Chapter Title: Conclusion: Merit beyond the Mirror

Book Title: Inside Graduate Admissions
Book Subtitle: Merit, Diversity, and Faculty Gatekeeping
Book Author(s): JULIE R. POSSELT
Published by: Harvard University Press

Stable URL: https://www.jstor.org/stable/j.ctvjghw8s.11

Conclusion

Merit beyond the Mirror

> My first thing about graduate admissions is that we do it as conscientiously as we can, but it is a crapshoot. It really is. We don't know who's going to blossom and who isn't. We have not found reliable predictors.
>
> —Senior professor of classics

IT IS NO WONDER that issues surrounding admissions are among the country's most controversial topics. It's not only educational credentials, but where they are earned, that increasingly shapes professional opportunities, both in academia and in the broader labor market. Two recent studies have found that faculty hiring occurs within largely closed networks and that most of the faculty who are hired into tenure-track positions possess PhDs from a small set of institutions.[1] And in industry, firms in some sectors only seriously recruit from a very small set of "super-elite" undergraduate institutions, creating what sociologist Lauren Rivera called a "golden pipeline" into society's most lucrative positions.[2] In a system like this, the transparency of selection mechanisms into selective colleges and doctoral programs is critical for equity in the system. At both the undergraduate and the graduate level, however, the basis for any applicant's acceptance or rejection is usually opaque to applicants and admissions personnel alike. Uncertainty about what admission requires, combined with the sense that it will play a determining role in their life outcomes, raises the stakes and anxieties for applicants and their families. For those tasked with making admissions decisions, the process may be known, but the basis for individual outcomes may be just as difficult to articulate.

Summary

We know this much about how faculty evaluate prospective graduate students: Test scores, college grade point average, and college reputation play a formative role in the initial review of applications, which explains their relationship with the probability of admission, generally. From the current study, I learned that faculty conceive of "merit" at this point in review principally as a matter of conventional achievement, although scholars across the disciplines may vary in the sections of the GRE they weigh and in considering overall GPA versus grades earned within one's major. They do not interpret every student's scores the same way. Due to China's deep-rooted culture of test preparation and history of admissions fraud, for example, many faculty believe that students should have very high GRE and TOEFL scores but that those scores cannot be trusted to reliably predict student skills. Intriguingly, although it is standard practice to contextualize test scores by national origin, and grades by institutional prestige, most faculty do not contextualize test scores in light of their distributions by race, gender, and socioeconomic status. Just two respondents, an economist and a philosopher, actively encouraged consideration of diversity and/or students' social identities in the initial round of review.

Why the double standard? The most common explanation I heard was that such an interpretation would introduce diversity into the conversation too soon. As a sociologist in the study put it: "First you have to be above a bar, then we can ask the diversity question." Setting extremely high bars on GRE scores and college prestige, however, disproportionately excludes the very populations whom university websites and mission statements claim they wish to attract—and who are already underrepresented in many fields at the levels of doctoral education and the professoriate. For example, 16 percent of Asian American high school graduates enrolled in highly selective colleges and universities in 2004 compared to 2 percent of African American high school graduates. In the physical sciences, 82 percent of Asian and white students earned a 700 or above on the Quantitative section of the GRE, compared to 5.2 percent of underrepresented minority students.[3] The informal admission standards that elite doctoral programs have established in many fields therefore make it extremely difficult for people of color to gain access.

As a matter of procedure, committees commonly maximize efficiency and minimize conflict by discussing only the cases in which initial ratings diverge from each other. As long as committee members rate an applicant similarly, the average rating can serve as a proxy for "the collective assessment of the committee." Through this process, they can quickly eliminate

a large share of the pool using common, academic criteria. But deliberations about borderline applicants are handled quite differently; these decisions often come down to hair splitting or making subtle distinctions between applicants based on novel criteria that were not considered relevant to the assessments of other candidates.

