech Clause of the First Amendment[8] and the "negative" or "dormant" implications of the Commerce Clause.[9] In effect, the Court turns to anticlassification rules to vindicate all constitutional nondiscrimination commands.

To many of the Justices, widespread adoption of anticlassification rules seems just a matter of "common[ ]sense."[10] Such rules seem a paragon of simplicity. They are parsimonious to state and straightforward to apply. They are buffered from the distorting tug of judicial bias. Their candid transparency sustains the rule of law.[11] Anticlassification rules further advance a broader jurisprudential commitment to formalism by the Roberts Court.[12] (Formalism is understood as a fidelity to "the literal mandate[] of the most locally applicable legal norm" to the exclusion of "a wider range of factors,"[13] such as legislators' intentions, administrators' behavior, or downstream consequences.)

This Article challenges this "commonsense"[14] view of anticlassification rules. By carefully mapping how the Court uses and justifies anticlassification rules, the Article reveals a series of difficult, and yet unavoidable, normative and analytic choices embedded within them. My aim is to explicate the hidden clockwork of anticlassification. For this purpose, I draw several technical terms from the philosophy of language. These include distinctions between "natural kinds" and "social kinds," [15] between "exemplar" and "prototypical" theories of

8 *Id.* amend. I.

9 *Id.* art. I, § 8, cl. 3.

10 *See* Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015) (citing Sorrell v. IMS Health Inc., 564 U.S. 552, 564 (2011)); *see also* City of Austin v. Reagan Nat'l Advert. of Austin, LLC, 142 S. Ct. 1464, 1482 (2022) (Thomas, J., dissenting) (quoting *Reed*, 576 U.S. at 163).

11 A common understanding of the rule of law demands that "all persons and authorities within the state, whether public or private, should be bound by and entitled to the benefit of laws publicly and prospectively promulgated and publicly administered in the courts." Lord Bingham, *The Rule of Law*, 66 CAMBRIDGE L.J. 67, 69 (2007).

12 *See, e.g.*, Stephen M. Feldman, *Free-Speech Formalism Is Not Formal*, 12 DREXEL L. REV. 723, 740 (2020) (noting the emergence of "a formal concept of free expression" in the Roberts Court); Adam B. Cox & Emma Kaufman, *The Adjudicative State*, 132 YALE L.J. 1769, 1778 (2023) (tracing a "through line in [recent structural constitutional] doctrinal developments": "a commitment to separation-of-functions formalism").

13 Frederick Schauer, *Formalism*, 97 YALE. L.J. 509, 522 (1988).

14 *See Reed*, 576 U.S. at 163.

15 *See* Ian Hacking, *A Tradition of Natural Kinds*, 61 PHIL. STUD. 109, 110–11 (1991) (setting out criteria of natural kinds); Rebecca Mason, *The Metaphysics of Social Kinds*, 11 PHIL. COMPASS 841, 841–42 (2016) (compiling different definitions of "social kinds"). There is debate on whether social and natural kinds are mutually exclusive categories. *Id.* at 841.

categorization,[16] between the "sense" and the "reference" of words,[17] and between "semantic" and "pragmatic" theories of meaning.[18] Leveraging these tools from the philosophy literature, the Article aims to develop a lucid, tractable framework for evaluating when and how anticlassification rules can and should be used in law.

To be clear at the threshold, my aim is not to show that the Court must abandon all anticlassification rules in constitutional law. Rather, by casting light on necessary technical choices impelled by anticlassification's logic, I rather "propound a framework for meaningful debate"[19] about their legitimate deployment. At present, anticlassification rules beg more questions than they answer. Controversial normative choices are ignored. The Justices oscillate erratically between rule-like and standard-like versions of anticlassification. Once underlying normative choices are pulled out into the cold light of day, I believe, the decision whether to retain or abandon anticlassification turns out to hinge on hidden normative choices. Because anticlassification cannot and does not advance several of its notional goals, my analysis further suggests that it should be used in more carefully tailored ways that account better for technical and linguistic complexities.

To motivate an inquiry into anticlassification, it is useful to offer at the start some examples of why anticlassification regimes are not quite as simple as they first seem The oral argument in the 2023 *SFFA* case surfaces two of the several analytic difficulties explored in this Article. These two points illustrate—but do not exhaust—the complex choices immanent in anticlassification regimes.

First, the initial questions from Chief Justice Roberts, Justice Kagan, and Justice Barrett about what counts as consideration of race nicely illustrate how a forbidden classification can turn up during the

16 *See* J. David Smith, Alexandria C. Zakrzewski, Jennifer M. Johnson, Jeanette C. Valleau & Barbara A. Church, *Categorization: The View from Animal Cognition*, BEHAV. SCIS., June 2016, at 1, 1 (identifying exemplar, prototype, and rule-based models of categorization in animals); *see also* Eliot R. Smith & Michael A. Zarate, *Exemplar and Prototype Use in Social Categorization*, 8 SOC. COGNITION 243, 243 (1990) (applying these frameworks to human cognition of groups).

17 This traces back to the work of Gottlob Frege, a German mathematician and logician from the turn of the century. *See* Gottlob Frege, *Über Sinn und Bedeutung*, 100 ZEITSCHRIFT FÜR PHILOSOPHIE UND PHILOSOPHISCHE KRITIK 25, 49 (1892) (Ger.). Future references are to Gottlob Frege, *Sense and Reference*, 57 PHIL. REV. 209 (1948).

18 Robyn Carston, *Legal Texts and Canons of Construction: A View from Current Pragmatic Theory*, *in* 15 LAW AND LANGUAGE: CURRENT LEGAL ISSUES 8, 9 (Michael Freeman & Fiona Smith eds., 2013) (distinguishing pragmatic presupposition from semantic presupposition).

19 Richard H. Fallon, Jr., *A Theory of Judicial Candor*, 117 COLUM. L. REV. 2265, 2271 (2017).

process of rule application even if the verbal specification of a rule does not itself mention or pick out that category.[20] As the Justices intimated, an admissions protocol might ask an admissions officer to account for an applicant's beliefs, achievements, or sources of pride.[21] Its text would not mention or direct attention toward race as such. Nevertheless, as the Justices' questions drew out, such questions could (and likely do) predictably elicit responses in which race plays a role today. Given this predictable outcome, why shouldn't the rule trigger anticlassification concern?[22] Perhaps the "opaque"[23] relationship between the rule's content and the suspect criterion dilutes any constitutional concern.

But why? Consider how a similar question is handled in related statutory contexts, such as Title VII of the Civil Rights Act. Applying that law, the Court asks whether an adverse action would have "taken place *but for*" a protected trait.[24] By analogy, when an admissions rubric directs attention to the way a successful candidate has overcome hurdles, including race-related ones, isn't race a but-for cause of their admission? Shouldn't an anticlassification rule align with the but-for test for discrimination?

Second, Justice Kavanaugh surfaced another difficulty hiding underneath the placid waters of anticlassification when he asked petitioner's counsel whether "a benefit to descendants of slaves" would be "race-based."[25] Norris, the petitioner's counsel, agreed that this would be a forbidden "proxy for race"—an answer that at least appeared to give some satisfaction to some of the Justices (since the line of questions was not extended).[26] This "proxy for race" argument, however, can be extended across other elements of the college admissions process. Indeed, soon after the *SFFA* ruling, civil rights groups challenged Harvard's legacy preferences as, in effect, proxies for race among

20 Transcript of Oral Argument, *supra* note 2, at 7–10.

21 Statements about an agent's attitudes—such as belief, intent, and desire—are known as propositional attitude ascriptions; the object of the attitude is called the "extension" or the reference. Jeff Speaks, *Theories of Meaning*, STAN. ENCYC. PHIL. (Apr. 23, 2014), https://plato.stanford.edu/Archives/Win2017/entries/meaning/ [https://perma.cc/6J69-HMQB].

22 It should not go unremarked that extending the anticlassification rule of *Students for Fair Admissions, Inc. v. President & Fellows of Harvard College*, 143 S. Ct. 2141 (2023), in this fashion would have an asymmetrically large effect on racial minorities. Facial neutrality hence elicits substantive inequalities.

23 WILLARD VAN ORMAN QUINE, *Reference and Modality*, *in* FROM A LOGICAL POINT OF VIEW 139, 144 (2d ed. 1961).

24 Bostock v. Clayton County, 140 S. Ct. 1731, 1744 (2020) (emphasis added).

25 Transcript of Oral Argument, *supra* note 2, at 15.

26 *Id.* at 16. Justice Gorsuch responded with a curt "[o]kay." *Id.*

applicants.[27] But in other contexts, the Court has adopted inconsistent views on "proxies for" suspect classifications. On the one hand, when the state uses pregnancy in ways that predictably and durably disadvantage women, long-standing and recently confirmed precedent holds that this is not "discrimination based upon gender as such."[28] On the other hand, in the First Amendment free speech context, it is a matter of blackletter doctrine that a law cannot on its face regulate speech on the basis of its content (an anticlassification rule), *and separately* cannot be "justified" by "reference to the content of the regulated speech."[29] The latter bolt-on phrase is generally glossed to sweep in proxies for content discrimination.[30] Or take rules that distinguish a city's residents against everyone else, whether they live in or out of state. It is well established that such a city-resident/non-city-resident rule is facial discrimination "against *out-of-state* residents" in violation of the Privileges and Immunities Clause of Article IV.[31] City residence is here taken as a proxy for state residence. There does not appear, in short, to be a clear and consistent answer to whether a state actor can substitute an impermissible classification with a parallel, but verbally distinct, synonym without the resulting enactment still being ranked as suspect. The Court instead reaches different answers in different domains without noticing or accounting for its inconsistency.

These complexities—how "classification" relates to the cognitive function of an impermissible category and when a "proxy" counts as a prohibited ranking—illustrate a broader range of difficulties lurking below the smooth surface of anticlassification logic. They hint at how anticlassification hinges on technical questions of linguistic theory resistant to the dissolving impress of common sense. My primary aim here is to clarify and analyze these complexities, and thus to offer a more secure, albeit narrower, foundation for the doctrine.

Part I offers a close reading of the Court's recent decisions in order to show that anticlassification is a coherent, distinct, and increasingly common doctrinal choice. Careful examination of precedent demonstrates that the Court draws on a common reservoir of normative justifications for anticlassification rules in different lines of cases.

---

27 Stephanie Saul, *Harvard's Admissions Is Challenged for Favoring Children of Alumni*, N.Y. TIMES (July 3, 2023), https://www.nytimes.com/2023/07/03/us/harvard-alumni-children-affirmative-action.html [https://perma.cc/XH4C-KPW8].

28 Geduldig v. Aiello, 417 U.S. 484, 496 n.20 (1974), *see also* Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228, 2245–46 (2022) (quoting *Geduldig*, 417 U.S. at 496 n.20).

29 McCullen v. Coakley, 573 U.S. 464, 477 (2014) (quoting Ward v. Rock Against Racism, 491 U.S. 781, 791 (1989)); *see also Ward*, 491 U.S. at 791 (quoting Clark v. Cmty. for Creative Non-Violence, 468 U.S. 288, 293 (1984)).

30 *See McCullen*, 573 U.S. at 477.

31 United Bldg. & Constr. Trades Council v. Mayor of Camden, 465 U.S. 208, 221 (1984) (emphasis added).

As a result, anticlassification rules implementing distinct constitutional provisions rest on convergent, overlapping normative foundations. They are, in short, a coherent *type* of legal rule.[32] Thus, they are distinct from the rule against constitutionally forbidden intentions, the but-for test for discriminatory intent, or the disparate impact rule.

Part II closely examines the first step of any anticlassificatory legal regime: the doctrine is predicated on the assumption there exist coherent *categories*. These are labeled "race," "interstate commerce," "content distinctions," etc. But where do they come from? The law itself does not explain how these categories are formulated, nor how their boundaries are chalked out.

Drawing on the philosophy of language and social science work on psychological categorization, this Part demonstrates that the categories of anticlassification jurisprudence have surprisingly varied, extralegal genealogies. The philosophy of language's distinction between "natural kinds" and "social kinds" helps organize this variety. It also works as a template for evaluating whether a category is likely to be judicially tractable. I further draw attention to lines of constitutional anticlassification doctrine where the Court lacks *any* social or natural kind to stabilize a category's bounds. This absence of a coherent category, however, leaves doctrine dangerously porous to the haphazard and unprincipled force of unspoken judicial priors.

Part III turns from how categories are built to how they are applied in practice—i.e., how do judges decide whether a legal text contains an instance of the impermissible category?[33] If a legal text refers to "descendants of slaves," for example, does it contain a racial class? Once again, I show that application of anticlassification rules is nothing like the mechanical process some Justices seem to think it is. Comparing categories to legal texts is both linguistically complex and normatively freighted. For one thing, legal texts often use proxies or euphemisms. How an anticlassification rule applies to these in practice turns upon how it is justified: disparate normative foundations conduce to divergent results.

To develop this last point, I set out technical distinctions drawn from the philosophy of language between semantic and communicative content, and between sense and referent. I then draw upon these distinctions to demonstrate how different justifications for anticlassification norms point toward different outcomes in practice. As in Part II, my aim here is not to identify a "right" or a "wrong" way to

32 *See infra* note 39 and accompanying text (defining "types").

33 This Article is focused on the two threshold steps of anticlassification. As I discuss in subsection I.A.3, anticlassification anticipates the application of strict scrutiny. I do not consider here the dynamics of that more broadly used doctrinal device.

apply anticlassification norms. My aim instead is to unpack the complex relationship between the normative foundations of that doctrine and outcomes on the ground. Without a clear understanding of these analytic ligaments, judicial inquiry into whether a particular text triggers anticlassification concerns is likely to be unstructured, even arbitrary.

In a brief Conclusion, I reflect on the implications of these so-far obscured normative and analytic complexities. Rather than calling for wholesale abandonment of anticlassification doctrine, I suggest that judges and scholars should ascertain whether a specific anticlassification rule rests on an empirically tractable and normatively plausible footing. While such reflection should lead to a retreat from the more aggressive ambitions of anticlassificatory jurisprudence, it can also shore up other use-cases of this increasingly commonplace doctrinal instrument.

I. THE ANTICLASSIFICATORY CONSTITUTION

Anticlassification is an instance of a more general class of constitutional decision rule. This is a "procedure . . . for determining whether to adjudge the operative proposition [of the Constitution] satisfied" under specific conditions.[34] Decision rules are necessary as a practical matter because courts rarely apply the Constitution's operative provisions directly.[35] Mediating doctrinal structures are instead pervasive in the jurisprudence.

To situate anticlassification in the larger tool kit of constitutional doctrine, this Part opens with a definition. I aim to distinguish the phenomenon from other species of constitutional decision rules. Next, this Part identifies and maps leading examples from equal protection, free speech, and dormant Commerce Clause doctrine. I also include here an interesting variant on these anticlassification rules that emerged recently in free exercise jurisprudence. This warrants a mention not only because it illuminates how anticlassification logic might be modulated, but also because it will serve in Parts II and III as a useful foil for evaluating some of the Court's normative justifications for anticlassification. This Part concludes by adumbrating briefly the main

34 Mitchell N. Berman, *Constitutional Decision Rules*, 90 VA. L. REV. 1, 10 (2004) (emphasis omitted); *see also* Henry P. Monaghan, *The Supreme Court, 1974 Term — Foreword: Constitutional Common Law*, 89 HARV. L. REV. 1, 2–3 (1975) (noting that "a surprising amount of what passes as authoritative constitutional 'interpretation' is best understood as something of a quite different order—a substructure of substantive, procedural, and remedial rules drawing their inspiration and authority from, but not required by, various constitutional provisions").

35 Berman, *supra* note 34, at 106.

normative justifications for anticlassification rules adduced in the jurisprudence.

A. Definition of Anticlassification Rules

The anticlassification rule is a doctrinal instrument applied in three distinct steps. These are (1) identifying the category, (2) determining whether an instance of the category is found in the text of a challenged legal text, and (3) determining whether the resulting classification survives strict scrutiny. Each step can be illustrated using Chief Justice Roberts's 2023 opinion in *SFFA*.[36] Judges trace the same series of steps, however, in other doctrinal contexts. At least in this respect, the logic of *SFFA* is exemplary, not idiosyncratic.

1. Identifying the Constitutionally Salient Category

First, the Court identifies a *category* of constitutional interest that comprises a number of *exemplars or instances* to be sought in legal texts. There is a "difference between talking about things and talking about categories."[37] A category is "a number of objects that are considered equivalent."[38] (In the philosophy literature, a related distinction is drawn using the terms "type" and "token." "Types are generally said to be abstract and unique; tokens are concrete particulars . . . ."[39]) I use the terms "category" and "type" interchangeably here. Likewise, I will alternate between "exemplar," "token," and "instance." I use the term "classification," in contrast, to refer to the larger analytic structure of judicial analysis.

Category creation is a pervasive feature of animal cognition.[40] Psychologists describe categories as "ecologically equivalent objects that have imperfect perceptual similarity."[41] Categories generally "map the

---

36 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141 (2023).

37 Jill C. Anderson, *Misreading Like a Lawyer: Cognitive Bias in Statutory Interpretation*, 127 HARV. L. REV. 1521, 1534 (2014).

38 Eleanor Rosch, *Principles of Categorization*, *in* COGNITION AND CATEGORIZATION 27, 30 (Eleanor Rosch & Barbara B. Lloyd eds., 1978).

39 Linda Wetzel, *Types and Tokens*, STAN. ENCYC. PHIL. (Apr. 28, 2006), https://plato.stanford.edu/entries/types-tokens/#Occ [https://perma.cc/WA9P-KTE9]. Consider the sentence, "Rose is a rose is a rose is a rose." It has three types and ten tokens. *Id.*; *see also* W.V. QUINE, QUIDDITIES: AN INTERMITTENTLY PHILOSOPHICAL DICTIONARY 216–17 (1987).

40 *See* Smith et al., *supra* note 16, at 1.

41 *Id.* There is some confusion in the social psychology literature over the terms "categories, schemas, and similar concepts," with "the terms used in often quite different ways by different scholars." Ronald Chen & Jon Hanson, *Categorically Biased: The Influence of*

perceived world structure as closely as possible" as a result of evolutionary fitness pressures.[42] In a pervasive just-so story of primate development, for example, a hominid cognitively equipped with a category of "tigers" is likely, all else being equal, to survive longer than one lacking that category for organizing the "perceived world."[43] To serve this function, category boundaries need not be precise. They are often "fuzzy," so it is "not always clear which instances belong in the category."[44] Cognitive use of categories also spills over into domains where differences are not readily observable.[45] A reason for this is that the mental act of categorization can reduce the cognitive cost of navigating the social as much as the presocial world.[46]

This two-level structure of categories and instances is illustrated by *SFFA*, where the Court drew a distinction between the overarching category itself and its instances. In *SFFA*, the category was "race."[47] The instances of this category, noted Chief Justice Roberts, included "Asian," "White," "African-American," "Hispanic," and "Native American."[48] "Race," then, is not itself an instance: it is the superordinate category.

2. Locating an Exemplar

Second, the Court must ask whether the verbal formulation of a government action contains an instance of the category. At a very abstract level, this is like examining two lists in order to figure out whether any of the items on the first list also appear on the second. At least in theory, there is no reason a judge needs to look beyond the text of the challenged legal entitlement in order to determine whether an exemplar of the relevant category is present in a law's words.

Knowledge Structures on Law and Legal Theory, 77 S. CAL. L. REV. 1103, 1131 (2004). I bracket that terminological debate.

42 Rosch, *supra* note 38, at 28, 28–29; *see also* Chen & Hanson, *supra* note 41, at 1132 ("[T]he absence of clear concepts and categories increases the cognitive energies required to process information and thus deters individuals from learning new ideas or processing new information.").

43 *Cf.* Rosch, *supra* note 38, at 29.

44 SUSAN T. FISKE & SHELLEY E. TAYLOR, SOCIAL COGNITION 106 (2d ed. 1991).

45 Social psychologists such as Henri Tajfel and A.L. Wilkes demonstrated in the 1960s that humans have a cognitive tendency to categorize even in the absence of actually observable differences between objects. Henri Tajfel & A.L. Wilkes, *Classification and Quantitative Judgement*, 54 BRIT. J. PSYCH. 101, 106–13 (1963).

46 Eleanor Rosch, *Human Categorization*, *in* 1 STUDIES IN CROSS-CULTURAL PSYCHOLOGY 1, 1–2 (Neil Warren ed., 1977).

47 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2162 (2023).

48 *Id.* at 2167 (citing Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 397 F. Supp. 3d 126, 137, 178 (D. Mass. 2019)).

Indeed, an anticlassification rule is distinctive in excluding such nontextual evidence.[49] In this respect, it tracks the "plain meaning" approach to statutory interpretation.[50]

Once again, *SFFA* provides an illustration. The Court in that case looked to the "only four pieces of information" considered in Harvard's "lop list" used for a second cut of candidates.[51] Because one of those items was explicitly denominated "race," the second step of the anticlassification rule had been made out.[52] It seems likely that the perceived mechanical quality of this step contributes substantially to the allure of anticlassification rules as "commonsense."[53]

Two points of tension, however, hint at the complexity of anticlassification rules. First, Chief Justice Roberts identified the presence of exemplars of the category "race" as constitutionally problematic; he did not suggest that the presence of the category "race" as such would be itself a trigger for constitutional scrutiny.[54] Antidiscrimination statutes such as Title VI of the Civil Rights Act, on the one hand, are not constitutionally problematic because they pick out the category of race, but not its types.[55] As a result, the Court could hold that Harvard violated Title VI, and not that Title VI (which explicitly mentions "race") itself violated the Equal Protection Clause. On the other hand, a measure that directs admissions offers to "account for race" (again, mentioning only the type, not its tokens), seems to elicit race-based thinking. This leaves then a question: why are a category's instances problematic, but not the category itself?