Academic preparation is just the tip of the iceberg when it comes to "merit," because faculty judge many more students smart and competent than their doctoral programs have capacity to accommodate. For the small number whose applications survive to receive full review, judgments are thus holistic, complex, and unpredictable. When comparing generally qualified students, seemingly small matters may matter very much to an applicant's chances because of the layers of inference involved. In the words of one tenured philosopher, "We are in the business of making fine distinctions." A single, ambiguous line in a letter of recommendation; the appeal of a writing sample's introduction; the poor reputation of a letter writer for speaking too highly of too many students; an applicant's weekend hobby or hometown—reviewers may read meanings and value judgments into each of these, in ways that can spell the difference between a candidate's admission or rejection. With few cases discussed and evaluations subject to a myriad of considerations, pinning down what counts as merit late in the process is more difficult than it is earlier in review, when a few key criteria go a long way in shaping ratings. "Everything matters, and nothing matters most," as one faculty member commented.

In this type of situation, *any* basis for comparing applicants can have the effect of a "preference." The individualistic analyses that are inherent to holistic review can elevate opportunities for one student in spite of reducing chances of another with different, albeit also desirable, qualities. When a small program like classics admitted one of two applicants from rural New England because a committee member envisioned it as a "pastoral" place for early socialization in classics, it impinged on opportunities for many applicants whose geographic origins were not perceived as salient. In political science, when a student with mediocre grades was admitted in part because a committee member thought it was "cool" that she wrote for an online magazine, it came at the cost of another borderline applicant who had written a book. Committees in several disciplines admitted applicants from China who disclosed creative hobbies over dozens of their fellow Chinese nationals who did not. What counts as merit when comparing borderline cases sometimes comes down to details that appear idiosyncratic and far afield from conventional achievement.[4]

Many of the grounds for judgment late in the review process are specific to one or two applicants rather than applied to the entire pool, but

those evaluations are all part of an effort to shape the future of their programs and disciplines. Leaving behind the focus on conventional achievement in most cases, deliberations revolved around how applicants "fit" with the program in the present and their idealized academic communities of the future. They sought students who offered fresh perspectives and strong grounding in discipline-specific research dispositions and skills. Diversity, broadly defined, was integral to the fresh perspectives they sought, and a majority strived to construct "balanced" cohorts of students. Rarely, however, did they discuss race, gender, or socioeconomic status in explicit or substantive ways. In contrast to the state of affairs that some worry about, in which race considerations drive admissions in ways that compromise the fairness of the entire process and its outcomes, I was more taken aback by the almost complete silence on these topics in the meetings that I observed.

Explaining the Gap between Principles and Practice

This observation returns us to the questions that motivated my research. With problems of inequality widely known, why do so many faculty rely upon selection criteria that obstruct access for women and underrepresented minorities? If they value diversity, why are they loathe to make racial diversity part of the conversation? I investigated the culture of faculty decision making in elite doctoral programs as a root cause. From there, the answers depend in part on the data and in part on the lenses through which we read it. I will synthesize my findings from three theoretical perspectives before turning to their implications for admissions practice.

Multiple Interests, Multiple Contexts

Assessments of the sort that happen in admissions committees are by nature an elaborate compromise, according to Luc Boltanski and Laurent Thévenot's theory of situated judgment.[5] Graduate programs choose their next cohorts of students with many social goods in mind—student success, prestige, and diversity, to name a few—and each of these will be more or less salient for a particular discipline, program, committee, or reviewer. Practically speaking, admissions decision making consists of negotiating multiple hierarchies of priorities (a heterarchy) that emerge from disciplinary logics, program values, committee dynamics, and personal identities. This perspective would argue that because faculty are trying to satisfy

perceived demands of multiple evaluative contexts that it is rare for any one interest or criterion to consistently hold up as decisive or determinative across reviewers and rounds of review. Only those interests constructed as core *across* contexts are likely to withstand the layers of compromise. A situated judgment angle on my data would conclude that diversity may be one in a constellation of interests that faculty would like to maximize, but especially with respect to race, it has not yet attained the status of a major priority when faculty are thinking about their own values, and those of their programs, committees, and disciplines. Here, it helps to contrast the role of diversity in admissions with that of prestige, an even more well-institutionalized interest.