Second, it is not clear that all of the terms mentioned in the Harvard admissions rubric are in fact instances of the category "race." There is, for example, continuing judicial and academic controversy over whether "Native American" is a token of the type "race."[56] Chief

---

49 As we will see, the firmness of this exclusionary rule can expand or contract over time.

50 William Baude & Ryan D. Doerfler, *The (Not So) Plain Meaning Rule*, 84 U. CHI. L. REV. 539, 542 (2017) ("[T]he plain meaning rule['s] . . . key feature is to deny the relevance of other interpretive data if the text's meaning is 'plain' or 'clear.'"). For a defense of such evidence-excluding rules for judges as well as juries, see Frederick Schauer, *On the Supposed Jury-Dependence of Evidence Law*, 155 U. PA. L. REV. 165, 196–97 (2006).

51 *Students for Fair Admissions*, 143 S. Ct. at 2155.

52 *Id.*

53 Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015).

54 *Students for Fair Admissions*, 143 S. Ct. at 2167.

55 42 U.S.C. § 2000d (2018) ("No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.").

56 *See* Haaland v. Brackeen, 143 S. Ct. 1609, 1638 (2023) (describing a challenge to the Indian Child Welfare Act under the Equal Protection Clause for its "racial discrimination," but declining to address the challenge on Article III standing grounds); *see also* Addie

Justice Roberts also criticized the term "Hispanic" as "arbitrary or undefined."[57] If so, this raises a further question as to whether it too is properly an example of "race."[58] Importantly, the existence of uncertainty or active controversy as to whether a specific item is within or outside a category, however, does not undermine the validity of that category[59]: it is possible to talk of "race," that is, even if we disagree as to the inclusion of "Native American" or "Hispanic" as an instance thereof.

3. Applying Strict Scrutiny

Third, it is not the case that judicial identification of a constitutionally suspect classification leads directly to a law's invalidity. Once a forbidden instance is found in a law, the Court asks whether it "is used to further compelling governmental interests" in a "narrowly tailored" way.[60] In *SFFA*, the crux of the Court's decision was its narrowing of the class of permissible "compelling state interest[s]"[61] that could justify employment of a racial classification. Two such interests were endorsed: "remediating specific, identified instances of past discrimination that violated the Constitution or a statute" and "avoiding imminent and serious risks to human safety in prisons."[62] Both the educational interests tendered by the two universities were rejected as "not sufficiently coherent for purposes of strict scrutiny."[63] This was enough, as a doctrinal matter, to resolve the case.

My focus in this Article is not on this third step. But it is worth noting here briefly that it is hardly self-evident how strict scrutiny more

---

C. Rolnick, *Indigenous Subjects*, 131 YALE L.J. 2652, 2664–65 (2022) (discussing and critiquing, per *Brakeen*, the extension of race as a "biological" category to include Native Americans, *id.* at 2665); *infra* text accompanying notes 252–61.

57 *Students For Fair Admissions*, 143 S. Ct. at 2167 (citing Mark Hugo Lopez, Jens Manuel Krogstad & Jeffrey S. Passel, *Who Is Hispanic?*, PEW RSCH. CTR. (Sept. 5, 2023), https://www.pewresearch.org/short-reads/2023/09/05/who-is-hispanic/ [https://perma.cc/AUB5-CD66]).

58 Despite the Chief Justice's criticism, the Census Bureau has asked individuals to self-identify as "Hispanic" or not since 1980 without much confusion or controversy. *See* Harvey M. Choldin, *Statistics and Politics: The "Hispanic Issue" in the 1980 Census*, 23 DEMOGRAPHY 403, 413–14 (1986) (discussing the process by which the question was included in the census).

59 See FISKE & TAYLOR, *supra* note 44, at 106.

60 *Students for Fair Admissions*, 143 S. Ct. at 2150 (first quoting Grutter v. Bollinger, 539 U.S. 306, 326 (2003), *abrogated by Students for Fair Admissions*, 143 S. Ct. 2141; and then quoting Fisher v. Univ. of Tex. at Austin, 570 U.S. 297, 311 (2013)).

61 *Id.* at 2164 (quoting *Grutter*, 539 U.S. at 325).

62 *Id.* at 2162 (first citing Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 720 (2007); then citing Shaw v. Hunt, 517 U.S. 899, 909–10 (1996); and then citing Johnson v. California, 543 U.S. 499, 512–13 (2005)).

63 *Id.* at 2166.

generally is supposed to work. A pivotal variable is the range of permissible "compelling state interests." In equal protection jurisprudence after *SFFA*, there seem to be only two. In other areas, however, the Court has not identified a parallel closed-off class of cognizable state interests. Instead, it has examined seriatim interests invoked by the state, considering each on the merits. For example, in a 2011 free speech challenge to Vermont's Prescription Confidentiality Law, the Court identified an impermissible facial classification.[64] It then recognized two cognizable state interests of protecting "medical privacy" and improving public health at lower costs.[65] Rather than rejecting either state interest as a categorical matter, the majority examined each in turn.[66] By necessary implication, neither of the two interests Vermont articulated was prima facie invalid. Neither, that is, were categorically flawed state interests of the sort proffered by the universities in *SFFA*.

All this is to say that it is not always obvious how the Court determines what counts as a compelling state interest.[67] Nor is it clear why the permissible range of such interests should expand or contract, accordion-like, as judges pivot from one provision of the Constitution to another. But it is plain that the range of compelling state interests varies dramatically from one doctrinal context to another. So while this Article focuses on the threshold structure of anticlassification, there are other locations where judicial discretion can be exercised.

* * *

To summarize, an anticlassification rule entails (1) identification of a category; (2) a search for instances of the category in the law's text; and (3) ascertaining whether a properly tailored, compelling state interest warrants deployment of such a classification. This Article focuses on steps (1) and (2).

B. The Paths of Anticlassification Rules in Constitutional Law

Anticlassification rules have emerged as the preferred doctrinal instrument for realizing the Constitution's many evenhanded-treatment commands. While most commonly associated with the equality context, judicial concern with improper distinctions can be discerned as early as the 1860s in cases dealing with federalism's horizontal component. Its crystallization in equal protection and free speech contexts

64 Sorrell v. IMS Health Inc., 564 U.S. 552, 567 (2011).

65 *Id.* at 572.

66 *Id.* at 572–79.

67 *Cf.* Richard H. Fallon, Jr., *Strict Judicial Scrutiny*, 54 UCLA L. REV. 1267, 1322 (2007) (noting that the Court often "labels interests as compelling on the basis of little or no textual inquiry").

occurred a century later. In the last five years, free exercise jurisprudence has also witnessed the emergence of a variant on anticlassification called the "most favored nation" rule.[68] The latter is akin to an "anti-anticlassification" rule.

1. Horizontal Federalism

A brief historical survey of anticlassification rules usefully begins with horizontal federalism, or the relations between a state and its citizens on the one hand, and other states and their citizens on the other. For anticlassification rules were used here before anywhere else.

In "negative" or "dormant" Commerce Clause jurisprudence, an anticlassification rule emerged as early as 1870.[69] The Court invalidated a Missouri statute taxing peddlers only if they sold out-of-state goods.[70] By imposing "burdens . . . by reason of [goods'] foreign origin," the Missouri law trenched on national commerce.[71] Reconstructing the doctrine after the jurisprudential convulsions of the New Deal,[72] the Court reverted to a "virtually *per se* rule of invalidity" when a state statute "on its face and in its plain effect . . . violates this principle of nondiscrimination."[73] The Court recently characterized the rule as triggered whenever a state law "*mandate[s]* 'differential treatment'" of "economic interests" inside and beyond the state.[74] That is, when the law "[b]y its plain terms" entails differential treatment of market

68 Nelson Tebbe, *The Principle and Politics of Equal Value*, 121 COLUM. L. REV. 2397, 2399 (2021).

69 *See* Brannon P. Denning, *Reconstructing the Dormant Commerce Clause Doctrine*, 50 WM. & MARY L. REV. 417, 441 (2008) (suggesting that the rule is implicit in caselaw from the 1860s). For an argument that the antidiscrimination rule is "perfectly consistent with framing-era concerns about economic balkanization, and no one at the Convention approved of these laws," see Barry Friedman and Daniel T. Deacon, *A Course Unbroken: The Constitutional Legitimacy of the Dormant Commerce Clause*, 97 VA. L. REV. 1877, 1927 (2011).

70 Welton v. Missouri, 91 U.S. 275, 278, 282 (1876).

71 *Id.* at 282. To be sure, *Welton*'s use of the phrase "by reason of" is ambiguous: it could mean "motivated by" or "by a rule that facially distinguishes using."

72 One study takes the doctrine's present form to "the late 1970s." *See* Friedman & Deacon, *supra* note 69, at 1926.

73 City of Philadelphia v. New Jersey, 437 U.S. 617, 624, 627 (1978); *see also* Or. Waste Sys., Inc. v. Dep't of Env't Quality of Or., 511 U.S. 93, 99 (1994) ("'[D]iscrimination' simply means differential treatment of in-state and out-of-state economic interests that benefits the former and burdens the latter. If a restriction on commerce is discriminatory, it is virtually *per se* invalid." (citing Chem. Waste Mgmt., Inc. v. Hunt, 504 U.S. 334, 344 & n.6 (1992))); Granholm v. Heald, 544 U.S. 460, 472 (2005) (same). The per se rule is subject to scope limits. It does not apply to states acting as market participants or performing a "government function." Dep't of Revenue of Ky. v. Davis, 553 U.S. 328, 341 (2008) (citing United Haulers Ass'n v. Oneida-Herkimer Solid Waste Mgmt. Auth., 550 U.S. 330, 343–44 (2007)).

74 *Granholm*, 544 U.S. at 472 (emphasis added) (quoting *Or. Waste. Sys.*, 511 U.S. at 99).

participants or goods based on their origin, the Constitution's commitment to a national market, embodied in the Commerce Clause, is violated.[75]

There is a second rule of horizontal, interstate equality issuing from the Privileges and Immunities Clause of Article IV. This entitles "Citizens of each State . . . to all Privileges and Immunities of Citizens in the several States."[76] Applying this command, the Court has leaned again on anticlassification logic. For example, the Court struck down in 1988 a New York tax rule that permitted only residents to deduct alimony payments as "facially inequitable."[77] The Privileges and Immunities Clause's rule against facial discrimination extends to measures that distinguish a specific city's residents, as well as laws that carve between in-state and out-of-state residents.[78]

The anticlassification "core"[79] of horizontal federalism has been stable now for a half century. At the same time, litigants have repeatedly pressed the Court to supplement it with an effects test. Thus, the dormant Commerce Clause anticlassification rule has been supplemented with a relatively permissive effects-focused test that weighs the costs and benefits of state measures with discriminatory repercussions.[80] Where, for example, a state regulatory scheme imposes parallel requirements of physical presence on both in-state and out-of-state market participants, but then imposes a "gauntlet" of steps only on the latter, the Court has had "no difficulty" finding discrimination.[81]

In 2023 a fragmented Court in *National Pork Producers Council v. Ross* declined to establish a more stringent "'almost *per se*' rule against

75 Healy v. Beer Inst., 491 U.S. 324, 341, 340–41 (1989). In a later case, the Court observed that "Tennessee's 2-year durational-residency requirement plainly favors Tennesseans over nonresidents," and so triggered strict scrutiny. Tenn. Wine & Spirits Retailers Ass'n v. Thomas, 139 S. Ct. 2449, 2462 (2019).

76 U.S. CONST. art. IV, § 2, cl. 1.

77 Lunding v. N.Y. Tax Appeals Tribunal, 522 U.S. 287, 310 (1998); *see also* Hicklin v. Orbeck, 437 U.S. 518, 520 (1978) (invalidating an "Alaska Hire" measure "'requiring the employment of qualified Alaska residents' in preference to nonresidents" (quoting ALASKA STAT. § 38.40.030(a) (1977))).

78 United Bldg. & Constr. Trades Council v. Mayor of Camden, 465 U.S. 208, 217–18 (1984) ("We conclude that Camden's ordinance is not immune from constitutional review at the behest of out-of-state residents merely because some in-state residents are similarly disadvantaged." (citing Zobel v. Williams, 457 U.S. 55, 75 (1982) (O'Connor, J., concurring in judgment))).

79 Nat'l Pork Producers Council v. Ross, 143 S. Ct. 1142, 1153 (2023) ("[T]he antidiscrimination principle [lies] at the core of this Court's dormant Commerce Clause cases . . . .").

80 Pike v. Bruce Church, Inc., 397 U.S. 137, 142 (1970).

81 Granholm v. Heald, 544 U.S. 460, 474, 476, 474–76 (2005).

state laws with 'extraterritorial effects.'"[82] It deflected the National Pork Producers Council's challenge to a California animal welfare measure on the ground that the state law imposed high compliance costs on both in-state and out-of-state farmers.[83] The Court thus declined to supplement the facial-classification rule with a strong effects-based test. *Ross*, nevertheless, left open questions as to how the Court will treat states' efforts to directly regulate their own residents' out-of-state activity, or shape private conduct unfolding beyond state boundaries due to moral objections. Derogation of the constitutional right of reproductive choices means these questions are likely to arise sooner rather than later.[84]

2. Equal Protection

Anticlassification rules have trodden an uneven path as instruments of equal protection. Early glosses on the Clause can, with some work, be read as turning on the presence of facial classifications.[85] Judicial embrace of Jim Crow, of course, sapped that gloss of plausibility. Facial classifications aimed at the subordination of Blacks and other racial minorities proliferated.[86] But after the federal courts slowly turned against Jim Crow in the 1950s, the dominant rule in equal protection was again unsettled. On the one hand, the post–World War II Court issued opinions that rejected motivation or intent as a

82 *Nat'l Pork Producers*, 143 S. Ct. at 1155 (quoting Brief for Petitioners at 19, 23, *Nat'l Pork Producers*, 143 S. Ct. 1142 (No. 21-468)).

83 *Id.* at 1155–56 (citing manageability costs and unintended consequences as a reason not to accept an extraterritoriality effect rule).

84 Josh Gerstein, *In New Supreme Court Decision, Abortion Lurks Just Below the Surface*, POLITICO (May 12, 2023, 6:46 PM EDT), https://www.politico.com/news/2023/05/12/supreme-court-california-pork-ruling-abortion-00096797 [https://perma.cc/6N7J-GSJF]. For a treatment of "different routes" by which a statute could extend civil regulation and digital surveillance of pregnant persons respectively, see Katherine Florey, Dobbs *and the Civil Dimension of Extraterritorial Abortion Regulation*, 98 N.Y.U. L. REV. 485, 497, 497–500 (2023).

85 *See, e.g.*, Strauder v. West Virginia, 100 U.S. 303, 308–10 (1880) (invalidating a state statute that limited jury service to "white male persons . . . twenty-one years of age," *id.* at 305), *abrogated by* Taylor v. Louisiana, 419 U.S. 522 (1975); Neal v. Delaware, 103 U.S. 370, 397 (1881).

86 *See* Plessy v. Ferguson, 163 U.S. 537, 551 (1896); *see also* Michael J. Klarman, *The* Plessy *Era*, 1998 SUP. CT. REV. 303, 333–39 (documenting the public and judicial aftermath of *Plessy*).

touchstone.[87] On the other hand, it handed down decisions difficult to explain without reference to intent.[88]

Only in 1964 did the Court point to "[c]lassification[s]" as a distinct problem for constitutional equality purposes.[89] Between the 1980s and the early 2000s, the doctrinal focus on classification as such gradually sharpened. The Court struggled in that period with the question whether "race-conscious measures" used to corrode historical patterns of marginalization received strict scrutiny, or something less, and ultimately whether any compelling state interest could justify them.[90] This line of cases converged on the proposition that any occasion on which "the government distributes burdens or benefits on the basis of individual racial classifications" leads to "strict scrutiny."[91] In this model of equal protection, "the level of scrutiny applied to alleged claims of discrimination turns on the identity category at issue."[92]

At the same time, the equal protection rule against racial classifications was not entirely static. It morphed from a porous and flexible standard to a rigid and impermeable rule. In 1995, Justice O'Connor could write a majority opinion rejecting "the notion that strict scrutiny is 'strict in theory, but fatal in fact.'"[93] By 2023, Chief Justice Roberts

---

87 *See, e.g.*, Palmer v. Thompson, 403 U.S. 217, 224 (1971) (finding "no case in this Court [holding] that a legislative act may violate equal protection solely because of the motivations of the men who voted for it").

88 *See, e.g.*, Griffin v. Cnty. Sch. Bd., 377 U.S. 218, 232 (1964) (invalidating the closing of a school board to avoid desegregation).

89 McLaughlin v. Florida, 379 U.S. 184, 190 (1964) ("Classification 'must always rest upon some difference which bears a reasonable and just relation to the act in respect to which the classification is proposed . . . .'" (quoting Gulf, Colo. & Santa Fé Ry. Co. v. Ellis, 165 U.S. 150, 155 (1897))); *see also* Michael Klarman, *An Interpretive History of Modern Equal Protection*, 90 MICH. L. REV. 213, 255 (1991) ("For the first time the Court in *McLaughlin* both articulated *and* applied a more rigorous review standard to racial classifications . . . .").

90 Metro Broad., Inc. v. FCC, 497 U.S. 547, 564 (1990), *overruled by* Adarand Constructors, Inc. v. Pena, 515 U.S. 200 (1995). Initially, the Court used a weaker form of scrutiny for "benign" classifications. *See id.* It retreated from this position in *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (plurality opinion). *See also Adarand Constructors*, 515 U.S. at 235 ("Federal racial classifications, like those of a State, must serve a compelling governmental interest, and must be narrowly tailored to further that interest.").

91 Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 720 (2007); *see also* Gratz v. Bollinger, 539 U.S. 244, 270 (2003) (describing the use of such classifications as "pernicious" (quoting Fullilove v. Klutznick, 448 U.S. 448, 537 (1980) (Stevens, J., dissenting))); Hunt v. Cromartie, 526 U.S. 541, 546 (1999) (holding that "all laws that classify citizens on the basis of race . . . are constitutionally suspect and must be strictly scrutinized"); Shaw v. Reno, 509 U.S. 630, 657 (1993) (same).

92 Lauren Sudeall Lucas, *Identity as Proxy*, 115 COLUM. L. REV. 1605, 1613 (2015).

93 *Adarand*, 515 U.S. at 237 (quoting *Fullilove*, 448 U.S. at 519 (Marshall, J., concurring in judgment)). For examples of a more flexible form of classification review, see *Fisher v. University of Texas at Austin*, 579 U.S. 365, 388 (2016), and *Grutter v. Bollinger*, 539 U.S. 306,

would describe the Court's "acceptance of race-based state action [as] rare."[94] After *SFFA*, of course, it will be yet rarer still.

So understood, an anticlassification regime now covers much of the waterfront of equal protection litigation.[95] Notably, the Justices have refrained from extending it into the domains of family law and criminal law.[96] By the time *SFFA* was decided, nevertheless, the basic doctrinal entrenching of skepticism of any and all racial classifications had taken deep roots.

3. Free Speech

Like equal protection jurisprudence, free speech law inducted anticlassification into its canon during the Burger Court. The main anticlassification regime under the Speech Clause of the First Amendment concerns content discrimination. In 1972, the Court struck down a Chicago ordinance that prohibited schools from being picketed at certain times, but that made an exception for labor-related protests.[97] This was, the Court said, a distinction based on speech's "message, its ideas, its subject matter, or its content,"[98] and as such subject to strict scrutiny. The majority opinion highlighted an unjustified distinction between labor and nonlabor disputes.[99] A decade later, it was blackletter law that all "[r]egulations which permit the Government to

---

343 (2003), *abrogated by* Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141 (2023).

94 *Students for Fair Admissions*, 143 S. Ct. at 2162.

95 At the same time, the Court maintains a separate doctrinal structure for cases where "invidious discriminatory purpose was a motivating factor [which] demands a sensitive inquiry into such circumstantial and direct evidence of intent." Village of Arlington Heights v. Metro. Hous. Dev. Corp., 429 U.S. 252, 266 (1977).

96 The Court has hence turned aside numerous chances to consider whether race-specific suspect selection rules are valid. *See, e.g.*, Monroe v. City of Charlottesville, 579 F.3d 380 (4th Cir. 2009), *cert. denied*, 559 U.S. 992 (2010); Brown v. City of Oneonta, 221 F.3d 329 (2d Cir. 2000), *cert. denied*, 534 U.S. 816 (2001). In contrast, the Court periodically allows intent-based challenges in criminal cases. *See, e.g.*, Pena-Rodriguez v. Colorado, 137 S. Ct. 855, 869 (2017). The Court has also declined to rule on the role pervasively played by race in family law. Katie Eyer, *Constitutional Colorblindness and the Family*, 162 U. PA. L. REV. 537, 540–41 (2014) ("Constitutional challenges to these race-based actions have generally fared poorly, with courts typically (albeit not always) applying de minimis constitutional scrutiny." *Id.* at 541.); *see also* R. Richard Banks, *The Color of Desire: Fulfilling Adoptive Parents' Racial Preferences Through Discriminatory State Action*, 107 YALE L.J. 875, 904–08 (1998).

97 Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95 (1972).

98 *Id.*; *see also* Leslie Kendrick, *Content Discrimination Revisited*, 98 VA. L. REV. 231, 242 (2012) (citing this language and noting that it is "hardly self-defining"). *Mosley*, for example, slips from a concern with "the ordinance itself" to a motivation-focused concern about "censorship." *Mosley*, 408 U.S. at 99.

99 *Mosley*, 408 U.S. at 101–02.

discriminate on the basis of the content of the message cannot be tolerated under the First Amendment."[100]

Akin to its counterpart in the equal protection context, the content-discrimination rule has scope conditions. It applies mainly to speech in a public forum.[101] And in its first decades, the prohibition on content discrimination alternated between a "strict test," pursuant to which all "facial content distinctions [were] *necessarily* content-based," and a "purpose-oriented" norm that looked to a regulation's justifications as well as its verbal content.[102] In other words, the Court treated the exclusionary effect of the content-discrimination rule on nontextual evidence as either more or less absolute. Yet the Court consistently treated facial discrimination on the basis of content as a trigger for strict scrutiny.[103]

In the Roberts Court, the doctrine had sharpened into a rule tightly focused on the verbal content of a challenged measure. In a pair of cases, the Court first adopted a strict version of this approach before pivoting to a weaker, more standard-like doctrinal norm. In the 2015 case of *Reed v. Town of Gilbert*, the Court insisted that strict scrutiny applied whenever "a regulation of speech 'on its face' draws distinctions based on the message a speaker conveys."[104] The majority distinguished content-neutral rules justified by reference to speech,

100 Regan v. Time, Inc., 468 U.S. 641, 648–49 (1984) (first citing Carey v. Brown, 447 U.S. 455, 463 (1980); and then citing *Mosley*, 408 U.S. at 95–96). But for an example of content-based discrimination being applied to speech not uttered in a public forum, see *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 563–64 (2011) (describing a Vermont law regulating physician detailing as "content- and speaker-based restrictions on the sale, disclosure, and use of prescriber-identifying information").

101 *Carey*, 447 U.S. at 461–62; *see also* Burson v. Freeman, 504 U.S. 191, 196–97 (1992) (plurality opinion) (examining the nature of the forum before applying content-discrimination analysis). For a more precise statement of content discrimination's scope conditions, see Leslie Kendrick, *Nonsense on Sidewalks: Content Discrimination in* McCullen v. Coakley, 2014 SUP. CT. REV. 215, 222 (noting the exception for government employment contexts).

102 Genevieve Lakier, Reed v. Town of Gilbert, Arizona*, and the Rise of the Anticlassificatory First Amendment*, 2016 SUP. CT. REV. 233, 246, 246–47. An example of the second category is *City of Cincinnati v. Discovery Network, Inc.*, which focused on the "relationship" between an ordinance's distinctions and "the particular interests that the city has asserted." City of Cincinnati v. Discovery Network, Inc., 507 U.S. 410, 424 (1993).

103 Kendrick, *supra* note 98, at 254 ("With two exceptions, for upward of thirty years the Court has found every facial classification on the basis of subject matter or viewpoint to be content based."). The main exception to this trend is a line of cases concerning adult movie theaters, where the Court declined to find facial distinctions dispositive on the theory that legislators were targeting "secondary effects" such that the content divide could be "'*justified* without reference to the content of the regulated speech.'" City of Renton v. Playtime Theatres, Inc., 475 U.S. 41, 47–48 (1986) (first emphasis omitted) (quoting Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc., 425 U.S. 748, 771 (1976)).

104 Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015) (quoting *Sorrell*, 564 U.S. at 564).

placing them into a separate doctrinal category.[105] The municipal signage ordinance challenged in *Reed*, the Court said, discriminated "[o]n its face," and so there was "no need to consider" evidence of the government's justifications or motives.[106] *Reed*, like *SFFA*, represents an apogee of formalist anticlassification—one where examination of the ordinance's text led directly to the exclusion of any other evidence.[107] Seven years later, however, a different majority of the Court in *City of Austin v. Reagan National Advertising of Austin, LLC* seemed to draw back from this strict formalism.[108] "[A]bsent a content-based purpose or justification," the Court ruled, a statutory distinction is content neutral.[109] *City of Austin*, unlike *Reed*, does not merely allow, but insists upon consideration of evidence beyond a statute's text. As such, it is a step back from the quintessence of anticlassification.

4. Anti-Anticlassification: Free Exercise

In 2020, the Supreme Court changed the decision rule applied to free exercise challenges lodged against enactments that did not mention religion on their face or that arrived trailing clouds of animus.[110] In a series of cases concerning COVID-19 restrictions, the Court ruled that whenever a general law creates an exception for a "secular" entity, activity, or motivation, it unconstitutionally discriminates against religion if it does not also offer an exemption to all "comparable" religious entities, activities, and motivations.[111] Analogizing to international trade law, scholars have labeled this a "most favored nation" (MFN) regime.[112]

In a leading scholarly treatment of this new free exercise jurisprudence, Nelson Tebbe has observed that it "can be violated even in the

105 *Id.* at 164 (holding that such regulations are also, separately, subject to strict scrutiny).

106 *Id.* at 164–65.

107 *See supra* text accompanying note 13 (discussing evidentiary exclusionary rules).

108 City of Austin v. Reagan Nat'l Advert. of Austin, LLC, 142 S. Ct. 1464 (2022).

109 *Id.* at 1471.

110 The leading precedent until then imposed a rule of facial neutrality, *see* Emp. Div. v. Smith, 494 U.S. 872, 878–79 (1990), and prohibited discriminatory gerrymandering of regulatory schemes, *see* Church of the Lukumi Babalu Aye, Inc. v. City of Hialeah, 508 U.S. 520, 545 (1993).

111 Tandon v. Newsom, 141 S. Ct. 1294, 1296 (2021) (per curiam). For other examples, see *Roman Cath. Diocese of Brooklyn v. Cuomo*, 141 S. Ct. 63, 66–67 (2020) (per curiam), *S. Bay United Pentecostal Church v. Newsom*, 140 S. Ct. 1613, 1613 (2020) (Roberts, C.J., concurring in denial of application for injunctive relief), *Gateway City Church v. Newsom*, 141 S. Ct. 1460 (2021) (mem.), and *S. Bay United Pentecostal Church v. Newsom*, 141 S. Ct. 716, 717 (2021) (Barrett, J., concurring in partial grant of application for injunctive relief).

112 Tebbe, *supra* note 68, at 2399. The idea of a most favored nation is commonly attributed to Douglas Laycock, *The Remnants of Free Exercise*, 1990 SUP. CT. REV. 1, 49.

absence of facial differentiation."[113] He distinguishes it from both the equal protection and the free exercise rules discussed above.[114] This is true as far as it goes—but it misses something too. For canonical anticlassification and MFN regimes share several important traits.

First, both turn on the verbal content of an enactment. In both regimes, judicial examination of verbal content is outcome determinative. MFN tests, that is, can have the same effect of excluding nontextual evidence (e.g., of intent or justification) as anticlassification rules.[115]

Second, the MFN rule treated the absence of a religious exemption in COVID-19 regulations as a failure of facial neutrality "because [those rules] single out houses of worship for especially harsh treatment."[116] As in anticlassification jurisprudence, the concern expressed here is that the state is not treating like things alike.[117] There is a clear parallel in the dormant Commerce Clause context, where the failure to grant an exception from a general rule to out-of-state actors who are differently situated from in-state actors counts as facial discrimination.[118]

Finally, the last step in each of these two regimes are mirror images of each other. Constitutionality hinges on the absence of an exemplar of a forbidden category in anticlassification rules. In contrast,

---

113 Tebbe, *supra* note 68, at 2424.

114 *Id.* at 2425.

115 *See Roman Cath. Diocese*, 141 S. Ct. at 66. But note that in determining what activities are "comparable," a court is likely to examine the justifications for statutory exceptions. *Tandon*, 141 S. Ct. at 1296.

116 *Roman Cath. Diocese*, 141 S. Ct. at 66. Although the Court's reasoning is not fulsome, it seems to turn on the idea that the absence of a religious exception is evidence of disparaging disregard for religion. *See* David Simson, *Most Favored Racial Hierarchy: The Ever-Evolving Ways of the Supreme Court's Superordination of Whiteness*, 120 MICH. L. REV. 1629, 1637 (2022) (explaining that MFN analysis aims to seek out instances in which "the government has implicitly made an impermissible value judgment that religious reasons for engaging in regulated activity are less valuable and thus less entitled to solicitude"). On this view, MFN analysis is congruent with the general rule that "denying a generally available benefit solely on account of religious identity imposes a penalty on the free exercise of religion that can be justified only by a state interest 'of the highest order.'" Trinity Lutheran Church of Columbia, Inc. v. Comer, 137 S. Ct. 2012, 2019 (2017) (quoting McDaniel v. Paty, 435 U.S. 618, 628 (1978) (plurality opinion)).

117 *See, e.g.*, *Tandon*, 141 S. Ct. at 1296; Gen. Motors Corp. v. Tracy, 519 U.S. 278, 299 (1997) (noting the "threshold question whether the companies are indeed similarly situated for constitutional purposes" in the context of the dormant Commerce Clause).

118 In *Granholm v. Heald*, the Court ranked a New York regulation of alcohol that required out-of-state actors to "establish a distribution operation in New York" as one that "grants in-state wineries access to the State's consumers on preferential terms." Granholm v. Heald, 544 U.S. 460, 474 (2005) (citing N.Y. ALCO. BEV. CONT. LAW §§ 3(37), 96 (McKinney 2005)). The *Granholm* Court's analysis here is not completely clear, but it seems to rank the New York law as a facially discriminatory one. *Id.* at 476.

the failure to mention religion indicates a statute's unconstitutionality under an MFN regime. Whereas the *presence* of the token is fatal in one case, its *absence* is fatal in another. For this reason, it is plausible to call MFN an anti-anticlassification rule.

This doctrinal parallel also highlights an easily missed puzzle: "Race" and "religion" are both constitutionally salient classifications. But one is linked to an anticlassification rule, and the other is linked to an anti-anticlassification rule. Why? In what respect are the categories "race" and "religion" diametrically opposed such that they would elicit mirror-image doctrinal protections? The puzzle is acute because on one reading of the Constitution, race and religion should be treated in identical ways. The federal government is not directly covered by the equal protection component of the Fourteenth Amendment. Instead, it is constrained by the equal protection component of the Due Process Clause of the Fifth Amendment.[119] This has been read to apply whenever the government uses "an unjustifiable standard such as race, *religion*, or other arbitrary classification."[120] Equal protection thus bans religious classifications—and yet free exercise demands them. However this doctrinal tension is resolved,[121] its very existence underscores the question of how anticlassification and anti-anticlassification rules can coherently coexist.

The emergence of MFN doctrine, in short, suggests the fecundity and the difficulty of anticlassification ideas as constitutional decision rules. It also shows that there are many ways to squeeze constitutional inferences from the presence or absence of distinctions on the face of legal enactments—although not all align in an obvious, logical way.

5. How Anticlassification Regimes Cycle

The doctrinal arc of anticlassification does not follow the same path in each of these different bodies of law. But even brief comparison reveals similar changes over time. On the one hand, anticlassification norms tend to calcify: they become more rule-like, less standard-like, and less porous to extratextual evidence. This is visible in the equal protection and free speech contexts. On the other hand, there is also an inflationary tendency within anticlassification regimes. That is, litigants tend to demand that the Court supplement a prohibition

119 Bolling v. Sharpe, 347 U.S. 497, 500 (1954).

120 United States v. Armstrong, 517 U.S. 456, 464 (1996) (emphasis added) (quoting Oyler v. Boles, 368 U.S. 448, 456 (1962)). There is dicta to the same effect construing the Fourteenth Amendment. *See, e.g.*, Burlington N. R.R. Co. v. Ford, 504 U.S. 648, 651 (1992); *cf.* Pena-Rodriguez v. Colorado, 137 S. Ct. 855, 883–84 (2017) (Alito, J., dissenting). But the point here has force even ignoring such dicta.

121 As a practical matter, the free exercise regime dominates.

on facial classifications in anticipation of (or in reaction to) government efforts at circumvention. The worry about proxies for race in *SFFA* and the push for an extraterritorial effect rule in *National Pork Producers Council* are examples of this inflationary tendency.

As a result of these competing pressures, the Justices show unsteady degrees of commitment to parsimonious anticlassificatory rules. They are constantly urged to smudge the boundaries of such rules by tacking on fuzzy-edged standards. Sometimes they yield to that temptation. Sometimes they resist. For this reason, there is not simply a secular trend toward more rigid and more rule-like norms of anticlassification manifest from the 1950s until the present. There is rather a pronounced tendency toward doctrinal "cycle[s]" between rules and standards.[122] The tendency is not absolute: its strength is felt differently across different jurisprudential contexts. Nevertheless, something about anticlassification legal regimes seems to render them wobbly in practice. It is a first hint that anticlassification rules may not be as straightforward as they first seem.

C. The Three Justifications of Anticlassification Legal Regimes

Anticlassification rules are justified on three different normative grounds. The first is an argument from administrability and transparency. The second hinges on the intrinsic, immediate effects of a classification. And the third links classifications to impermissible psychological states of those enacting or enforcing the law. These justifications for anticlassification are cross-cutting: often, the same argument is offered (with small modifications) in different lines of precedent. Yet the three kinds of justifications rest on distinct empirical and normative premises. This yields a tangled skein of argumentation. The Court promiscuously slides between a common yet heterogeneous class of quite divergent normative grounds in justifying formally parallel rules across distinct lines of rights and structural constitutional cases.

Consider these justifications in turn. *First*, anticlassification rules are praised on the ground that they are easy for judges to administer and for the public to understand. They are, as *Reed* said, a matter of "common[ ]sense."[123] Anticlassification rules are indeed informationally sparse in operation. By delimiting severely the universe of relevant

122 *Cf.* Carol M. Rose, *Crystals and Mud in Property Law*, 40 STAN. L. REV. 577, 585 (1988).

123 Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015); *see also* City of Austin v. Reagan Nat'l Advert. of Austin, LLC, 142 S. Ct. 1464, 1482 (2022) (Thomas, J., dissenting) (same). *But see Reed*, 576 U.S. at 167 (suggesting that "[i]nnocent motives" do not obviate the constitutional problem).

evidence, they are formalist and not "functionalist" in character.[124] This has a salutary effect, it is sometimes said, in narrowing the scope of judicial discretion.[125] As Justice Scalia said in another context, judgments are "less subjective" when they turn on "evidence susceptible of reasoned analysis rather than a variety of vague ethico-political First Principles whose combined conclusion can be found to point in any direction the judges favor."[126] Further, as an instrument of antidiscrimination norms, anticlassification rules can appeal to a simple and powerful intuition about the nature of discrimination: if a measure that discriminates (in the sense of drawing a distinction) is not (legally) discrimination—then what is? This intuition is captured by Chief Justice Roberts's famous dictum to the effect that "[t]he way to stop discrimination on the basis of race is to stop discriminating on the basis of race."[127]

Second, anticlassification rules are justified by the harmful effects they immediately and proximately cause. Such *intrinsic* effects are different from the arguments based on the intent of legislators or officials.[128] Even holding constant the behavior and motives of enactors and implementors, the argument goes, a distinction causes harm. The Court has identified two problematic downstream consequences, which can be called "dignity" and "balkanization" effects.

Racial classifications exemplify the first possibility: they are, the Court contends, "perceived by many as resting on an assumption that those who are granted this special preference are less qualified in some respect."[129] Their official use is disapproved because it "demeans the dignity and worth of a person to be judged by ancestry instead of by his or her own merit and essential qualities."[130] Such classifications

124 *City of Austin*, 142 S. Ct. at 1483 (Thomas, J., dissenting); *id.* (praising anticlassification rules as "a clear and firm rule" (quoting *Reed*, 576 U.S. at 171)).

125 *See* Lakier, *supra* note 102, at 253 (noting the argument that content-discrimination rules may be "a more effective means of constraining the repressive and censorial impulses of both legislators and judges").

126 McDonald v. City of Chicago, 561 U.S. 742, 804 (2010) (Scalia, J., concurring). For a similar point, in respect to social-science evidence, see *Roper v. Simmons*, 543 U.S. 551, 617 (2005) (Scalia, J., dissenting).

127 Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 748 (2007).

128 This is the distinction between arguments turning on "the individual's right to respect as an individual" and those resting on a policy's "undesirable consequences." Julie C. Suk, *Quotas and Consequences: A Transnational Re-Evaluation*, *in* PHILOSOPHICAL FOUNDATIONS OF DISCRIMINATION LAW 228, 234 (Deborah Hellman & Sophia Moreau eds., 2013).

129 Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 229 (1995) (quoting Fullilove v. Klutznick, 448 U.S. 448, 545 (1980) (Stevens, J., dissenting)).

130 Rice v. Cayetano, 528 U.S. 495, 517 (2000); *see also* Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2202 (2023) (Thomas, J., concurring) ("Only that promise can allow us to look past our differing skin colors and identities and see each other for what we truly are: individuals with unique thoughts, perspectives,

directly inflict harm on those to whom they apply since being "forced to live under a state-mandated racial label is inconsistent with the dignity of individuals in our society. And it is a label that an individual is powerless to change."[131]

A different variety of intrinsic effect is vividly expressed in cases concerning the race-based gerrymandering of congressional or state legislative districts. This is the worry that the government's use of race "may balkanize us into competing racial factions."[132] Outside the districting context, the Justices have suggested that racial classifications "cause a new divisiveness" and produce a "corrosive discourse."[133] Roughly, the idea here seems to be that explicit racial classifications lower the transaction costs of organizing along racial lines, and so entrench rather than dissipate race-based distinctions.[134] A further unstated assumption here is that such race-based divisions in politics do not already exist.

A parallel concern about balkanization in a different key emerges in the dormant Commerce Clause context. The latter's very existence in the absence of a clear hook in constitutional text has been justified

---

and goals, but with equal dignity and equal rights under the law."); *Parents Involved*, 551 U.S. at 797 (2007) (Kennedy, J., concurring in part and concurring in judgment) ("To be forced to live under a state-mandated racial label is inconsistent with the dignity of individuals in our society."); Miller v. Johnson, 515 U.S. 900, 912 (1995) (criticizing such measures because they "treat individuals as the product of their race" (quoting Metro Broad., Inc. v. FCC, 497 U.S. 547, 604 (1990) (O'Connor, J., dissenting), *overruled by Adarand Constructors, Inc.*, 515 U.S. 200)). Benjamin Eidelson summarizes (and then critiques) this argument as positing that "race-based state actions show a fundamental kind of disrespect for each person's standing as an autonomous, self-defining individual." Benjamin Eidelson, *Respect, Individualism, and Colorblindness*, 129 YALE L.J. 1600, 1604 (2020). Cass Sunstein has distinguished between expressive arguments that turn on law's effect on norms from expressive arguments resting on "the individual interest in integrity." Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2026 (1996). The dignity argument here is of the first kind.

131 *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring in part and concurring in judgment).

132 Shaw v. Reno, 509 U.S. 630, 657 (1993); *accord* Allen v. Milligan, 143 S. Ct. 1487, 1517 (2023).

133 *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring in part and concurring in judgment) ("Governmental classifications that command people to march in different directions based on racial typologies can cause a new divisiveness."). For a further elaboration of this idea, see Reva B. Siegel, *From Colorblindness to Antibalkanization: An Emerging Ground of Decision in Race Equality Cases*, 120 YALE L.J. 1278, 1300 (2011) (elaborating that "pervasive racial stratification can engender anomie and leave some groups feeling like outsiders or nonparticipants").

134 Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 VA. L. REV. 1649, 1713 (2000) (explaining how "clarifying the racial category by statutory definitions facilitated conventions based on race" during the Jim Crow era).

by an appeal to "the Framers' distrust of economic Balkanization."[135] In the First Amendment context, content-based classifications are criticized on the ground that they "may interfere with democratic self-government."[136] While this concern is not further unpacked in the caselaw, the idea seems to be that the content-based restrictions are more likely than other forms of speech restrictions to balkanize public debate into acceptable and unacceptable speech.[137] Government hurdles to flows of information thus impede the creation of a national public good: a free market of truthful information.

Third, anticlassification rules are thought to serve an evidentiary purpose in relation to impermissible intent. That is, they are extrinsically salient. Such justifications come in two forms. These can be called *tracking* and *provocation* logics. First, the presence of a facial classification is said to track the presence of an impermissible legislative motive or an impermissible effect. In the equal protection context, for example, the anticlassification rule is said to "smoke out" illegitimate considerations upstream at the enactment stage.[138] Similarly, in the First Amendment context, content discrimination is explained in terms of the "purposes"[139] served by a regulation, and in particular "improper censorial motive[s]."[140]

Second, a facial classification can also be constitutionally problematic because it can be said to provoke an unconstitutional state of mind in those tasked with implementing the law. That is, an impermissible classification will "require[]" downstream officials to deploy an impermissible consideration in respect to a specific person.[141] This sort of a "provocation" logic bleeds into an argument about effects. In the dormant Commerce Clause context, for example, courts worry about

135 Dep't of Revenue of Ky. v. Davis, 553 U.S. 328, 338 (2008); *see also* Or. Waste Sys., Inc. v. Dep't of Env't Quality of Or., 511 U.S. 93, 98 (1994) (articulating the interest in "avoid[ing] the tendencies toward economic Balkanization"); *accord* Comptroller of the Treasury of Md. v. Wynne, 575 U.S. 542, 548 (2015); Hughes v. Oklahoma, 441 U.S. 322, 325–26 (1979).

136 Reed v. Town of Gilbert, 576 U.S. 155, 174 (2015) (Alito, J., concurring).

137 *See, e.g.*, Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 WM. & MARY L. REV. 189, 198 (1983) (arguing that First Amendment doctrine is primarily concerned with "the extent to which the law distorts public debate").

138 Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 226 (1995) (quoting City of Richmond v. J.A. Croson Co., 488 U.S. 469, 493 (1989) (plurality opinion)); Johnson v. California, 543 U.S. 499, 506 (2005) (same).

139 Ward v. Rock Against Racism, 491 U.S. 781, 791 (1989).

140 Ark. Writers' Project, Inc. v. Ragland, 481 U.S. 221, 228 (1987). In the dormant Commerce Clause context, the Court talks of distinguishing "*bona fide*" reasons from unconstitutional ones. Tenn. Wine & Spirits Retailers Ass'n v. Thomas, 139 S. Ct. 2449, 2464 (2019) (quoting Scott v. Donald, 165 U.S. 58, 91 (1897)).

141 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2155 (2023).

"the pernicious effect on interstate commerce" that arises when a facial distinction is used.[142] In the First Amendment context, even facial classifications enacted with "[i]nnocent motives" are suspect because they may "one day" be used "to suppress disfavored speech."[143] In all of these cases, a statutory classification is viewed with concern because it is assumed to be infused with a risk that subsequent application of the law will be constitutionally problematic. This provocation logic differs from the tracking logic of anticlassification rules because the unconstitutional motives follow, rather than coming before, the law's enactment.