Prestige. The powerful influence of pedigree and prestige can be understood from two related angles. Pierre Bourdieu's perspective argues that the power of pedigree reflects a broader tendency for elite educational institutions to organize their activities and define quality in ways that preserve their status in the field.[6] Because institutional reputation ("peer assessment") and enrolled doctoral students' mean GRE scores and grade point averages contribute to popular university-ranking systems, prestige-oriented graduate programs use admissions to boost their academic profile.[7] Often this enrollment management decision comes at significant cost to other espoused priorities, such as broadening access for underrepresented groups.

However, status competition does not explain the many individual preferences that also shape admissions judgments. These more complicated patterns signal locally defined—even self-defined—ideas about quality and the tendency for faculty to judge prospective students in the same domains they judge themselves.[8] Michèle Lamont explains how drawing such identity-based boundaries can contribute to social reproduction:

> Exclusion is often the unintended consequence or latent effect of the definition by the upper middle class of its values and indirectly of its group identity and its nature as a community . . . Only when boundaries are widely agreed upon (i.e., only when people agree that some traits are better than others) can symbolic boundaries take on a widely constraining (or structural) character and pattern social interaction in an important way.[9]

I observed faculty using informal conversation to express shared tastes, identities, and goals, thereby constructing symbolic boundaries that guided their assessments of fit and belonging. The student from a politically conservative, religious college was debated as a possible "nutcase." The student with a strong but conventional file was mockingly compared to a Ford. In physics, a lack of research experience was called the "kiss of death."

Prestige thus drives graduate admissions in programs like these because it helps them maintain their status and because it is central to their identity. More than admitting a group whose average characteristics preserve the program's position in nationally defined status orders, it helps graduate programs create the sort of community they feel they are and aspire to be. For such organizations, which already see themselves as successful, the status quo therefore represents a powerful default. This is particularly the case, as I will discuss, if changing the basis for admission may require some soul searching about flaws in their collective identity or goals, and if admission already requires more time and effort than is desirable.

Diversity. In contrast to prestige, diversity is relatively new among admissions priorities. How has it emerged as an interest at all? According to the organizational theory of institutionalism, organizations survive by adapting their practices and priorities to changes in the institutional and political environment. Shifting values in society can bring about new policy in organizations like colleges and universities by motivating changes that might otherwise have been avoided.[10] This institutionalist angle helps explain why educational equity came to be recognized as a social imperative during the Civil Rights Movement, and why diversity has come to be viewed as a practical interest for higher education today.[11]

A very brief history can situate the current state of diversity as an interest for higher education stakeholders. Civil Rights protests changed the political environment in the 1950s and 1960s to support demands for improved access to higher education and employment for African Americans, Latinos, Native Americans, and, in spaces where they were still excluded, white women. The Civil Rights Act of 1964 and executive orders for affirmative action under Presidents Lyndon Johnson and Richard Nixon made these shifting values visible, and they presented them to the public as means of redressing the effects of long-standing policies of segregation, discrimination, and exclusion. These macro-level influences, coupled with acute pressures that campus-based student movements placed on university administrators, compelled the adoption and diffusion of race-conscious admissions policies by selective colleges and universities across the country.

The difference that scholars and institutional researchers found diversity to make for student learning and development provided an empirical foundation for another affirmative action rationale—the educational benefits of diversity. Proponents sought to broaden the appeal of affirmative action by emphasizing its benefits for all students' development, namely white students, not only for those who bore the "present effects of past injustice." Thus, whereas universities implemented affirmative action due to changes

in the political environment, university efforts since the 1970s have been to protect the legitimacy of race-conscious admissions by elevating the rationale with the widest public support.