With the exception of dignitary concerns distinctive to the equal protection context, these administrability, extrinsic, and intrinsic arguments for anticlassification rules are offered across different lines of doctrine. Yet they rest on heterogenous empirical and normative foundations. For example, an antibalkanization argument about race is bottomed on assumptions about individual behavior. The homologous argument in the dormant commerce context rests on distinct assumptions about state governments' behavior and incentives. The administrability justification, in contrast, turns on how judges behave and the perils of judicial discretion. Finally, the argument from impermissible intent turns on assumptions about the correlations between statutory text and legislative behavior, as well as the possibility of intentions being ascribed to collective institutional actors such as multicameral legislatures.

The very fact that anticlassification rules rest on plural criteria provides an important clue as to why they might prove unstable in practice. Studies of collective choice have long flagged the possibility of cycling over different decisions when a collective body disagrees about how to apply plural criteria of decision—as is the case with anticlassification rules.[144] This kind of "Arrovian cycling,"[145] as it is known, can also "occur when individuals apply multicriterial predicates."[146] Where a single term (or category) can be justified in a multiplicity of ways, that is, a group of judges who disagree among these grounds can collectively oscillate in ways that show up as doctrinal shifts. Hence, the existence

142 Camps Newfound/Owatonna, Inc. v. Town of Harrison, 520 U.S. 564, 579 (1997) (noting that the law in question was "facially discriminatory," *id.* at 579 n.13).

143 Reed v. Town of Gilbert, 576 U.S. 155, 167 (2015).

144 Matthew L. Spitzer, *Multicriteria Choice Processes: An Application of Public Choice Theory to* Bakke*, the FCC, and the Courts*, 88 YALE L.J. 717, 719–20 (1979).

145 *See* WILLIAM H. RIKER, LIBERALISM AGAINST POPULISM: A CONFRONTATION BETWEEN THE THEORY OF DEMOCRACY AND THE THEORY OF SOCIAL CHOICE 115–19 (1982) (explaining the Arrovian theorem). The canonical work is KENNETH J. ARROW, SOCIAL CHOICE AND INDIVIDUAL VALUES (2d ed. 1963).

146 Ryan D. Doerfler, *Who Cares How Congress Really Works?*, 66 DUKE L.J. 979, 1002 (2017).

of plural justifications for anticlassification rules can lead, thanks to the emergence of Arrovian cycling, to inconsistencies in the way in which doctrine is applied in practice across different cases.

* * *

Constitutional anticlassification rules implement the dormant Commerce Clause, the Article IV Privileges and Immunities Clause, the Equal Protection Clause, and the First Amendment's Free Speech and Free Exercise Clauses. Each line of doctrine, to be sure, has followed a different historical arc—perhaps because of subterranean cycling among disjoint justifications. But each has ended up in a similar place today.

Anticlassification rules hence dominate the law on the ground. The question now is what difficulties they conceal.

II. DELINEATING CONSTITUTIONALLY SALIENT CATEGORIES

We can begin to chart the complexities of anticlassification by asking how a judge embarks on the analysis sketched in Section I.A. At this threshold stage, the questions before a court seem simple: what is the constitutionally relevant category, and how are its constituent instances to be identified? But neither question is quite so unadorned. To illuminate the difficulties of answering these questions, this Part starts by introducing terminology from the philosophy of language—specifically, the distinction between "natural kinds" and "social kinds," as well as variations between three different sorts of "social kinds." With these definitions in hand, I turn to the different categories relevant to constitutional analysis—interstate commerce, content, race, and religion in particular—and explore the difficulties of category creation in each domain.

It is worth noting here that the questions pursued in this Part are distinct from the questions raised at the second step of anticlassification analysis: the latter considers how a judge determines whether one of the category's instances is present on the face of a challenged legal text. I take up that second step in Part III, having explored carefully how constitutionally salient categories are fashioned in the first instance.

A. Natural Kinds and Social Kinds

The first step of an anticlassification analysis is the identification of a *constitutionally salient category*. The leading categories used in the cases discussed in Section I.B are race, content distinctions, interstate commerce, and state citizenship. On initial viewing, these categories look to be self-evident. On closer inspection, none can be applied without controversial linguistic and normative choices.

Outside of law, philosophers of language recognize two importantly distinct kinds of categories: natural kinds and social kinds. There are further distinctions drawn between three different sorts of social kinds. So it is important to understand whether a constitutionally salient category is a natural or a social kind—and if the latter, what sort of social kind—because the election between these kinds conduces to distinct evidentiary protocols and varying levels of stability in the nature of categories. It is worth drawing out these distinctions before applying them to the problem of identifying constitutionally salient categories.

1. Natural Kinds

The idea of a "natural kind" is associated with the philosophers Saul Kripke and Hilary Putnam.[147] They in turn drew on ideas first posited by John Stuart Mill and Bertrand Russell.[148] A natural kind is a term that "exhibits a causal-historical relation to a kind" and that also "refers to all and only members of the kind."[149] Water, for example, is a natural kind that includes all H_2O molecules regardless of whether they are encountered as gas, liquid, or ice. Whether a specific extension[150] or instance is properly called "water" depends on a fact of the matter about the world, and not upon our concepts or systems of mental representation. A natural kind instead directly "links word to world," and so "allows science to be the final arbiter of meaning."[151] The possibility and operation of a natural kind are philosophically contested.[152] On one view, there are arbitrarily many ways of carving up the natural world at the joints. An appeal to "science" does not explain which one counts.[153] Those critiques can be bracketed here, so long as the general idea is grasped.

147 *See* SAUL A. KRIPKE, NAMING AND NECESSITY 5 (1980); HILARY PUTNAM, MIND, LANGUAGE AND REALITY, at xiii (1975).

148 *See, e.g.*, KRIPKE, *supra* note 147, at 26–27; PUTNAM, *supra* note 147, at 10, 51, 57.

149 Stephen Laurence & Eric Margolis, *Concepts and Cognitive Science*, *in* CONCEPTS: CORE READINGS 3, 23 (Eric Margolis & Stephen Laurence eds., 1999) (discussing KRIPKE, *supra* note 147, and PUTNAM, *supra* note 147); *see also* Hacking, *supra* note 15, at 110–22 (offering several criteria for natural kinds, and then charting the development of that term); Kimberly Kessler Ferzan, *A Planet by Any Other Name . . .*, 108 MICH. L. REV. 1011, 1017–18 (2010) ("[S]omething (like water, tigers, and gold) is thought to be a natural kind when the term rigidly designates the 'real' essence of the item.").

150 The "extension" of a term is the set of things the term is true of. PUTNAM, *supra* note 147, at 216.

151 Ferzan, *supra* note 149, at 1020.

152 *See, e.g.*, Richard Rorty, *Texts and Lumps*, 39 NEW LITERARY HIST. 53, 56–60 (2008).

153 Hence, Hacking offers "utility" as one defining trait of a natural kind. Hacking, *supra* note 15, at 110. Clearly, a natural kind that lumped carbon dioxide and oxygen together wouldn't meet this criterion.

2. Social Kinds

Standard examples of social kinds include "marriage," "money," and "presidents." While these examples are uncontroversial, the definition of "social kind" is contested. In an influential recent treatment, Muhammad Ali Khalidi defines them as follows: "[F]or some social kind, *x*, to be *x* is simply to be regarded as, used as, and believed to be *x*."[154] In contrast, Rebecca Mason offers two definitions. She notes that a kind may be social "if the conditions for kind membership involve social properties and relations" or, alternatively, if they "depend[] on collective intentions or other attitudes for [their] existence or nature."[155] There is also disagreement in the philosophy of language as to whether natural kinds and social kinds are mutually exclusive of each other.[156] Social kinds, however, are generally understood to be distinct from natural kinds insofar as they are "ontologically subjective in the sense that their very existence depends on our propositional attitudes towards them."[157] That said, Amie Thomasson has pointed out that some social kinds are not directly "ontologically subjective" in the sense that their existence does not depend on people having certain attitudes or thoughts about specific instances.[158] She has offered the examples of "recessions" and "racism."[159] These are "mind-dependent" in the sense that some "human mental states need to be in place for the kind to exist at all."[160] But there would be "poverty" even if we lacked specific attitudes toward the poor.

Supplementing these definitions, Khalidi offers a typology of social kinds useful for evaluating constitutionally salient categories.[161] First, as Thomasson argued, there are social kinds requiring no attitudes or thoughts, but that depend on other social kinds (e.g.,

154 Muhammad Ali Khalidi, *Three Kinds of Social Kinds*, 90 PHIL. & PHENOMENOLOGICAL RSCH. 96, 98 (2015) (citing JOHN R. SEARLE, THE CONSTRUCTION OF SOCIAL REALITY 32 (1995)).

155 Mason, *supra* note 15, at 841–42.

156 For a summary of the debate, see MUHAMMAD ALI KHALIDI, NATURAL CATEGORIES AND HUMAN KINDS: CLASSIFICATION IN THE NATURAL AND SOCIAL SCIENCES, at xiii (2013); and Mason, *supra* note 15, at 842 (arguing for overlap).

157 Khalidi, *supra* note 154, at 98–99. John Searle, who is credited with catalyzing research on social kinds, distinguished "observer-independent" from "observer-relative" phenomena. John R. Searle, *Social Ontology and the Philosophy of Society*, 20 ANALYSE & KRITIK 143, 147–48 (1998). Social kinds, he suggested, turn on propositional attitudes in a way that natural kinds do not. *Id.*

158 Amie L. Thomasson, *Foundations for a Social Ontology*, 18–19 PROTOSOCIOLOGY 269, 271 (2003) (emphasis omitted).

159 *Id.* at 276.

160 Khalidi, *supra* note 154, at 104.

161 *Id.* at 99–104.

"racism," "recession").[162] Second, there are social kinds that are "*partly* dependent on specific attitudes," where specific attitudes do not need to be in place for an instance to count as an instance of the kind.[163] Khalidi offers the example of "war" here: armed hostilities can count as "war" even if the label is denied by the participants.[164] Third, there are some social kinds "whose existence and that of their instances are both dependent on attitudes that human beings have towards them."[165] Take the term "pope." Between 1378 and 1417, there was a "Western Schism" leading to two separate papal courts and two different men calling themselves "pope."[166] There was a shared social kind of "pope," but disagreement about the individualized extensions of that kind.[167] In this third form, social kinds are "response-dependent"; that is, "something constitutes *z* just in case it is taken to constitute *z*," and "anything that constitutes *z* does so in virtue of being taken to constitute *z*."[168] The philosopher Kit Fine offers the example of "cool" as a social kind that is response dependent: something is cool, that is, just in case we take it to be "cool."[169] I will call the last two sorts of social kinds *response-independent social kinds* and *response-dependent social kinds*. That distinction will become relevant when discussing "race" as a social kind in Section II.D.

B. Using Natural Kinds or Social Kinds to Build Constitutionally Salient Categories

Distinctions between natural kinds and social kinds provide a way of sorting among constitutionally salient categories and explaining why

---

162 *Id.* at 99.

163 *Id.* at 100 (emphasis added).

164 *Id.*

165 *Id.* at 101.

166 JOËLLE ROLLO-KOSTER, THE GREAT WESTERN SCHISM, 1378–1417: PERFORMING LEGITIMACY, PERFORMING UNITY 1–2 (2022).

167 *Id.* at 5 (noting how the term "usurper" was used to signal during the crisis).

168 Asya Passinsky, *Social Objects, Response-Dependence, and Realism*, 6 J. AM. PHIL. ASS'N 431, 434 (2020) (emphasis omitted). This is not the only way in which this relation is described. Sveinsdóttir suggests that response-dependence implies "there is something in the object that induces or causes the response in question." Ásta Kristjana Sveinsdóttir, *The Social Construction of Human Kinds*, 28 HYPATIA 716, 721 (2013). She proposes an alternative "conferralism" account, in which there is a "grounding property" and "also on top of it . . . the social property." *Id.* at 728. I do not read Khalidi or Passinsky, however, to require causation. Sveinsdóttir's concept also seems vague in ways that may undermine the legal analysis. Hence my decision to rely on response-dependence here.

169 Kit Fine, The Structure of Joint Intention 16–17 (2010) (unpublished manuscript), https://www.academia.edu/42971332/The_Structure_of_Joint_Intention [https://perma.cc/45BY-7G3H]. Fine identifies two conditions, safety and harmony, that must be met for this definition to be "good." *Id.* at 17–18.

only some raise distinctive and unusual difficulties of legal analysis. To see the roles played by natural kinds and social kinds in anticlassification legal regimes, it is useful to start with a basic question: Why not look to the law itself for the categories employed at step one? Why do we need the distinction between natural and social kinds at all?

The most intensely litigated categories used in anticlassification doctrine are not picked out in the Constitution's text. To be sure, Article IV does mention "Citizens of each State."[170] But the Equal Protection Clause does not mention race.[171] The First Amendment talks of "speech" but does not mention "content."[172] And while Article I, Section 8, does talk of "Commerce . . . among the several States,"[173] it does so for the limited purpose of defining congressional power, not drawing up an impermissible zone of state action. Most categories hence are extrinsic to the Constitution's text.

Constitutionally salient categories, moreover, have not been a constant fixture of the jurisprudence. They have instead waned or flourished at various times. For instance, the Court once rejected the salience of "race" as a category under the Equal Protection Clause.[174] At least according to some scholars, it looked askance at content neutrality as a First Amendment ideal until the 1970s.[175] Its first articulation of the preemptive effect of the Commerce Clause turned upon the distinction between the matters "in their nature national" and those "local and not national,"[176] even though this language restates (and does not resolve) the problem to be addressed.[177] With one

170 U.S. CONST. art. IV, § 2, cl. 1.

171 *Id.* amend. XIV, § 1.

172 *Id.* amend. I.

173 *Id.* art. I, § 8, cl. 2.

174 *See, e.g.*, The Civil Rights Cases, 109 U.S. 3, 25 (1883) ("When a man has emerged from slavery, . . . there must be some stage in the progress of his elevation when he . . . ceases to be the special favorite of the laws . . . .").

175 Jud Campbell, *The Emergence of Neutrality*, 131 YALE L.J. 861, 865 (2022) ("From the Founding through the mid-twentieth century, the freedom of speech entailed a limited right of toleration, not neutrality."). *But see* Genevieve Lakier, *A Counter-History of First Amendment Neutrality*, 131 YALE L.J.F. 873, 875 (2022) (tracing "a demand that the government act neutrally with respect to the content of the speech" back to "the eighteenth century").

176 Cooley v. Bd. of Wardens, 53 U.S. (12 How.) 299, 319 (1852), *abrogated by* N.J. Bell Tel. Co. v. State Bd. of Taxes & Assessments, 280 U.S. 338 (1930), *as recognized in* Okla. Tax Comm'n v. Jefferson Lines, Inc., 514 U.S. 175 (1995).

177 5 CARL B. SWISHER, THE OLIVER WENDELL HOLMES DEVISE: HISTORY OF THE SUPREME COURT OF THE UNITED STATES; THE TANEY PERIOD 1836–64, at 406 (Paul A. Freund ed., 1974) ("[*Cooley*] left unanswered the question whether the Court would find to be local any interstate or foreign commerce other than that which Congress had designated as such . . . ."). In any case, the *Cooley* formulation did not endure. The subsequent adoption of a direct/indirect test in dormant Commerce Clause doctrine is charted in Barry Cushman, *Formalism and Realism in Commerce Clause Jurisprudence*, 67 U. CHI. L. REV. 1089 (2000).

exception, therefore, the verbal contents of the Constitution supply no evidence of how the constitutionally salient category ought to be drawn. Instead, the categories used in anticlassification rules must draw substance from beyond the Constitution's four corners.

Bereft of an anchor in constitutional text, several lines of cases haphazardly oscillate between different verbal formulations of the category itself. The First Amendment and the dormant Commerce Clause illustrate the tendency most clearly. The leading content-discrimination precedent, for example, speaks of "discrimination" on the basis of an expression's "message, its ideas, its subject matter, or its content."[178] Leslie Kendrick has collected further synonyms for content from the secondary literature, including "message," "substance," "meaning," or "communicative significance."[179] None of these terms are precisely defined; many are vague. The result, she fairly notes, is that the very nature of the category "content discrimination" is "somewhat of a chameleon."[180]

In the dormant Commerce Clause context, the Court alternates between talking of "patent discrimination against interstate trade"[181] and the advantaging of "local consumers" at the expense of "consumers in other States."[182] In many instances, this distinction will not matter. But in other moments, it does. In *General Motors v. Tracy*, for example, the Court upheld a use tax on natural gas imposed on out-of-state but not in-state firms because of the latter's distinctive function in the heavily regulated in-state energy market.[183] The "absence of actual or prospective competition" between the two kinds of entities meant there simply was "no local preference."[184] *Tracy* suggests that nonlocal market participants can be disfavored without an effect on "interstate commerce" as such. But *Tracy* did not lead to a general

178 Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95–96 (1972) (first citing Cohen v. California, 403 U.S. 15, 24 (1971); then citing Street v. New York, 394 U.S. 576, 592 (1969); then citing N.Y. Times Co. v. Sullivan, 376 U.S. 254, 269–70 (1964); then citing NAACP v. Button, 371 U.S. 415, 445 (1963); then citing Wood v. Georgia, 370 U.S. 375, 388–89 (1962); then citing Terminiello v. Chicago, 337 U.S. 1, 4 (1949); and then citing De Jonge v. Oregon, 299 U.S. 353, 365 (1937)). On the pathbreaking role of *Mosley* in First Amendment doctrine, see Paul B. Stephan III, *The First Amendment and Content Discrimination*, 68 VA. L. REV. 203, 223–27 (1982).

179 Kendrick, *supra* note 98, at 244.

180 Kendrick, *supra* note 101, at 216.

181 City of Philadelphia v. New Jersey, 437 U.S. 617, 624 (1978).

182 Brown-Forman Distillers Corp. v. N.Y. State Liquor Auth., 476 U.S. 573, 580 (1986) ("Economic protectionism is not limited to attempts to convey advantages on local merchants; it may include attempts to give local consumers an advantage over consumers in other States." (citing New Eng. Power Co. v. New Hampshire, 455 U.S. 331, 338 (1982))).

183 Gen. Motors Corp. v. Tracy, 519 U.S. 278, 300–03, 310 (1997).

184 *Id.* at 300.

correction in the Court's verbal formulations: Justices still inveigh against state laws advantaging only "local consumers."[185]

Verbal instability in constitutionally salient categories is hard to square with the doctrine's administrability justification. If the Justices are not consistent in describing the type at issue when examining an enactment's text in search of a specific token, it is reasonable to expect legal uncertainty. Efforts to stabilize the doctrine, moreover, remain vulnerable to cycling between different specifications of salient categories. In the First Amendment context, for instance, the Court in 2015 tried to rationalize content discrimination in *Reed v. Town of Gilbert*.[186] Yet that decision still left lower courts with "considerable discretion to determine how broadly the decision applies."[187] Seven years later, however, the Court retreated from *Reed*'s strict formalism in favor of a more porous understanding of content discrimination.[188] In this way, the underlying verbal pluralism of constitutionally salient categories is one enabler of the cycling between standards and rules identified in subsection I.B.5.

All this means both natural and social kinds now play a role in the doctrine. If categories cannot be derived from legal sources, they must be taken from the world external to law. All the key categories used by the contemporary Court—"content," "race," "interstate commerce," and "religion"—have this quality. For this reason, anticlassification in practice rests on the judicial identification of a natural kind or a social kind. Either one or the other must play a necessary role as a relevant category when an anticlassification rule is deployed.

With one important exception, the Court appears to lean on social kinds rather than natural kinds when it draws up constitutionally salient categories. There is no natural kind of "content" or "interstate commerce" or "religion." That is, one cannot turn to "science" as a way of defining any of these categories. All instead depend, directly or remotely, upon the existence of beliefs, attitudes, or orientations among individuals. The question then becomes how such propositional attitudes of a regulated community can or should be cashed out in a doctrinal form.

But the categories used in the equal protection context, such as "race" and "gender," raise distinctive concerns: Taking race as the leading example, it is possible to identify cases in which the Court has treated it as a natural kind. At other moments, it has been taken as a social kind. Equal protection classification, hence, warrants a distinct

185 Camps Newfound/Owatonna, Inc. v. Town of Harrison, 520 U.S. 564, 578 (1997) (quoting *Brown-Forman Distillers*, 476 U.S. at 580).

186 Reed v. Town of Gilbert, 576 U.S. 155, 168–69 (2015).

187 Lakier, *supra* note 102, at 254.

188 City of Austin v. Reagan Nat'l Advert. of Austin, LLC, 142 S. Ct. 1464, 1471 (2022).

and separate treatment. Therefore, I offer first a general analysis of how anticlassification doctrine uses social kinds, and then drill down on equal protection.

C. Social Kinds as Constitutionally Salient Categories: Interstate Commerce, Content, and Religion

A careful parsing of how the categories "interstate commerce," First Amendment "content," and "religion" are used points to three rather different ways of building a constitutionally salient category out of various social kinds. In drawing out these differences, I direct attention to how a social kind is used to fix the boundaries of a category. *How* a category is defined, or "scoped," determines *what* it comprises. Analytic method, in this way, filters constitutional meaning.

1. Interstate Commerce

Consider first the category of "interstate commerce." This is a social kind: interstate borders exist because people hold certain attitudes toward different territories and the legal demarcation of their difference.[189] But "interstate commerce" is not "ontologically subjective" in the sense that its existence does not depend on people having certain attitudes or thoughts about the specific goods or services at issue.[190] It is "mind-dependent" in the thinner sense that some "human mental states need to be in place for the kind to exist at all."[191] "Interstate commerce" might be metaphorically said to lie on top of a set of propositional attitudes about borders. This sort of more limited mind dependency might be thought to be reasonably tractable because it leans on a relatively abstract level of social consensus. Indeed, the dormant Commerce Clause does evince a certain stability when it comes to the architecture of categories.

Recognizing "interstate commerce" as a "social kind" casts light on why the doctrine is malleable and does not exclude all judicial discretion. "Interstate commerce" describes an activity, not a finite, countable set of tangible items. It is a "mass noun[]," that is, and not a "count noun[]."[192] Mass nouns are vague in a way that count nouns are not. As Gennaro Chierchia puts it, "heaps and mountains are vague and context dependent in very different ways than rice or

189 Passinsky, *supra* note 168, at 435–36 (showing this with the example of state borders).

190 *See* Thomasson, *supra* note 158, at 276; Khalidi, *supra* note 154, at 104.

191 Khalidi, *supra* note 154, at 104.

192 Gennaro Chierchia, *Mass Nouns, Vagueness and Semantic Variation*, 174 SYNTHESE 99, 100–02 (2010) (terming this an "object versus substance contrast," *id.* at 102).

sand."[193] Dormant Commerce Clause doctrine, indeed, evinces sensitivity to the presence of mass-related vagueness. Hence, the *Tracy* Court ignored a facial distinction when it perceived market differentiation between in-state and out-of-state actors.[194] But the Justices elsewhere rejected constitutional claims when an insufficient fraction (count) of out-of-state actors were affected.[195] They also toggle between embracing and rejecting discrimination arguments based on the extent to which out-of-state actors are impacted.[196] The mass-related vagueness of "interstate commerce," that is, offers the Court a degree of flexibility that renders case-level outcomes unpredictable. It invites the exercise of judicial discrimination, albeit in a sub-rosa way.

2. Content Discrimination

"Content" discrimination in First Amendment jurisprudence evinces a different kind of challenge for judges. Unlike "interstate commerce," there is no social kind of "content neutrality" independent of legal discourse to anchor this category. As Kendrick perceptively notes, "the term ['content'] could mean any number of things."[197] As a result, there is no set of social or empirical facts that judges can use as a benchmark to evaluate whether a law is content discriminatory or content neutral. In the absence of a stable, external benchmark to define "content discrimination," vague terms for the category proliferate in caselaw and secondary commentary. Absent such an external benchmark, "content" becomes a "social kind" that originates from, and derives its force exclusively from, the social practice of lawyers and judges during the course of constitutional litigation.[198] The term lacks resolving power without a recourse to some independent normative theory of the First Amendment.

Two examples illustrate this point. A first is found in the 2014 opinions in *McCullen v. Coakley*, in which the Court struck down a

193 *Id.* at 123.

194 Gen. Motors Corp. v. Tracy, 519 U.S. 278, 300–03 (1997); *see supra* text accompanying notes 183–84.

195 Exxon Corp. v. Governor of Md., 437 U.S. 117, 127 (1978).

196 *Compare* Kassel v. Consol. Freightways Corp. of Del., 450 U.S. 662, 677–78 (1981) (plurality opinion) (invalidating a facially neutral highway safety rule on the ground that "Iowa's statute may not have been designed to ban dangerous trucks, but rather to discourage interstate truck traffic," *id.* at 677), *with* Minnesota v. Clover Leaf Creamery Co., 449 U.S. 456, 460, 471–72 (1981) (rejecting trial court's finding of impermissible protectionist purpose and finding no discrimination).

197 Kendrick, *supra* note 98, at 242.

198 For a suggestion that official determinations of a social kind's scope will often be dispositive, see Passinsky, *supra* note 168, at 436 (describing these as "authority-involving social concepts").

Massachusetts statute that made it a crime to knowingly stand on a public way or a sidewalk within 35 feet of a facility in which abortions were being performed.[199] While all the Justices agreed that the statute was invalid under the First Amendment, they disagreed as to whether it discriminated on the basis of content. Taking the law to be "facially neutral," the majority opinion of Chief Justice Roberts trained on how the "stated purpose" of the law was "content neutral."[200] Concurring in the judgment, Justice Scalia replied with characteristic acid: it "blinks reality," he said, to treat "a blanket prohibition . . . on only one politically controversial topic" as content neutral.[201]

This disagreement suggests that the Justices disagreed about the scope of the category "content." On the one hand, Chief Justice Roberts understood it as being defined in juxtaposition to time, place, and manner restrictions. On the other hand, Justice Scalia perceived overlap between content restrictions and time, place, and manner restrictions. "[S]cience" does not help us discern who is right.[202] There is also no linguistic or social practice outside the legal system to which we can appeal to discern who has the better argument. We have to have a normative theory of the First Amendment to broker the disagreement.

The Court's more recent efforts to clarify the scope of content discrimination offer a second example of the same difficulty. In *Reed v. Town of Gilbert*, the Court crafted a more rule-like doctrine focused on whether a law "is content based on its face" without regard to legislative motive.[203] *Reed* holds that "laws that employ content distinctions are always content-based, regardless of the purposes they serve."[204] But the Court quickly fractured over exactly what "content based" meant. In *Reed*, the Court held that a "speaker based" distinction is a kind of content-based discrimination.[205] Concurring, Justice Kagan expressed concerns about the breadth of that rule. She suggested that the doctrine should be administered "with a dose of common sense" so as not to reach "laws that in no way implicate its intended function," including some speaker-based laws.[206] Seven years later, a different majority of the Court in *City of Austin v. Reagan National Advertising of Austin,*

199 McCullen v. Coakley, 573 U.S. 464, 469–72, 497 (2014).

200 *Id.* at 480.

201 *Id.* at 501 (Scalia, J., concurring in judgment).

202 *See* Ferzan, *supra* note 149, at 1020.

203 Reed v. Town of Gilbert, 576 U.S. 155, 165 (2015) (citing City of Cincinnati v. Discovery Network, Inc., 507 U.S. 410, 429 (1993)). The Court also installed a secondary inquiry into "the purpose and justification for the law." *Id.* at 166.

204 Lakier, *supra* note 102, at 234 (emphasis omitted).

205 *Reed*, 576 U.S. at 169, 169–71.

206 *Id.* at 183 (Kagan, J., concurring in judgment).

LLC rejected *Reed*'s categorical treatment of speaker-based regulations as content-based, provided that the state's consideration of speakers was "only in service of drawing neutral, location-based lines."[207] In effect, *City of Austin* rejected *Reed*'s position that *all* speaker-based distinctions are instances of content-based distinctions. It said that only *some* are based on Justice Kagan's earlier worries about the "intended function" of content-discrimination review.

Neither the *Reed* majority nor the *City of Austin* majority could point to a definition of "content" independent of constitutional law. And neither leaned on an extrinsic benchmark grounded in empirical evidence or social practice. There simply is no immediately available social kind of "content distinctions" upon which a judge can rely. As a result, it is again difficult to see how the Court *could* resolve the question of how to taxonomize speaker distinctions without making a normatively freighted judgment about what the First Amendment is for.

The evolution of content-discrimination doctrine, in short, demonstrates at multiple points how an anticlassification rule can demand an appeal to a nonlegal social kind in order to delimit its bounds. But it also points to how the absence of a readily available social kind means that the Court must instead resort to the sort of first-order normative reasoning that the "commonsense" doctrine of anticlassification is supposed to make redundant.[208]

3. Religion

The third example of a social kind being used as a constitutionally salient category is religion in MFN cases.[209] The underlying assumption of the anti-anticlassification rule used in those decisions is that it is possible to carve out an exception for religion when there is an exception for a "comparable" secular activity.[210] The Court assumes that legislators can carve up the world into "religious" and "secular" activities.[211] Of course, as judges and scholars have long been aware, it is painfully difficult to draw this line. Courts have, for example, denied

207 City of Austin v. Reagan Nat'l Advert. of Austin, LLC, 142 S. Ct. 1464, 1471 (2022).

208 *See Reed*, 576 U.S. at 163.

209 *See supra* subsection I.B.4.

210 Roman Cath. Diocese of Brooklyn v. Cuomo, 141 S. Ct. 63, 66 (2020).

211 Judicial recognition of the distinction long predates the emergence of MFN status. *See, e.g.*, Hosanna-Tabor Evangelical Lutheran Church & Sch. v. EEOC, 565 U.S. 171, 189 (2012) (rejecting as "untenable" the notion that religion should be assimilated into a general right of association for constitutional purposes); Wisconsin v. Yoder, 406 U.S. 205, 215 (1972) ("A way of life, however virtuous and admirable, may not be interposed as a barrier to reasonable state regulation of education if it is based on purely secular considerations; to have the protection of the Religion Clauses, the claims must be rooted in religious belief.").

the status of religion to "New Age" practitioners[212] and sincere, if parodic, faith claims.[213] Meanwhile, at least one other court has found atheists as such to be shielded under the Free Exercise Clause.[214]

Despite the "famously difficult" challenge of defining religion,[215] there is a longstanding technique for scoping that category for legal purposes. In a series of cases concerning conscientious objectors to the draft, the Court has asked whether a particular set of beliefs has "in the life of that individual 'a place parallel to that filled by . . . God' in traditionally religious persons."[216] Earlier judgments relied on other, also putatively prototypical, features of religion as a social practice to distinguish it from the secular. In these cases, Justices implicitly deploy a mechanism of category formation identified as common to human and animal cognition called "prototype theory," by which people "average their experiences with disparate exemplar to form the schema, the prototype, the central tendency of the category."[217] A range of cognitive studies suggests that prototyping usually offers the "clearest signals" for classification, but that "poorly-structured" categories can "defeat prototype processing and demand exemplar memory."[218] The cognitive efficacy of a prototype hence depends on the existence of regular "structure in the perceived world."[219]

Reliance on religious prototypes leads the Court to apply MFN analysis only when familiar, but not unusual, religious practices are burdened by a law. Because the Court takes certain forms of Christianity as "central" (i.e., prototypical) to the category of religion, the strength of judicial solicitude for a religious practice often turns on its

212 *See, e.g.*, Moore-King v. County of Chesterfield, 708 F.3d 560, 564, 570–71 (4th Cir. 2013) (finding that a "New Age" "psychic" with "a strong belief in the 'words and teachings of Jesus'" did not rank as religious).

213 *See, e.g.*, Cavanaugh v. Bartelt, 178 F. Supp. 3d 819, 830 (D. Neb. 2016) (finding that the Church of the Flying Spaghetti Monster is not entitled to First Amendment protections).

214 Kaufman v. McCaughtry, 419 F.3d 678, 681–82 (7th Cir. 2005) (citing, inter alia, United States v. Seeger, 380 U.S. 163, 184–88 (1965)). Scholars also offer a range of definitions of religion. *See* Nelson Tebbe, *Nonbelievers*, 97 VA. L. REV. 1111, 1130–40 (2011) (outlining various academic and legal definitions of religion).

215 Mark L. Movsesian, *The New Thoreaus*, 54 LOY. U. CHI. L.J. 539, 543 (2022).

216 Welsh v. United States, 398 U.S. 333, 340 (1970) (plurality opinion) (omission in original) (quoting *Seeger*, 380 U.S. at 176); *accord Seeger*, 380 U.S. at 166. Kent Greenawalt has summarized the ensuing jurisprudence as a search for "family resemblance" in novel fact patterns. 1 KENT GREENAWALT, RELIGION AND THE CONSTITUTION: FREE EXERCISE AND FAIRNESS 139, 139–40 (2006).

217 Smith et al., *supra* note 16, at 2; *accord* Rosch, *supra* note 38, at 31 (explaining that categories are defined by "the weighted sum of the measures of all of the common features within a category minus the sum of the measures of all of the distinctive features").

218 Smith et al., *supra* note 16, at 9, 11.

219 Rosch, *supra* note 38, at 29.

similarity to the most familiar, majoritarian form of religiosity. The analytic structure of category scoping, therefore, implicitly embeds a (notionally impermissible) doctrinal preference.[220] It identifies religious practices by looking for idiosyncratic features of a canonical faith.[221] In this way, the Court's analysis is selectively porous to "real-world knowledge" about some faiths and not others.[222] Nothing in the Court's recent MFN jurisprudence suggests recognition of this concern. To the contrary, decisions applying MFN doctrine to COVID-19 regulations consistently preferred "religious groups and beliefs that involve gathering in large groups [a feature of Christianity, but not of all faiths] over those that do not."[223]

While the law's reliance on a social kind allows the Justices' priors about what counts as a religion to seep into the doctrine, the social kind of "religion" generates little analytic friction to prevent the resulting distortions. It is commonplace in the social science literature to observe that there is no central case of "religion."[224] Rather, "religion" is structurally varied in terms of individual actions, beliefs, and forms of collective sociality (if any) such that there is no clear core case. But if there is no way of constructing the category of religion through prototypical theory without embracing an impermissible doctrinal preference, then there is a risk that the ensuing classification will be unprincipled or arbitrary in scope. This risk is acute because prototype theory is "unable to account for the phenomenon of compositionality."[225] That is, when a phenomenon is pervasively characterized by internal, structural complexity (or "compositionality"[226]), the psychological

220 *See* Larson v. Valente, 456 U.S. 228, 244 (1982) ("The clearest command of the Establishment Clause is that one religious denomination cannot be officially preferred over another.").

221 For instance, the Court allowed the government to exempt from the military draft objectors to all wars rather than only particular ones on administrability grounds. *See, e.g.*, Gillette v. United States, 401 U.S. 437, 454–55 (1971). Embedded in the administrability analysis is a prototypical assumption that a religious view on war must track just war theory in the Christian tradition. *See id.* at 455–56.

222 *See* Laurence & Margolis, *supra* note 149, at 41.

223 Simson, *supra* note 116, at 1644.

224 *See, e.g.*, ELIZABETH SHAKMAN HURD, BEYOND RELIGIOUS FREEDOM: THE NEW GLOBAL POLITICS OF RELIGION 6 (2015) ("Religion is too unstable a category to be treated as an isolable entity . . . ."); CÉCILE LABORDE, LIBERALISM'S RELIGION 20 (2017) ("There is no feature, or set of features, that all religions share.").

225 Laurence & Margolis, *supra* note 149, at 37. The reason appears to be that prototyping works by averaging across the frequency of specific features, and does not account for how those features can or do fit together. *Id.* at 38–39.

226 Zoltán Gendler Szabó, *Compositionality*, STAN. ENCYC. PHIL. (Aug. 17, 2020), https://plato.stanford.edu/entries/compositionality/ [https://perma.cc/3HUP-9A3F] ("The meaning of a complex expression is determined by its structure and the meanings of its constituents.").

mechanisms of prototype creation and extension are likely to break down.

4. The Difficulty of Social Kinds for Law

Constitutionally salient classification can lean upon a variety of social kinds. These dependencies, however, lead to a range of hard but unrecognized analytic and normative choices. Justices now glide over these questions. The complexities created by the law's reliance on social kinds should be a matter of more sharply focused debate.

D. Equal Protection: Race as Natural Kind or Social Kind

One kind of constitutionally salient category merits especially close examination. Equal protection caselaw oscillates between natural kinds and social kinds as the touchstone of constitutionally salient categories. Neither tack, however, can be followed without considerable analytic difficulty and tricky normative choices. Yet they are neither recognized nor expressly addressed by the Court. Here, I focus on "race" as a lens to illustrate this phenomenon, although a similar exercise might be pursued with "gender."

1. How the Court Oscillates Between Race as a Natural and a Social Kind

The Court oscillates between understanding race as a natural kind and as a social kind. It does so without recognizing that it is waffling over a key element of anticlassification analysis. The dominant approach—treating race as a natural kind—is at odds with both leading scientific and social-scientific theory. It is also more likely to exacerbate than abate the very harms that the Court purports to be addressing. The judicial choice of category construction, in other words, has profound and wide-ranging effects on the Equal Protection Clause's goals. These, no doubt, merit more extended treatment than I offer here. And my focused aim in this section is simply to demonstrate the stakes of scoping the category of race as either a natural or as a social kind.

The Roberts Court often assumes race is a natural kind. In a series of cases culminating in *SFFA*, the Court has described race as a distinction "between citizens solely because of their ancestry."[227] On this

227 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2162 (2023) (quoting Rice v. Cayetano, 528 U.S. 495, 517 (2000)); *accord* Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 214 (1995); Hirabayashi v. United States, 320 U.S. 81, 100 (1943). There is language in *Rice v. Cayetano* suggesting that "[a]ncestry *can* be a proxy for race." *Rice*, 528 U.S. at 514 (emphasis added); *see also* Davis v. Guam, 932 F.3d

view, the category of race is comprised of classifications that turn upon the fact of "ancestry." Although the Court does not define "ancestry," it normally refers to a fact that can be verified using scientific methods: i.e., identification of those from whom a person is descended.[228] Implicit in the idea of "ancestry" as a parameter in classification is the further notion that there are group-level differences between descendants that can be cashed out as instances of categories in law.[229] In other words, "race" as *SFFA* operationalizes it is a type realized through a set of biological facts amenable to scientific inquiry into potential population-level differences. Consistent with this view, the Court posits race as ancestry—a biological kind with no space for individual agency—to explain its dignity-related justification of anticlassification rules in the equal protection context.[230] In so doing, it takes as given that races are defined "independent of human belief"—which is simply to say that it assumes "races are natural kinds."[231]

Illustrating the judicial construal of race in biological terms, the *SFFA* majority defined impermissible racial classifications by reference to the "ignoble history" of "state-mandated segregation" at the end of the nineteenth century.[232] By implication, what government cannot lawfully do now is rank people in the way, say, Louisiana did at the turn of the twentieth century. Then, the state classified a person by asking whether they had more than one in thirty-two parts Black ancestry.[233] So construed, the Equal Protection Clause places out of bounds legal classifications that, like those from the Jim Crow era, track the biological fact of ancestry. This takes race as a natural kind.

But in other, mostly statutory, contexts, the Court has explicitly abjured a biological understanding of race as a natural kind.

822, 834 (9th Cir. 2019) ("Our first inquiry is whether . . . *Rice* held *all* classifications based on ancestry to be impermissible proxies for race. It did not."). But this distinction appears to collapse by *SFFA*.

228 *See* Rolnick, *supra* note 56, at 2685 ("Ancestry refers to the genetic or historical connection between a living person and those progenitors who preceded the person, sometimes through generations. It connects a living person to their parents, grandparents, or other relatives who came before.").

229 That is, ancestors are characterized in racial terms too. Note the potentially infinite regress packed into this definition: One's race depends on the race of one's ancestors. But how is their race defined?

230 *Rice*, 528 U.S. at 517 ("[Racial classification] demeans the dignity and worth of a person to be judged by ancestry instead of by his or her own merit and essential qualities.").

231 CHARLES W. MILLS, "*But What Are You* Really*?": The Metaphysics of Race*, *in* BLACKNESS VISIBLE: ESSAYS ON PHILOSOPHY AND RACE 41, 46 (1998).

232 *Students for Fair Admissions*, 143 S. Ct. at 2149.

233 Act No. 46, 1970 La. Acts 167 (repealed 1983); *see also* Paul Finkelman, *The Color of Law*, 87 NW. U. L. REV. 937, 955 n.96 (1993) (reviewing ANDREW KULL, THE COLOR-BLIND CONSTITUTION (1992)) (listing "what fraction of ancestry that is black would lead to someone being legally" Black in the various southern states in the early twentieth century).

Construing a Civil Rights era antidiscrimination statute in 1987, the Court stated that "genetically homogeneous populations do not exist and traits are not discontinuous between populations; therefore, a population can only be described in terms of relative frequencies of various traits. Clear-cut categories do not exist."[234] Glossing the 1790 Naturalization Act's limitation of citizenship to "free white person[s],"[235] the Court in 1922 recounted how "federal and state courts, in an almost unbroken line, ha[d] held that the words 'white person' were meant to indicate only a person of what is *popularly known* as the Caucasian race."[236] Rejecting the alien plaintiff's arguments from scientific definitions of race in a subsequent case, it held that "the word [Caucasian] by common usage has acquired a popular meaning, not clearly defined to be sure, but sufficiently so to enable us to say that its popular as distinguished from its scientific application is of appreciably narrower scope."[237] The Court ruled that "'free white persons' are words of common speech, to be interpreted in accordance with the understanding of the common man."[238]

The latter "popular" understanding of "who was white" is in tension with the Court's embrace elsewhere of a biological conception.[239] It takes racial kinds as social rather than natural kinds. Race, on this view, is "ontologically subjective in the sense that [its] very existence depends on our propositional attitudes towards [it]."[240] But it lacks "ontological depth" that can be traced in blood.[241]

2. The Costs of Classificatory Inconsistency

Variation in how race is treated, and the doctrinal dominance of race as a natural kind, is hard to explain in a principled way. It creates both analytic and normative puzzles. Spelling those out here, my aim is to demonstrate the value of recognizing the role of natural and social kinds as the analytic foundations of constitutionally salient

234 Saint Francis Coll. v. Al-Khazraji, 481 U.S. 604, 610 n.4 (1987).

235 Ozawa v. United States, 260 U.S. 178, 192 (1922) (emphasis omitted) (quoting An Act to Establish an Uniform Rule of Naturalization, ch. 3, 1 Stat. 103 (1790)).

236 *Id.* at 197 (emphasis added). In *United States v. Thind*, Justice Sutherland defined white persons as "the descendants of these [Europeans], and other immigrants of like origin, who constituted the white population of the country." United States v. Thind, 261 U.S. 204, 213–14 (1923). The opinions make no attempt to integrate these different conceptions of race.

237 *Thind*, 261 U.S. at 209.

238 *Id.* at 214.

239 Trina Jones & Jessica L. Roberts, *Genetic Race? DNA Ancestry Tests, Racial Identity, and the Law*, 120 COLUM. L. REV. 1929, 1965, 1964–65 (2020) (also noting the tension).

240 Khalidi, *supra* note 154, at 98–99.

241 MILLS, *supra* note 231, at 42.

classifications. That is, I aim to draw out analytic difficulties rather than offer a theory of equal protection.

First, the Court employs race as a *natural kind* in cases where a challenged practice uses race as a *social kind* without seeing the salience of the gap between the two concepts. Once again, *SFFA* provides a useful example. Recall that the Court in that case characterized race as a natural kind, and flagged dignitary concerns when the state divides individuals by "ancestry," a biological kind over which they have no meaningful control. But neither Harvard nor UNC solicited or possessed information on ancestry.[242] Both universities instead elicited applicants' self-defined racial identities[243]—i.e., they asked about race as a social kind that applicants *chose* to embrace. Indeed, Harvard in particular had been earlier "prepared to accept the most tenuous act of self-identification as proof positive of racial status."[244] Race, so far as affirmative action goes, is a self-ascribed social kind.

Tellingly, the *SFFA* majority was silent about *whence* the universities glean race information. In contrast, the district court in the *Harvard* case explained that the Common Application or Universal College Application includes a question on race.[245] Individual applicants assign themselves race. There is empirical reason to think they don't use "ancestry" to do so. As Camille Gear Rich has explained, Americans "tend to see racial identity primarily as a result of individual racial self-identification decisions," rather than a matter of indefeasible ascriptions.[246] She usefully labels this "elective race."[247] Empirical studies reveal large gaps between the choice of one's elective minority race and others' perceptions of one's race.[248] As many parents with

242 *See* Students for Fair Admissions, Inc. v. President & Fellows of Harvard. Coll., 143 S. Ct. 2141, 2193 (2023) (Thomas, J., concurring).

243 *See id.* at 2254 (Sotomayor, J., dissenting).

244 Camille Gear Rich, *Affirmative Action in the Era of Elective Race: Racial Commodification and the Promise of the New Functionalism*, 102 GEO. L.J. 179, 183 (2013) (describing classification of Professor Elizabeth Warren by the law school for faculty demographic reporting).

245 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 397 F. Supp. 3d 126, 136 (D. Mass. 2019).

246 Camille Gear Rich, *Elective Race: Recognizing Race Discrimination in the Era of Racial Self-Identification*, 102 GEO. L.J. 1501, 1512 (2014).

247 *Id.*; *see also* AMY GUTMANN, IDENTITY IN DEMOCRACY 24 (2003) (distinguishing ascriptive and elective identities).

248 For examples, see Andrew M. Penner & Aliya Saperstein, *Disentangling the Effects of Racial Self-Identification and Classification by Others: The Case of Arrest*, 52 DEMOGRAPHY 1017, 1018–21 (2015) (finding such gaps in study of arrestees), and Aliya Saperstein, *Double-Checking the Race Box: Examining Inconsistency Between Survey Measures of Observed and Self-Reported Race*, 85 SOC. FORCES 57, 58 (2006) (demonstrating that self-reported race and third-party classification often yield dissimilar results despite government policies that treat these methods as producing similar findings). *See also* Michelle Fine & Cheryl Bowers, *Racial Self-Identification: The Effects of Social History and Gender*, 14 J. APPLIED SOC. PSYCH. 136, 137 (1984).

children from interracial marriages can attest, children can and do make their own choices about race. Many "dissent from traditional cultural norms," and instead craft their own identity and path.[249]

These complex dynamics cast doubt on the "dignity"-based account of anticlassification in *SFFA*.[250] If the Constitution prohibits racial classification on the theory that it "tells each student he or she is to be defined by [the natural kind] race,"[251] then it should have not had an objection to regulatory or allocative schemes tracking the social kind of elective race—like Harvard's. At a minimum, it might be thought that an explanation was owed for treating elective race (a social kind) as if it was a biological fact (a natural kind).

Second, the application of the "race" classification in edge cases will often depend on whether race is understood as a natural or a social kind. Not only does the Court lack a principled way of electing between those two understandings, but the Justices are likely to make that choice without acknowledging or explaining what they are doing. The result is that the scope of anticlassification turns on arbitrary, or at least unreasoned, grounds.

The 2023 case of *Haaland v. Brackeen*, for example, framed the question whether provisions in the Indian Child Welfare Act (ICWA) allocating adoption priorities by native tribal status violated the Equal Protection Clause.[252] Petitioners challenged ICWA as impermissible "racial discrimination."[253] The same question had been reserved some nineteen years earlier.[254] In 2013, a majority of the Court identified "equal protection concerns" in the possibility that "a biological Indian father . . . could play his ICWA trump card at the eleventh hour to override the mother's decision and the child's best interests."[255] The district court in *Haaland*, echoing that logic, found ICWA invalid on the ground that it allocated children to families depending on whether they were "related to a tribal ancestor by blood."[256] On its view, Congress's decision to base "tribal membership" on "ancestry, rather than

249 Madhavi Sunder, *Cultural Dissent*, 54 STAN. L. REV. 495, 498 (2001).

250 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2170 (2023).

251 Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 789 (2007) (Kennedy, J., concurring in part and concurring in judgment).

252 Haaland v. Brackeen, 143 S. Ct. 1609, 1638 (2023).

253 *Id.* (explaining why the Court dismissed the equality challenge on Article III standing grounds).

254 United States v. Lara, 541 U.S. 193, 209 (2004) (not addressing the equal protection argument because it was "simply beside the point"). An earlier case had authorized classification by Indian status. Morton v. Mancari, 417 U.S. 535, 555 (1974).

255 Adoptive Couple v. Baby Girl, 570 U.S. 637, 656 (2013).

256 Brackeen v. Zinke, 338 F. Supp. 3d 514, 533, 536 (N.D. Tex. 2018).

actual tribal affiliation" meant that ICWA "use[d] ancestry as a proxy for race."[257]

The relationship between equal protection and statutes such as ICWA, which classify based on tribal status, turns on whether race is understood as a natural kind or a social kind. If race is a natural kind—defined by the biological fact of ancestry—then the challenge to ICWA gains force. ICWA contains this sort of classification. If race is a social kind, however, a more complex question arises. Recall that social kinds come in different flavors. Race is either a response-independent or a response-dependent social kind. That is, it may be "partly dependent on specific attitudes" (such as kinds like "war").[258] Or its "existence and that of [its] instances" may both be "dependent on attitudes that human beings have towards them" (such as "prime minister" or "pope").[259] Deciding between these versions of social kind, we must ask: Is it enough to have attitudes to the category of "race," or must there also be attitudes to its specific instances? And if the latter, the determination whether "Native American" classifications are impermissible "racial" categories would further hinge upon whether the appropriate attitudes to that exemplar in fact existed.[260] To answer this question, a judge would also have to decide whose attitudes count—Native Americans? Nonnative citizens? The federal government? The tribes qua sovereign? But it is simply not clear which of these groups matters for legal purposes.[261]

Again, my aim is not to answer these questions definitively. It is rather to illustrate their complexity. Deciding whether "race" is a natural or a social kind (and, if the latter, what sort of social kind) is likely outcome determinative of the equal protection issue framed but not answered by *Haaland*. Yet judges who have grappled with the question do not recognize that they need to define race as a natural or a social kind. They do not realize that social kinds come in different varieties, and they have not thought about whose responses to a potential racial token matter. In short, they have ignored the normative and analytic

257 *Id.* at 533–34; Gregory Ablavsky, *"With the Indian Tribes": Race, Citizenship, and Original Constitutional Meanings*, 70 STAN. L. REV. 1025, 1031 n.18 (2018) (collecting cases addressing the same question); *see also* Matthew L.M. Fletcher & Wenona T. Singel, *Lawyering the Indian Child Welfare Act*, 120 MICH. L. REV. 1755, 1784 (2022) (noting a "gloss on Indian law introduced by . . . Judge Kozinski, who theorized that all classifications based on Indian status not rooted in tribal membership are unconstitutional").

258 Khalidi, *supra* note 154, at 100.

259 *Id.* at 101, 103.

260 In the late 1700s, "Anglo-Americans . . . defined 'Indians' . . . [s]ometimes . . . as nonwhites, 'red' people defined by racial difference. Other times, especially in diplomacy and law, they classified Indians as noncitizens marked by their allegiance to another sovereign." Ablavsky, *supra* note 257, at 1033–34.

261 *See* Passinsky, *supra* note 168, at 436.

difficulties that flow from anticlassification as applied to equal protection.

That said, at least some empirical evidence points toward race being a *response-dependent* social kind. Under that description, an instance is not within the category unless people have the correct attitudes toward the instance as well as the category. What counts as a non-white race has never been fixed. Migrant groups such as Italian-Americans, for example, campaigned through the 1940s and 1950s against being viewed as "culturally, and perhaps racially, undesirable."[262] Irish Americans, similarly, were once malignly caricatured as "'low-browed,' 'brutish,' and even 'simian.'"[263] Neither "Irish" nor "Italian" is understood as a racial classification today: they have sailed across the racial meridian. Yet we still talk of "Irish" or "Italian" Americans, and we also obviously still use the general type of "race." The possibility that exemplars can migrate across the boundary around the type of "race" implies that both the category and its instances are necessarily "dependent on attitudes that human beings have towards them."[264]

Occasionally, such fluidity elicits judicial derision. In *SFFA*, for example, Chief Justice Roberts criticized the "Hispanic" category as "arbitrary or undefined."[265] Reflection on the dynamics of identity-related social kinds in the United States, however, suggests that this posture is unwarranted. The use of "Hispanic" and "Latino" labels is subject to what philosophers of language call "semantic indecision."[266] This occurs when a group of language users "ha[s] neither decided on one of [several] plausible disambiguations of [a term] nor accepted that it is ambiguous."[267]

For example, the class "Latino" is treated in the U.S. Census as "an ethnic group whose members can be of any race."[268] Respondents

262 Danielle Battisti, *The American Committee on Italian Migration, Anti-Communism, and Immigration Reform*, 31 J. AM. ETHNIC HIST. 11, 19 (2012); *see also* Mae M. Ngai, *The Architecture of Race in American Immigration Law: A Reexamination of the Immigration Act of 1924*, 86 J. AM. HIST. 67, 69 (1999) (explaining how, in the early twentieth century, "race and nationality—concepts that had been loosely conflated since the nineteenth century—[became] disaggregated and realigned in new and uneven ways").

263 MATTHEW FRYE JACOBSON, WHITENESS OF A DIFFERENT COLOR: EUROPEAN IMMIGRANTS AND THE ALCHEMY OF RACE 48 (1998); *see also* Ngai, *supra* note 262, at 67–69 (also discussing the trajectory of Irish Americans).

264 Khalidi, *supra* note 154, at 101.

265 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2167 (2023).

266 Muhammad Ali Khalidi & Liam Murphy, *Disagreement About the Kind* Law, 12 Juris. 1, 7 (2021).

267 *Id.*

268 Christopher Lewis, *Latinos and the Principles of Racial Demography*, 16 DU BOIS REV. 63, 64 (2019).

to the census in practice diverge as to which race they select.[269] The proportion of responses picking any given race fluctuates sharply over time.[270] Criticism of the "Hispanic" or "Latino" label as "arbitrary or undefined," in other words, suggests that the Court has not understood how specific racial tokens dynamically emerge and recede in American society. But the consequential question of how to draw bounds on the constitutionally salient category of "race" is likely better answered by careful attention to the dynamics of social-kind formation, with particular attention to how semantic indecision is either dissolved or preserved.

Third, and relatedly, it is widely agreed among social scientists that race is not a natural kind, but a set of categories that emerge through contentious and contestable social and political practices.[271] The category's pernicious effects turn upon the complex way in which it is nested in larger social and economic dynamics that lock a group into a fixed, subordinate site within a social hierarchy.[272] Race has historically been the "centerpiece of a hierarchical system that produces differences" not because it can be defined scientifically as a natural kind—it cannot—but precisely because its vagueness can be leveraged

269 *Id.*

270 For evidence of shifts in terms of "Latino" identity, see Atiya Kai Stokes-Brown, *America's Shifting Color Line? Reexamining Determinants of Latino Racial Self-Identification*, 93 SOC. SCI. Q. 309, 310–12 (2012).

271 For exemplary statements of this position, see MICHAEL OMI & HOWARD WINANT, RACIAL FORMATION IN THE UNITED STATES 110 (3d ed. 2015) (defining race as "a concept that signifies and symbolizes social conflicts and interests by referring to different types of human bodies"); PAUL GILROY, THERE AIN'T NO BLACK IN THE UNION JACK: THE CULTURAL POLITICS OF RACE AND NATION 35 (Routledge Classics 2002) (1987) ("[Racial formation] refers both to the transformation of phenotypical variation into concrete systems of differentiation based on 'race' and colour and to the appeals to spurious biological theory which have been a feature of the history of 'races.'"). For a leading account in the legal scholarship, see Ian F. Haney López, *The Social Construction of Race: Some Observations on Illusion, Fabrication, and Choice*, 29 HARV. C.R.-C.L. L. REV. 1, 7 (1994) (defining race as "a vast group of people loosely bound together by historically contingent, socially significant elements of their morphology and/or ancestry" and "a *sui generis* social phenomenon in which contested systems of meaning serve as the connections between physical features, races, and personal characteristics").

272 Sally Haslanger, *Tracing the Sociopolitical Reality of Race*, *in* WHAT IS RACE? FOUR PHILOSOPHICAL VIEWS 4, 24–31 (Joshua Glasgow et al. eds., 2019) (discussing theories of race that link it as a necessary or as a contingent matter to the organization of social hierarchies). In other, joint work, Haslanger has endorsed a definition of race in terms of whether individuals are "positioned as subordinate or privileged along some dimension—economic, political, legal, social, etc.— . . . and the group is 'marked' as a target for this treatment by observed or imagined bodily features." Sally Haslanger & Jennifer Saul, *Philosophical Analysis and Social Kinds*, 80 PROC. ARISTOTELIAN SOC'Y, SUPPLEMENTARY VOLUMES 89, 93 (2006).

for purposes of maintaining stratification under shifting social and economic circumstances.[273]

Treating race as a natural kind for the purposes of law and then prohibiting its statutory use will not prevent a racial hierarchy from persisting. Quite the contrary. As Justice Ketanji Brown Jackson's dissent in *SFFA* extensively explored, racial reductionism is more likely to insulate social and economic grounds of stratification from reform.[274] To offer just one striking example, in 2019 the modal Black household had twelve cents for every dollar in wealth held by the typical white household.[275] A racial wealth gap of this kind has existed since household wealth data has been gathered.[276] One cause of this persisting gap is the tendency for increased investment in education to yield lower economic payoffs for Blacks than for whites.[277] It is hard to see how these dynamics—with obvious harmful health, welfare, and dignity effects—can be identified, let alone addressed, without using race as a lens for analysis and action.

Instead, a legal and social imaginary world in which the idea of "racism" is assimilated to the notion of anticlassification is one in which malign biological explanations for racial differences in outcomes are likely to flourish because their proponents can point to ambient social reality as "confirmatory" evidence.[278] This is the world to which the Roberts Court is leading us.

---

273 STUART HALL, THE FATEFUL TRIANGLE: RACE, ETHNICITY, NATION 33, 67–68 (Kobena Mercer ed., 2017).

274 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2263–71 (2023) (Jackson, J., dissenting) ("Gulf-sized race-based gaps exist with respect to the health, wealth, and well-being of American citizens." *Id.* at 2263.); *id.* at 2271 ("To demand that colleges ignore race in today's admissions practices—and thus disregard the fact that racial disparities may have mattered for where some applicants find themselves today—is not only an affront to the dignity of those students for whom race matters.").

275 Grieve Chelwa, Darrick Hamilton & Avi Green, *Identity Group Stratification, Political Economy & Inclusive Economic Rights*, 152 DÆDALUS 154, 155 (2023); *see also* Neil Bhutta, Andrew C. Chang, Lisa J. Dettling & Joanne W. Hsu, *Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances*, FEDS NOTES (Sept. 28, 2020), https://www.federalreserve.gov/econres/notes/feds-notes/disparities-in-wealth-by-race-and-ethnicity-in-the-2019-survey-of-consumer-finances-20200928.html [https://perma.cc/NQ4J-2M4W].

276 Chelwa et al., *supra* note 275, at 156; Darrick Hamilton & Regine O. Jackson, *An Absent Asset-Based Black American Middle Class: The Iterative Role of Hard Work, Education, and Intergenerational Poverty*, *in* THE MIDDLE CLASS IN WORLD SOCIETY: NEGOTIATIONS, DIVERSITIES AND LIVED EXPERIENCES 201, 201–05 (Christian Suter et al. eds., 2020).

277 Chelwa et al., *supra* note 275, at 156; Hamilton & Jackson, *supra* note 276, at 210–12 (collecting studies).

278 As indeed they are doing. For a piercing account of how biological ideas of racial hierarchy are starting to re-emerge, see Quinn Slobodian, *The Rise of the New Tech Right: How the Cult of IQ Became a Toxic Ideology in Silicon Valley and Beyond*, NEW STATESMAN (Sept. 13, 2023), https://www.newstatesman.com/ideas/2023/09/rise-new-tech-right-iq-cognitive-elite [https://perma.cc/AP29-W7CE].

* * *

This Part has explored how categories are fashioned and deployed in anticlassification jurisprudence. Because almost none of the relevant categories are to be found in the law itself, judges must derive them from extralegal sources. They must use either natural or social kinds. And this brings with it a host of complexities.

But once the role of natural and social kinds in anticlassification doctrine is foregrounded, it quickly becomes apparent that the threshold act of forging a classification is not mechanical. It cannot be resolved by a naked appeal to the "commonsense."[279] To the contrary, it is often unclear whether the Court has in mind a natural or a social kind (and if the latter, whether it's homed in upon a specific species of social kind). These choices are currently discretely folded into the seemingly singular and unitary act of creating a constitutionally salient category such as race. But they are irreducible, unavoidable, and eminently contestable on normative grounds.

III. FINDING CONSTITUTIONALLY SALIENT CATEGORIES IN LEGAL TEXTS

It would be a relief if the act of delineating a category fixed the set of its instances and so dissolved the difficulties of anticlassification doctrine. But even once a category has been fixed—resolving the problematics identified in Part II—there remains the second doctrinal step of anticlassification analysis.[280] A judge must determine whether an instance of the prohibited category is present in the enactment's text. Once again, this task of isolating the presence of an impermissible exemplar of a forbidden type is beset by unappreciated analytic and normative difficulties. These cannot be understood without leaning on technical distinctions drawn from the philosophy of language. My aim here again is to tease out those problematics and show that their challenge has not been properly appreciated.

This Part takes as given that we have in hand a constitutionally salient category, and that we have plotted its outer bounds. It focuses on the subsequent task of applying that benchmark to legal text. To motivate the inquiry, I start by mapping how courts struggle when asked to say whether an instance of a constitutionally salient category is present on the face of a legal enactment. In order to clarify what seems like confusion in the doctrine's application, I draw upon two additional technical distinctions in the philosophy of language: the semantic/communicative content distinction and the sense/referent

279 Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015).

280 *See supra* subsection I.A.2.

distinction. These technical terms are useful because they allow us to frame and articulate what superficially seems puzzling or incoherent in the law.

With these distinctions in hand, I then flesh out the central claim of this Part: the way in which an anticlassification rule works in practice, I argue, depends on *how* it is normatively justified. Different normative groundings lead to different choices of semantic or communicative meaning (sense or referent). Indeed, it is not possible to understand the seemingly haphazard practice of judges without accounting for how their different normative commitments are leading to different choices between such analytic frames—and so divergent outcomes.

In developing this point, I draw on the distinctions between the three justifications for anticlassification sketched in Section I.C, and in particular the possibility of either intrinsic or extrinsic justifications for anticlassification.[281] Intrinsic justifications treat classifications as directly harmful because of their putative dignitary or balkanizing effects. Extrinsic justifications, in contrast, take classifications as signals of problematic intentions.

A. The Opacity of Classificatory Practice

It ought to be straightforward: take the list of exemplars under a constitutionally salient category, and then examine a legal enactment to determine whether one is present. The following examples, partly anticipated in the Introduction, suggest that matters are not so simple:

- A university's admissions rubric asks about what a candidate is proud of, or obstacles they have overcome. Minority candidates frequently take these as prompts to discuss their race, and admissions staff necessarily reflect and take account of the role of a candidate's race in their life course.[282]
- One university responds (hypothetically!) to *SFFA* by creating an admissions preference for the descendants of slaves; another responds by enlarging its legacy preferences (which, it knows, benefits almost exclusively white parents).[283]
- A police sergeant directs officers going on patrol to stop "the right people," and later concedes that she and her officers

281 I put to one side momentarily administrability justifications. As I show through Sections III.B and C, these are substantially undermined by an excavation of hidden analytic and normative choices of anticlassification's application to enactment text.

282 Transcript of Oral Argument, *supra* note 2, at 9–11.

283 *Id.* at 48, 103 (discussing legacy preferences).

alike understood this to mean "young men of color in their late teens, early 20s."[284]

- A state allows both in-state and out-of-state wineries to make direct sales to consumers, so long as they have a "branch office" in state.[285] State regulations make this very hard to do for out-of-state wineries, but not those in state.[286]
- A state bans protests on sidewalks, but only near facilities that provide abortion—reasoning that these facilities among all other possible protest sites attract the most vituperative public demonstrations.[287]

In each of these examples, there is a plausible argument that an anticlassification rule is not triggered because there is no instance of the forbidden category in the dispositive legal directive. But in each of these examples, there is also a plausible argument that an anticlassification rule has been violated, triggering strict scrutiny. The first two examples are hypotheticals drawn (with some tweaking) from the *SFFA* oral argument. The third and fourth are drawn from cases in which the state practice was invalidated on anticlassification grounds. In the fifth case, the measure was invalidated—but not as content discrimination. In short, these illustrate that the second step in anticlassification analysis is a potentially complex one.

To help clarify the difficulty of these examples, it is helpful to start once again with some terminology from the philosophy of language. I first set out two technical ways of framing the core problematic here: the semantic/pragmatic distinction and the sense/referent distinction. I then use these technical terms to explore how anticlassification's justifications shape its application to real-world cases, such as those listed above. I show that resolution of these questions turns on the underlying justification for anticlassification mapped in Section I.C, focusing on intrinsic and extrinsic justifications.[288]

B. Technical Frames for Reading Legal Texts

Two technical distinctions drawn from the philosophy of language are useful to surface here as tools for the doctrinal analysis to follow.

---

284 Floyd v. City of New York, 959 F. Supp. 2d 540, 603–04 (S.D.N.Y. 2013) (emphasis omitted) (first quoting Transcript of Proceedings re: Trial Held on 4/10/2013, at 3034, *Floyd*, 959 F. Supp. 2d 540 (No. 08-cv-1034), ECF No. 311; and then quoting *id.* at 3029).

285 *See* Granholm v. Heald, 544 U.S. 460, 474 (2005).

286 *Id.* at 474–75.

287 McCullen v. Coakley, 573 U.S. 464, 469–70, 482 (2014).

288 Philosophers divide volubly about the proper units for the analysis of language. I draw here on what seem to be reasonably uncontroversial understandings as scaffolding. I do not, however, aim to contribute to debates on the many hard questions in the philosophy of language.

First, there is a distinction between the "semantic" content of a text and its "communicative" content.[289] The former concerns "the context-independent meaning of words, phrases, and sentences."[290] The latter can be defined in several different ways. These include (a) what the speaker intended to communicate, (b) what an audience reasonably would have taken her to communicate, and (c) what the speaker reasonably expects the audience to understand.[291] Communicative content is thus ambiguous, but its potential specifications all involve different aspects of intersubjective human conversation.

There are many reasons why semantic content and communicative content (however disambiguated) can peel apart. Most importantly, communicative content can be "pragmatically enriched" insofar as a reasonable speaker supplements or modifies semantic meaning in light of certain features of the actual context in which a statement is made.[292] For example, when we meet a friend and they tell us, "I've not had breakfast," we usually understand them to mean, "I've not had breakfast [yet today]," and not, "I've never had breakfast in my life." Pragmatic enrichment can draw upon many sources. These include "the semantic content of the sentence uttered, the communicative intentions of the speaker, the shared presuppositions of speaker-hearers, and obvious features of the context of utterance."[293] In particular, pragmatic enrichment is said to operate through the "maxims" identified by the philosopher Paul Grice, such as the

289 Mark Greenberg, *Legislation as Communication? Legal Interpretation and the Study of Linguistic Communication*, *in* PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW 217, 217 (Andrei Marmor & Scott Soames eds., 2011).

290 Richard H. Fallon Jr., *The Meaning of Legal "Meaning" and Its Implications for Theories of Legal Interpretation*, 82 U. CHI. L. REV. 1235, 1246 (2015); *accord* Lawrence B. Solum, *The Interpretation-Construction Distinction*, 27 CONST. COMMENT. 95, 98 (2010) ("The semantic content of a legal text is simply the linguistic meaning of the text.").

291 Greenberg, *supra* note 289, at 220–21. As Solum rightly warns, "communicative content" and "legal content" are distinct: we can, for example, sign a contract intending to communicate certain obligations—but the contract can violate a legal norm in a way that renders it invalid. Lawrence B. Solum, *Communicative Content and Legal Content*, 89 NOTRE DAME L. REV. 479, 486 (2013). I do not rely on the concept of "legal meaning" here.

292 1 SCOTT SOAMES, *The Gap Between Meaning and Assertion: Why What We Literally Say Often Differs from What Our Words Literally Mean*, *in* PHILOSOPHICAL ESSAYS: NATURAL LANGUAGE; WHAT IT MEANS AND HOW WE USE IT 278, 282 (2009); *accord* Solum, *supra* note 291, at 488 ("The full communicative content of a legal writing is a product of the semantic content (the meaning of the words and phrases as combined by the rules of syntax and grammar) and the additional content provided by the available context of legal utterance.").

293 Andrei Marmor & Scott Soames, *Introduction*, *in* PHILOSOPHICAL FOUNDATIONS OF LANGUAGE IN THE LAW, *supra* note 289, at 1, 8; *see also* Carston, *supra* note 18, at 12–13 (describing how semantic meaning can be either narrowed or expanded); Robyn Carston, *Linguistic Communication and the Semantics/Pragmatics Distinction*, 165 SYNTHESE 321, 323 (2008).

directive to "[m]ake [a] conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged."[294]

The importance of pragmatic enrichment for actual communication is hard to overstate. On one powerfully argued view, it is "an inherent property of language systems that the sentences they generate do not (cannot) fully determine the meaning expressed or communicated by a speaker."[295] One does not have to accept a strong "linguistic underdeterminacy" thesis[296] to recognize that pragmatic enrichment is pervasive in everyday communication. It also plays a central role when searching legislative text for exemplars of constitutionally salient categories.

Second, a different line of work in the philosophy of language, initiated by Gottlob Frege, draws attention to the distinction between "sense" and "reference." In Frege's account, "a sign (name, combination of words, letter)" is connected to "that to which the sign refers, which may be called the referent of the sign," and also "the sense of the sign, wherein the mode of presentation is contained."[297] The referent of a sign is the specific, concrete, real-world thing(s) it picks out, or its "extension."[298] Frege's notion of "sense" is, in contrast, elusive and extensively criticized.[299] It is enough for present purposes here to say that he uses the term "sense" to capture the "intuitive difference in cognitive significance" between pairs of words.[300] A simple example can clarify: The signs "Mark Twain" and "Samuel Clemons" have the same referent, but the statement "Mark Twain is Samuel Clemens" is not empty of content.[301] In consequence, "Mark Twain" and "Samuel Clemens" have different senses. Further, "sense determines reference," but signs with different senses can have the same reference (as in the Twain example).[302] Senses, on Frege's view, are also "objective,

294 PAUL GRICE, STUDIES IN THE WAY OF WORDS 26 (1989); *see also* Carston, *supra* note 18, at 13–16 (documenting the role of Gricean maxims in pragmatic enrichment).

295 Carston, *supra* note 18, at 13.

296 *Id.*

297 Frege, *Sense and Reference*, *supra* note 17, at 210 (emphasis omitted).

298 David J. Chalmers, *On Sense and Intension*, 16 PHIL. PERSPS. 135, 135–36 (2002).

299 *Id.* at 135 (noting criticisms); *see also* Tyler Burge, *Sinning Against Frege*, 88 PHIL. REV. 398, 398 (1979).

300 Chalmers, *supra* note 298, at 138.

301 Another way Frege puts this is to say that sentences have the same "meaning," but express different "thoughts." Richard G. Heck & Robert May, *Frege's Contribution to Philosophy of Language*, *in* THE OXFORD HANDBOOK OF PHILOSOPHY OF LANGUAGE 3, 20 (Ernie Lepore & Barry C. Smith eds., 2008).

302 Speaks, *supra* note 21. Hilary Putnam has challenged Frege's view that meaning is "public," but "grasped" by specific individuals, arguing that the extension of a term is "not a function of the psychological state of the speaker," but has an indexical quality. Hilary Putnam, *Meaning and Reference*, 70 J. PHIL. 699, 699, 702, 710 (1973) (emphasis omitted).

in that more than one person can express thoughts with a given sense."[303]

Both of these technical distinctions help clarify the five examples set forth at the beginning of this Part.[304] In each of the examples, it is plausible to say that the *semantic meaning* of the law does not include an exemplar of a problematic category. But it is simultaneously possible to say that at least one version of the *communicative content* of a law, appropriately pragmatically enriched, includes an impermissible classification. Take the example of stopping "the right people."[305] This phrase is an "indexical," or "a linguistic expression whose reference can shift from context to context."[306] Only when pragmatically enriched using the context of the sergeant's morning peroration does it come to have communicative content that includes a racial classification.

It is also possible to say that the *referent* of the law in the five scenarios described above is identical to the referent of a law whose *sense* does probably violate the anticlassification command: shifting senses while preserving referents might be seen as intentionally end-running anticlassification in an especially blatant way. For example, the state law requiring in-state branch offices has as a referent only out-of-state wineries. When admissions staff ask about applicants' background in the first example, they pick out a referent that will (often and predictably) demand consideration of race.

The technical distinctions between semantic and communicative content, and between sense and reference, therefore, are useful in parsing how anticlassification operates in practice. Each distinction helps frame and clarify conflict over whether instances of a constitutionally salient category appear in an enactment's text.

In the balance of this Part, I use these technical distinctions to analyze how the second step of anticlassification works in practice. The discussion is organized around the different normative grounds of anticlassification. My aim is to show how the normative justification for an anticlassification regime powerfully shapes when a category is perceived in a legal text. I hence return to the different justifications for anticlassification rules identified in Section I.C, in particular intrinsic reasons (dignity or balkanization effects) and extrinsic reasons (related to impermissible intent). The application of anticlassification rules to legal texts, I argue here, turns on which of these justifications is accepted. In this Part, I focus on intrinsic and extrinsic justifications

303 Speaks, *supra* note 21.

304 *See supra* text accompanying notes 282–87.

305 Floyd v. City of New York, 959 F. Supp. 2d 540, 603 (S.D.N.Y. 2013).

306 David Braun, *Indexicals*, STAN. ENCYC. PHIL. (Jan. 16, 2015), https://plato.stanford.edu/entries/indexicals/ [https://perma.cc/U2BU-RL9N].

because they lead to different choices between semantic and pragmatic meaning in particular. I return to administrability justifications in the Conclusion.

C. Intrinsic Justifications for Anticlassification Rules

Anticlassification rules have been defended on the ground that they advance dignitary or antibalkanization goals.[307] These can be labeled "intrinsic" justifications because they follow directly as necessary and proximate effects of the text itself. Intrinsic justifications point logically toward the communicative content of an enactment, not its semantic content. The Court's doctrine, however, does not offer a clear or consistent way of disambiguating what sort of communicative content counts. While the resulting doctrine maintains an air of easy administrability, this aura of clarity is achieved by suppressing or denying hard underlying analytic choices.

1. Dignity and Balkanization Harms from Communicative Content

Intrinsic justifications for anticlassification rules hinge upon the effect of language on audiences. Racial classifications, for instance, are said to directly infringe a "'personal right[]' to be treated with equal dignity and respect"[308] by taking "individuals as the product of their race" alone.[309] The doctrine is clear that this problematic effect flows from a statute's text.[310] Under the dormant Commerce Clause, state laws targeting out-of-state market participants are similarly condemned because they send a balkanizing message. Measures are facially discriminatory, that is, when they "invite a multiplication of preferential trade areas"[311] or could foment "low-level trade war[s]."[312] Intrinsic justifications hence turn on the effect of an enactment upon others. They turn, that is, on the words' communicative content.[313]

---

307 *See supra* text accompanying notes 129–37.

308 City of Richmond v. J.A. Croson Co., 488 U.S. 469, 493 (1989) (plurality opinion) (quoting Shelley v. Kraemer, 334 U.S. 1, 22 (1948)).

309 Miller v. Johnson, 515 U.S. 900, 912 (1995).

310 *See* Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 240 (1995) (Thomas, J., concurring in part and concurring in judgment) ("[Racial] classifications ultimately have a destructive impact on the individual and our society.").

311 Dean Milk Co. v. City of Madison, 340 U.S. 349, 356 (1951).

312 Granholm v. Heald, 544 U.S. 460, 473 (2005).

313 Elizabeth Anderson and Richard Pildes distinguish between expressive and communicative effects. *See* Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1565 (2000) ("To express a state of mind is, among other things, to manifest it in action. To communicate a state of mind is to act with the intention of inducing others to recognize that state of mind by recognizing that very communicative intention. Communicative acts are only a small subset of all expressive

Intrinsic justifications hence imply that the semantic content of a rule should not determine whether it is facially discriminatory. This is likely why the hypothetical admissions preference in favor of slaves' descendants raised in the *SFFA* oral argument seemed to some Justices a potential trigger for strict scrutiny.[314] The phrase's referent is not a present set of persons who share the same racial identity: people with different racial identities can trace their lineage back to formerly enslaved persons. Its semantic meaning is not identical to "Black." But the Justices seem to think that its communicative content, pragmatically enriched, would be to pick out only "Black" applicants.[315]

If semantic content is not dispositive for anticlassification rules, the doctrine's administrability suffers. Legal analysis necessarily becomes far more complex if it is not the semantic meaning of an enactment that determines whether it contains an instance of a constitutionally salient category. If the Court needs to go beyond the most readily available point of semantic reference—to decide which kind of communicative content matters—it is plunging into murkier, more contested waters.

This understanding of intrinsically justified anticlassification rules also brings to light a number of practical and analytic questions. Communicative content is ambiguous. There are many ways in which the communicative content of the law can be understood, including what the enactor meant to communicate, what a reasonable audience would have understood, and what a reasonable enactor would have expected an audience to understand.[316] As Mark Greenberg has rightly observed, it is hard "to adjudicate between [these] different notions of communicative content" without "some deeper understanding of the underlying rationale or purpose that lies behind the appeal to the notion of communication."[317]

None of the decisions leaning upon intrinsic justifications for anticlassification, however, identifies *which* kind of communicative

---

acts."). They understand expression in terms of how "an action or a statement (or any other vehicle of expression) manifests a state of mind." *Id.* at 1506. And they interpret equal protection and dormant Commerce Clause doctrine in terms of such "expressive" effects. *Id.* at 1533–45, 1551–56. My use of the term "communicative content" is closer to their idea of "expression" and their idea of "communication." I think my terms are helpful for understanding analytic choices that an expressive-effects lens does not surface.

314 *See* Transcript of Oral Argument, *supra* note 2, at 16.

315 Put differently, "Black" and "slaves' descendants" have different senses but the same referent—at least on the view of some Justices.

316 Greenberg, *supra* note 289, at 220; *see also* Lawrence B. Solum, *Intellectual History as Constitutional Theory*, 101 VA. L. REV. 1111, 1117 (2015) (recognizing this ambiguity).

317 Greenberg, *supra* note 289, at 249.

content is at issue.[318] The Court has offered no such general account in the cases applying an intrinsically justified anticlassification rule. There are, in consequence, no guidelines to stabilize or channel the exercise of first-order normative judgments in court. Untethered by semantic meaning, the Justices seem free to impute whatever communicative content to an enactment they see fit to find. This discretion, moreover, is layered onto the threshold choice whether dignity demands an anticlassification rule (for race) or an anti-anticlassification rule (for religion). Put these choices together, and it is apparent that the Court has a large yet largely invisible zone in which to exercise untrammeled normative discretion.

2. The Challenge of Applying Anticlassification Rules: The Dormant Commerce Clause

It is useful to offer a specific example. In the dormant Commerce Clause context, the Court has found discrimination in a facially neutral Massachusetts tax coupled to a subsidy scheme solely for in-state market participants.[319] But in another case, it has authorized "subsid[ies] for the in-state members of the industry, funded from the State's general revenues."[320] The latter rule, indeed, covers much of what a state does in maintaining infrastructure, funding education, and supplying basic social services. Across these domains, the state often offers subsidies to only in-state market participants. This distinction between facial discrimination and the ordinary diet of state regulation turns upon the Court's sense of what will *seem to be* a "tariff" to other state governments—i.e., some notion of communicative content. But without any empirical basis for making such a judgment about such cross-border perceptions, the law must turn entirely on the Court's (unstated) priors. More bluntly stated, we simply have no basis for concluding that one or another kind of state law will, in fact, precipitate a "low-level trade war."[321] Perhaps the Court's intuitions are credible. But since

318 In other work, Greenberg has stressed the absence of any stable judicial framework for sorting between semantic content and different varieties of communicative content, and the want of a framework for explaining why it is relevant or should trump other putative determinants. Mark Greenberg, *The Moral Impact Theory of Law*, 123 YALE L.J. 1288, 1327–28 (2014). My point here tracks that observation.

319 W. Lynn Creamery, Inc. v. Healy, 512 U.S. 186, 197 (1994) (describing this as a "tariff-like barrier[]").

320 *Id.* at 210–11 (Scalia, J., concurring in judgment) (first citing New Energy Co. of Ind. v. Limbach, 486 U.S. 269, 278 (1988); and then citing Hughes v. Alexandria Scrap Corp., 426 U.S. 794, 809–10 (1976)).

321 Granholm v. Heald, 544 U.S. 460, 473 (2005). The literature on trade wars, unsurprisingly, focuses almost exclusively on the global system of rules-based trade between nations, and not on the internal dynamics of federations. But even in this literature, recent

they are offered without explanation of grounding, it is hard to know why or why not.

The doctrine also contains internal tensions. In the dormant Commerce Clause context (as in the equal protection context), the intentions of enactors are said to be irrelevant.[322] Yet enactor intent seems to matter when it comes to finding impermissible categories in legal texts. Cases such as the challenge to the Massachusetts tax seem to turn upon some account of either what enactor intentions actually are or what they are reasonably taken to be. Put another way, the account of communicative content relevant in the cases seems to rest on identifying the assumptions speakers and listeners share, how they anticipate different terms will be used, and what the "widely understood applications and nonapplications" of a term are.[323] But if that is correct, the intentions of enacting legislators are outcome determinative—and certainly not irrelevant.

3. The Challenge of Applying Anticlassification Rules: Equal Protection

In the equal protection context, there is a related set of problems that arise in the process of nailing down communicative content. But the Court smooths over apparent differences in facial classifications' communicative content. It also lacks a clear account of when a proxy for such an impermissible classification should trigger judicial skepticism.

The first dynamic is this: the Court treats all racial classifications as if they had the same communicative content as "the First Regulation to the Reichs Citizenship Law"[324] and the "Population Registration Act

---

work points to the importance of structural asymmetries between nations' productive capacity and consumer markets as driving trade wars. MATTHEW C. KLEIN & MICHAEL PETTIS, TRADE WARS ARE CLASS WARS: HOW RISING INEQUALITY DISTORTS THE GLOBAL ECONOMY AND THREATENS INTERNATIONAL PEACE 3 (2020). By extension, variance in the productive capacity of American states may be more important than tariff-like policies in catalyzing trade wars.

322 *See, e.g.*, Grutter v. Bollinger, 539 U.S. 306, 326 (2003) ("[A]ll racial classifications imposed by government 'must be analyzed by a reviewing court under strict scrutiny.'" (quoting Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 227 (1995))), *abrogated by* Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141 (2023).

323 Fallon, *supra* note 290, at 1248.

324 Metro Broad., Inc. v. FCC, 497 U.S. 547, 633 n.1 (1990) (Kennedy, J., dissenting) (quoting Fullilove v. Klutznick, 448 U.S. 448, 534 n.5 (1980) (Stevens, J., dissenting) (citing Erste Verordnung zum Reichsbürgergesetz [First Regulation to the Reichs Citizenship Law], Nov. 14, 1935, REICHSGESETZBLATT, Teil I [RBG$_L$ I] at 1333, *translated in* 4 OFF. OF U.S. CHIEF OF COUNS. FOR PROSECUTION OF AXIS CRIMINALITY, NAZI CONSPIRACY AND AGGRESSION 8 (1946))), *overruled by Adarand Constructors*, 515 U.S. 200.

No. 30 of 1950, Statutes of the Republic of South Africa."[325] But it has never offered evidence to the effect that the communicative content of even indisputable racial classifications has the same message as these instruments of iniquity. To many in the Court's audience, these comparisons likely seem strained at best. At the least, the Justices are helping themselves to a degree of pragmatic enrichment that is hard to justify on empirical grounds, just as they did in talking of trade wars and protectionism in the dormant Commerce Clause context.

It is plausible to think that the communicative content the Court assigns to racial classifications is partly a function of the Court's own rhetoric. After all, Justices were among the first major participants in American public life in the 1970s and early 1980s to condemn affirmative action schemes because of the "injustice" they inflicted on "losers."[326] The communicative content attributed to all racial classifications may hence be a bootstrap product of the Justices' own ideological entrepreneurship. Today, Benjamin Eidelson has noted, it is the Court's own logic of equivalence that currently "tells people" that any racial classification entails "disrespect, when in fact it ordinarily does not."[327] In this way, the Court produces the very harm it purports to mitigate.

Second, and in addition to this smoothing-out problem, there is a mirror-image dynamic in cases where the state substitutes a clear instance of a facial classification with what might be called an intentional proxy. In the third example offered above in Section III.A, for example, the instruction to stop "the right people"[328] is (by concession at trial) a proxy for stopping "young men of color."[329] In the absence of a concession by a defendant official, however, it will often be unclear when a potential proxy has the same communicative content as an explicit racial classification. It is necessary to have a consistent way of determining the communicative content of putative "proxies."

---

325 *Id.* (citing Population Registration Act No. 30 of 1950 (S. Afr.)); Shaw v. Reno, 509 U.S. 630, 647 (1993) (stating that majority-minority voting districting "bears an uncomfortable resemblance to political apartheid").

326 Johnson v. Transp. Agency, 480 U.S. 616, 677 (1987) (Scalia, J., dissenting); *accord* United Steelworkers of Am. v. Weber, 443 U.S. 193, 254 (1979) (Rehnquist, J., dissenting). These statements were tendered in the context of a "public debate over affirmative action in the late 1970s and 1980s [in which] whites increasingly framed whiteness as a liability." Erinn Brooks, Kim Ebert & Tyler Flockhart, *Examining the Reach of Color Blindness: Ideological Flexibility, Frame Alignment, and Legitimacy Among Racially Conservative and Extremist Organizations*, 58 SOCIO. Q. 254, 256 (2017).

327 Eidelson, *supra* note 130, at 1660.

328 Floyd v. City of New York, 959 F. Supp. 2d 540, 604 n.284 (S.D.N.Y. 2013).

329 *Id.* at 604 (emphasis omitted) (quoting Transcript of Proceedings, *supra* note 284, at 3029).

But the Court's approach to this question has been erratic. In the *SFFA* oral argument, some Justices suggested that an admissions preference for "descendants of slaves" might trigger strict scrutiny.[330] Yet, in *Hernandez v. New York*, a plurality of the Court held that a prosecutor's dismissal of native Spanish-speaking jurors was race-neutral because it was based (perhaps implausibly) on their supposed inability to accept official court translations.[331] In the Fourth Amendment context, a suspect's presence in a high-crime area is relevant in determining reasonable suspicion for a stop.[332] But in practice, police citations to a "high-crime area" are not predicted by crime rates, but instead by suspects' race and a neighborhood's racial demographics.[333] That is, "high-crime" neighborhood is a de facto proxy for race. The examples can be multiplied: because of its shaping influence on American life, race can be switched out for many other traits, such as "marital history and employment status."[334] The Court, however, lacks a theory of communicative content that can make consistent sense of such cases by saying when a racial meaning is communicated and when it is not.

The specter of potential proxies seeds instability in the doctrine's operation over time. Thus, Eidelson has charted the possibility of a "recursive tendency toward inflation in respect's demands."[335] This happens when a word denotes disrespect toward a certain group; when people avoid that word and also close substitutes; when the substitutes come to denote disrespect too; and when the new substitutes *also* come to denote disrespect.[336] The logic of anticlassification has a parallel inflationary tendency driven by a hermeneutics of suspicion.

Here is how it could work: In the aftermath of *SFFA*, universities likely experience pressure to avoid not just explicit racial classifications but also close substitutes for such classifications. This is because as litigants and interest groups probe their admissions protocols, post-*SFFA* changes will be intensely and skeptically scrutinized on the belief that universities will likely strive to find other ways to diversify their classes. Opponents of affirmative action will be prone to argue expansively that post-*SFFA* changes are substitutes for impermissible classifications. In

330 *See* Transcript of Oral Argument, *supra* note 2, at 16.

331 Hernandez v. New York, 500 U.S. 352, 361 (1991) (plurality opinion).

332 Illinois v. Wardlow, 528 U.S. 119, 124 (2000).

333 Ben Grunwald & Jeffrey Fagan, *The End of Intuition-Based High-Crime Areas*, 107 CALIF. L. REV. 345, 351 (2019).

334 Jennifer L. Skeem & Christopher T. Lowenkamp, *Risk, Race, and Recidivism: Predictive Bias and Disparate Impact*, 54 CRIMINOLOGY 680, 681 (2016); *see also* Anna Offit, *The Character of Jury Exclusion*, 106 MINN. L. REV. 2173, 2209 (2022) (noting that "aspects of demeanor . . . can serve as proxies for race" at trial).

335 Benjamin Eidelson, *The Etiquette of Equality*, 51 PHIL. & PUB. AFFS. 97, 100 (2023) (emphasis omitted).

336 *See id.* at 120.

effect, they will leverage Frege's sense/reference distinction: when a university's admission rubric is changed by a shift in sense, without a change in its referent—they will argue—the Constitution remains just as offended as if there was still an explicit facial classification. Because the sense-referent relation has a many-one structure, this sort of argument can be made iteratively. Given this dynamic, a first set of alternative admissions rubrics will come to be widely viewed with skepticism. But this in turn will lead universities and critics of affirmative action to seek out alternatives for the initial post-*SFFA* moves. Suspicion, imperfect information, and the inability of universities to credibly commit to legal compliance will drive a widening gyre of communicative content being treated as a racial "classification." This is Eidelson's "inflationary" tendency at work.

The same dynamic is once again likely to elicit "cycling" between rules and standards. As subsection I.B.5 explained, such instability can already be observed in the historical paths taken by anticlassification doctrine. The ambiguity of communicative content of "race" and "interstate commerce" offers another reason why anticlassification doctrine may not be stable over time: instability is knitted into the analytic foundations of this sort of legal regime. As we shall see in Section III.D, it is not the only possible engine of cycling in the substance of anticlassification rules.[337]

* * *

A more disciplined and less ideologically freighted anticlassification doctrine would rest explicitly upon a clear and justified choice among the different understandings of communicative content. It would then demand an empirically nuanced accounting of relevant shared suppositions and inferences. To say the least, this doctrine would not be straightforward to craft: it would demand many contestable normative and analytic judgments. But it would have the merit of being more defensible than the freewheeling way in which intrinsically justified anticlassification rules are presently cashed out.

D. Extrinsic Justifications for Anticlassification Rules

Could these difficulties be avoided if courts adopted an extrinsic justification for anticlassification rules? Would courts avoid contestable, extralegal judgments if they explained and applied anticlassification rules as if they were devices for "smok[ing] out" the presence of a

337 It may also explain the disagreement between the majority and the concurrence in *McCullen v. Coakley*, 573 U.S. 464 (2014). This turned on whether a protective zone around abortion clinics was a time, place, and manner restriction or content discrimination. *Compare id.* at 480–81, *with id.* at 501 (Scalia, J., concurring in judgment). In effect, the majority was focused on the sense of the statute, while the concurrence trained on its referent.

constitutionally forbidden species of intent on the part of those who enacted a measure?[338] Or if classifications were prohibited because they induced unconstitutional reasoning on the ground? Alas, no.

Recall that the extrinsic justifications for anticlassification rules can be understood as a tracking logic or a provocation logic.[339] The first turns on the possibility that classifications play an evidentiary function by tracking the presence of unconstitutional intentions. This *tracking logic* of anticlassification rules can be found in equal protection doctrine,[340] free speech doctrine,[341] and dormant Commerce Clause doctrine.[342] A similar thought is expressed in the MFN line of free exercise cases, where the Court expresses concern for government "singl[ing] out [religion] for especially harsh treatment,"[343] by implication, intentionally. In contrast, the *provocation logic* of anticlassification rules looks forward, rather than backward, in time to the point at which a law is applied. The person applying that rule, it is said, must account for an impermissible criterion. For a forbidden factor is embedded in the decision procedure stipulated by a law. The resulting state actions are discriminatory, therefore, because they would not have been taken "but for" the illicit factor.[344] The tracking logic and the provocation logic of anticlassification rules both turn on the intentions of government actors. But in the first case, the relevant states of mind precede the law's application, and in the second case they are side effects of applying the law.

Like the intrinsic justifications, these theories of anticlassification rules raise many questions. Again, they invite unstable cycling between rules and standards. It is not clear that either the logic of tracking or of provocation captures a stable relation between enactment text and improper intent.

1. Tracking Logics of Anticlassification Doctrine

To begin with, there is no consensus among judges or scholars about the role that legislative intent should play in determining the

338 Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 226 (1995) (quoting City of Richmond v. J.A. Croson Co., 488 U.S. 469, 493 (1989) (plurality opinion)).

339 *See supra* text accompanying notes 138–43.

340 *Adarand*, 515 U.S. at 226.

341 *See* Ark. Writers' Project, Inc. v. Ragland, 481 U.S. 221, 228 (1987).

342 *See* Tenn. Wine & Spirits Retailers Ass'n v. Thomas, 139 S. Ct. 2449, 2464 (2019).

343 Roman Cath. Diocese of Brooklyn v. Cuomo, 141 S. Ct. 63, 66 (2020) (per curiam); *accord* Simson, *supra* note 116, at 1636–37 (explaining MFN as an effort to smoke out "impermissible value judgment[s]," *id.* at 1637).

344 *See, e.g.*, Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll. 143 S. Ct. 2141, 2155–56 (2023).

constitutionality of legal rules.[345] The Roberts Court often resists the idea that judges should "displace the plain meaning of the law in favor of something lying beyond it" called "legislative intent."[346] Scholars disagree about the coherence and desirability of an intent-based understanding of constitutional rules.[347] At a minimum, therefore, the use of anticlassification rules to track and reveal improper legislative intent rests upon the highly contested premise that such intent exists and matters in the first place.

Even assuming that such intent is relevant, it is not clear facial classifications offer a reliable means of tracking improper intent. Applying an anticlassification rule to this end, the Supreme Court must identify in advance which instances of a category matter to guide lower courts and governmental actors. Without such specification, anticlassification rules could not guide lower-court adjudication. But the very act of picking out impermissible terms creates an incentive for legislators to innovate by seeking out alternate verbal formulations. For the anticlassification rule to play its tracking function, the Court would have to respond by repeatedly enlarging the set of prohibited verbal formulas. Doctrinal inflation would likely be pervasive.

Judicial efforts to manage this inflationary dynamic are apparent in cases applying the First Amendment prohibition on content-based regulation. In its first application of content-based discrimination, *Police Department of Chicago v. Mosley*, the Court held that laws "may not be based on content alone," but also "may not be justified by reference to content alone."[348] Hence, at the very inception of the doctrine, the risk of circumvention induced a measure of doctrinal ambiguity. The felt need to buffer content-based discrimination doctrine with a

345 I use the term "legislative intent" in this passage as a shorthand for the intent of whoever enacted a legal rule.

346 Bostock v. Clayton County, 140 S. Ct. 1731, 1750 (2020).

347 For descriptions pointing to confusion in the caselaw, see Richard H. Fallon, Jr., *Constitutionally Forbidden Legislative Intent*, 130 HARV. L. REV. 523, 526–29 (2016) (collecting cases); *see also* Aziz Z. Huq, *What Is Discriminatory Intent?*, 103 CORNELL L. REV. 1211, 1215 (2018) (observing that "the federal judiciary has not homed in upon a single definition of discriminatory intent" or "a consistent approach to the evidentiary tools through which discriminatory intent is substantiated"); Caleb Nelson, *Judicial Review of Legislative Purpose*, 83 N.Y.U. L. REV. 1784, 1786 (2008) (same). For defenses of the role of legislative intent, see W. Kerrel Murray, *Discriminatory Taint*, 135 HARV. L. REV. 1190, 1196–97 (2022) ("The Supreme Court's antidiscrimination doctrine is widely understood as requiring specific, subjective intent to harm because of a protected trait."); Joseph Landau, *Process Scrutiny: Motivational Inquiry and Constitutional Rights*, 119 COLUM. L. REV. 2147, 2150 (2019) ("A number of commonly used procedures—such as the quality or duration of deliberation, the involvement of experts, the facilitation of regular public hearings and open debate, and the documentation of studies and reasoning behind various policies—provide useful indicators in deciphering political branch motivation.").

348 Police Dep't of Chi. v. Mosley, 408 U.S. 92, 96 (1972).

safeguard against circumvention persists. In the 2015 case of *Reed v. Town of Gilbert*, the Court insisted on a more rule-like doctrine focused on whether a law "is content based on its face" without regard to legislative motives.[349] But even the *Reed* Court acquiesced to an auxiliary inquiry into whether "the purpose and justification for the law are content based" as a precautionary backstop.[350] *Reed* suggests that even judges operating within a highly formalist register recognize the need to craft constitutional decision rules with an eye to potential inflationary dynamics.[351] An inevitable result is that anticlassification rules lack clarity: very quickly, impermissible verbal formulations become fuzzy. What is first celebrated as crisp rule devolves soon into murky and unpredictable standard.

2. Provocation Logics of Anticlassification Doctrine

The second extrinsic justification for an anticlassification rule homes in upon the effect of classifications "downstream" on the behavior of officials implementing the law. The logic here is well illustrated by statutory discrimination cases, such as *Bostock v. Clayton County*, where the Court held that Title VII of the Civil Rights Act of 1964's prohibition on discrimination "because of" sex is violated "whenever a particular outcome would not have happened 'but for' the purported cause."[352] The same account of discrimination has been extended to the Age Discrimination in Employment Act[353] and other statutory antidiscrimination regimes.[354] One scholar has suggested more ambitiously that "the but-for principle ought to reside at the core of the constitutional anti-discrimination inquiry."[355] One way of

349 Reed v. Town of Gilbert, 576 U.S. 155, 165 (2015).

350 *Id.* at 166.

351 *City of Austin v. Reagan Nat'l Advert. of Austin, LLC*, 142 S. Ct. 1464 (2022) can be glossed as the next step of the rules-standard cycle. In effect, worried about the asperity of a formalist content-based rule, the majority throttled back the doctrine to a more fluid standard. *Id.* at 1471–73.

352 Bostock v. Clayton County, 140 S. Ct. 1731, 1739 (2020) (citing Gross v. FBL Fin. Servs., Inc., 557 U.S. 167, 176 (2009)); *accord* City of L.A. Dep't of Water & Power v. Manhart, 435 U.S. 702, 711 (1978); McDonald v. Santa Fe Trail Transp. Co., 427 U.S. 273, 282 n.10 (1976) (Title VII race discrimination claim); *see also* Univ. of Tex. Sw. Med. Ctr. v. Nassar, 570 U.S. 338, 362 (2013) (applying the same formulation to a Title VII retaliation claim).

353 *Gross*, 557 U.S. at 180.

354 Comcast Corp. v. Nat'l Ass'n of Afr. Am.-Owned Media, 140 S. Ct. 1009, 1018–19 (2020) (42 U.S.C. § 1981 claim).

355 Katie Eyer, *The But-For Theory of Anti-Discrimination Law*, 107 VA. L. REV. 1621, 1652 (2021). For powerful criticisms, however, see Andrew Verstein, *The Failure of Mixed-Motives Jurisprudence*, 86 U. CHI. L. REV. 725, 788 (2019) (arguing that but-for tests for intent are "unstable and . . . highly sensitive to the court's definition of the action to be analyzed").

understanding anticlassification, therefore, is as an application of the but-for test in antidiscrimination law. It covers instances in which a law mandates official consideration of an impermissible criterion.

Facial classifications, however, are a highly imperfect proxy for impermissible considerations on the part of officials. The set of cases in which decisionmakers use a notionally impermissible criterion is much larger than the set of cases in which a facial classification is in play. An anticlassification rule, therefore, is radically underinclusive as a proxy for but-for causation. Moreover, nothing distinguishes the use of an impermissible criterion provoked by an explicit classification from other, plainly lawful, uses. Anticlassification rules, therefore, do not pick out a distinctively harmful set of cases in which but-for causation is present.

Despite the recent judicial embrace of but-for causation in statutory antidiscrimination law,[356] notionally forbidden factors still play a necessary and pervasive role in official decisionmaking. A judge or an administrator confronted by an individual's claim of race discrimination, for example, usually accounts for the race of a complainant when deciding whether to treat their accusations as credible. The first step in Title VII antidiscrimination analysis poses the question whether a person is a "member of a protected class."[357] As a result, plaintiffs who do not fit the racial kinds commonly litigated under that statute, such as mixed-race persons, "remain largely unacknowledged."[358] Similarly, courts have narrowly construed the set of cases in which "whites may sue over minority-targeted racism," even when racial animus is a but-for cause of an adverse employment action.[359] In both cases about mixed-race and majority-race plaintiffs, the plaintiff's racial identity plays a necessary role in the decisionmaking procedures—and results in the denial of a judicial forum or a remedy. Nevertheless, neither judges nor scholars have suggested that an anticlassification rule against racial discrimination is triggered by statutes such as Title VII itself.[360]

---

356 *See* Eyer, *supra* note 355, at 1623–25 (summarizing trends).

357 Littlejohn v. City of New York, 795 F.3d 297, 307 (2d Cir. 2015).

358 Nancy Leong, *Judicial Erasure of Mixed-Race Discrimination*, 59 AM. U. L. REV. 469, 476 (2010).

359 Camille Gear Rich, *Marginal Whiteness*, 98 CALIF. L. REV. 1497, 1501 (2010) (explaining that "*only* when their primary motive is to advance the social project of racial equality or promote diversity" can white plaintiffs sue when an employer takes an action motivated by animus against Black employees that has a collateral effect on the white plaintiff).

360 Scholars, however, have suggested that "[d]isparate treatment law is capable of determining when discriminators have acted on forbidden grounds such as race, sex, and religion without requiring plaintiffs to show that they belong to a particular class." Jessica A. Clarke, *Protected Class Gatekeeping*, 92 N.Y.U. L. REV. 101, 106 (2017). Clarke recognizes the

Even in cases that bear traces of the provocation theory of anticlassification, such as *SFFA*,[361] the Court makes analytic moves that cannot easily be squared with that theory. Recall that in *SFFA*, the majority explicitly held that admissions officers could account for how a student "overcame racial discrimination" in a way that showed "courage and determination."[362] This means that race as part of an individual student's life course can be considered. So there will be a meaningful class of cases in which race is a but-for ground for an admissions decision without a constitutional problem. This carve-out to *SFFA*'s general rule undermines the explanatory force of the but-for model of discrimination. That is, admissions officials could notice and employ an applicant's racial identity, so long as they did not assign certain kinds of significance to that trait. This implies there is an "intuitive difference in cognitive significance"[363] between two senses of "race," only one of which raises constitutional concerns. The actual legal effect of *SFFA*, in short, is hard to square with the provocation theory of anticlassification.

Finally, an anticlassification rule is badly fitted to proxy for a but-for causation model. Often, the way in which decisions are made is sufficiently complex that it is not straightforward to ascertain when a factor indeed plays a determining role. The but-for model of causation assumes a linear, algorithm-like process of reasoning that can be decomposed into discrete factors. But real decisions do not come packaged in such neatly tied parcels. Judges need to isolate a "decision," and in doing so can "scale up or scale down the granularity of analysis of the action" in outcome-determinative ways.[364] The judicial analysis of racial gerrymandering, for example, can either focus on a whole district, or it can zoom in upon one discrete stretch of that district's borders.[365] Studies of workplace discrimination covered by Title VII flag the prevalence of "racial emotion," or the "emotions related to race that people experience when they engage in interracial interaction."[366] Empirical research into workplace interactions suggests that "fear, anger, frustration, and anxiety" often follow interracial interactions, both for minorities and majority race groups.[367] Whether race is

present pervasiveness of such gatekeeping, even as she decries it on policy grounds. *Id.* 103–06.

361 Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2155–56 (2023).

362 *Id.* at 2176.

363 Chalmers, *supra* note 298, at 138.

364 Verstein, *supra* note 355, at 788.

365 *Id.* at 790 (discussing Miller v. Johnson, 515 U.S. 900 (1995)).

366 Tristin K. Green, *Racial Emotion in the Workplace*, 86 S. CAL. L. REV. 959, 961–62 (2013).

367 *Id.* at 971, 972–74.

a but-for cause of action in these cases depends on the unit of analysis. Again, anticlassification is a problematic device for identifying but-for causation on the ground.

* * *

Neither a tracking logic nor a provocation logic, in short, offers a satisfactory functional explanation of how anticlassification should be applied. Both theories posit an alignment between facial classifications and a set of constitutionally troublesome facts in the external world. Yet once that posited alignment is closely examined, it turns out to dissolve: anticlassification rules cannot keep up with the fluid dynamics of intentional circumvention. Nor do they map onto the complex ways in which a forbidden factor causes ultimate decisions. Extrinsic justifications for anticlassification rules stand on fragile ground, just like their intrinsically motivated counterparts.

CONCLUSION

The core claim of this Article is easily restated: anticlassification rules are very far from "commonsense."[368] Instead, anticlassification hinges on technical questions of linguistic theory. Exploration of these technical choices exposes complex linkages between doctrine and leading justifications for anticlassification rules. When closely evaluated, these foundations turn out to be weaker than commonly supposed. Some ring hollow. Others must be hedged with caveats and hesitations. Mapping them out casts doubt on whether anticlassification is a tractable doctrinal choice in all circumstances.

Thus, a judge using an anticlassification rule usually must start by building a "category" around a natural kind or social kind drawn from outside the law. The relevant natural kinds have dubious ontological standing, as a close examination of the Court's treatment of the category of "race" shows. At minimum, that "natural kind" is hotly contested in ways the Court blinks. Social kinds often have better grounding in our shared practices—although not always, as the Court's perambulations around the concept of "content discrimination" suggest. But social kinds are also inevitably fuzzy at the edge: there are often no agreed-upon means of deciding which instances are inside or outside the category. What results is a series of effectively arbitrary allocations of instances within or outside the scope of constitutional protection.

Once a category has been stabilized, the judge's challenges have not finished: it is not always easy to say when a forbidden exemplar is present on the face of a legal enactment. Language produces complex

368 Reed v. Town of Gilbert, 576 U.S. 155, 163 (2015).

effects in the arc from semantic to communicative content. It relates in complex ways to underlying beliefs and dispositions. Judges can parse these tangled relations in different ways, depending on whether they hew to extrinsic or intrinsic justifications for an anticlassification legal regime. Once that choice is made, however, there are still serious analytic problems thanks to inflationary dynamics. Judicial formalism, in short, comes at a steep price in judicial candor.

The secondary aim of this Article, indeed, has been to clarify the analytic and normative choices that must be made for anticlassification rules to provide stable rules of judicial decision. A mapping of those choices should not be mistaken for an argument against anticlassification rules. My ambition has not been to show that the formalist strategy they employ is categorically beyond reach.[369] By anatomizing choices implicit in an anticlassificatory legal regime, I hope to have illuminated where its normative justifications have traction. Its foundations are not illusory. But judges and scholars have not paid sufficient attention to when and how formalist doctrinal tools are justified.

When, then, are anticlassification rules warranted? To begin with, the linguistic complexity of anticlassification regimes implies that they cannot be justified on administrability or transparency grounds. To the contrary, anticlassification doctrines more commonly obfuscate. They hide the real normative and analytic choices of judges from scrutiny or criticism. Further, recognition of the necessary benchmarking role of natural kinds and social kinds points to the inevitability of complex normative and analytic choices before any anticlassification regime can get off the ground. It is reasonable to be skeptical of *any* judicial reliance on natural kinds. Incorporating by reference of social kinds, in contrast, seems more tractable. But a social kind must exist prior to its legal use: "content" for First Amendment purposes fails this test. Finally, judges can grasp and apply that social kind with less difficulty if only its existence, and not that of its instances, is "dependent on attitudes that human beings have towards them."[370] "Interstate commerce," that is, is a more tractable social kind than "religion" or "race" because it is not response dependent.

The definition of a legal category still leaves open questions of how it is applied to specific legal texts. In practice, judges arbitrarily move between intrinsic and extrinsic justifications. The latter weakly motivate a decision to use an anticlassification rule because they are radically underinclusive or unstable. The former depend on empirical

369 For an argument to that effect in the statutory interpretation context, see Frank H. Easterbrook, *The Absence of Method in Statutory Interpretation*, 84 U. CHI. L. REV. 81, 83, 90 (2017).

370 Khalidi, *supra* note 154, at 101.

claims about the expressive effects of verbal formulations on behavior. At least these can be tested. A more modest iteration of anticlassification doctrine based on testable extrinsic ground is imaginable.

This sort of revisionist approach to anticlassification demands a pause in the seemingly inexorable adoption of anticlassification rules. It calls on judges and scholars to parse more carefully their own rules and reflect the empirical and justificatory grounds on which they stand. What follows from this may not have such a satisfying air of simplicity. But it will at least have the virtues of coherence, candor, and rigor.