Institutions have also focused on diversity's educational benefits because the courts have established narrow parameters for the constitutionality of race-conscious admissions. In 1978 in *Regents of the University of California v. Bakke* the U.S. Supreme Court ruled racial quotas unconstitutional, and in 2003 in *Gratz v. Bollinger* it also struck down the awarding of "automatic, predetermined point allocations" to underrepresented students.[12] Justice Lewis Powell's controlling opinion in *Bakke* ruled consideration of race permissible as a "plus factor" in admissions and financial aid decisions, but this consideration requires narrow tailoring to a compelling state interest.[13] What counts as a compelling state interest? In *Bakke,* whose precedent has been upheld in such recent decisions as *Fisher v. University of Texas,* the Court rejected three of the four interests in affirmative action asserted by the University of California–Davis Medical School. However, Powell's opinion affirmed the educational benefits of diversity, broadly defined, citing promotion of diversity in higher education as a "compelling governmental interest." Powell also upheld educational institutions' discretion to determine the selection procedures that best suit their needs, associating it with academic freedom and the First Amendment.[14] In the years since *Bakke,* voters have banned affirmative action through ballot initiatives in Arizona, California, Michigan, Nebraska, Oklahoma, and Washington, and the legislature and governor passed anti-affirmative policies in New Hampshire and Florida. These bans are responsible for declining racial/ethnic diversity in selective undergraduate institutions and in many graduate fields of study, amid increasing racial and ethnic diversity in the population.[15] For academic institutions in the other forty-two states, however, principles from Powell's opinion are still the law of the land, and are important for those engaged in admissions to know. A brief from the Civil Rights Project at UCLA summarizes them:

1. Reserving seats or proportions of seats specifically for underrepresented students is not permissible.
2. Reviewers should use a common process of review for all applicants.
3. Race should be one of several individual characteristics assessed as a plus factor in the effort to promote diversity.
4. Every applicant should be evaluated as an individual, and should not be assumed to represent a broader identity category.
5. Programs should not single out specific racial/ethnic groups, but consider the contribution that all groups make to diversity.[16]

Bakke focused on admission to medical school and *Grutter* on law school, but the Supreme Court has not specifically examined graduate school admissions among the arts and sciences. However, economists, psychologists, sociologists, as well as higher education and legal scholars have all advised that selection processes and their rationales be tied to educational mission.[17] In addition to ethical and economic rationales, the distinct mission of doctoral education may therefore elicit interests in equity and diversity that are distinct from those in undergraduate and professional education. As discussed earlier, research by teams possessing diverse social identities is associated with core characteristics of scholarly quality. Heterogeneous research teams have demonstrated advantages in creativity and problem solving, and publications that result from ethnically and gender diverse research collaborations are cited more often.[18] Relatively few people I interviewed associated increasing social diversity with intellectual excellence or the health of their fields. Important new analyses also assert that, if reducing inequities in higher education is the goal, then "diversity" itself merits critical evaluation as a rhetorical strategy for organizational behavior.[19]

From the perspective of situated judgment, what counts in practice as merit is as an institutionalized compromise across multiple interests associated with the multiple social contexts that decision makers represent. Professors strive to simultaneously maximize individual, committee, program, and disciplinary interests, and to uphold their personal values, interpersonal relations, and institutional and wider policy. Prestige and diversity may both be organizational interests, but only if something is conceived as a core interest *across* those contexts will it be likely to survive as a priority in the negotiation process. We can explain continuing inequality in spite of diversity's institutionalization through (1) the entrenched value that status and prestige hold for organizations and participants like the ones in this study, (2) the more tenuous place that diversity holds (legally, politically, and discursively) as a value in their disciplines, programs, committees, and for many of the individual reviewers, and (3) evaluative scripts that associate very high GRE scores and attending elite institutions with intelligence and belonging (two additional shared values) in prestigious communities like theirs, and that construe lower GRE scores or less selective college enrollment as a trade-off with excellence. However, same deliberative processes by which faculty collectively define merit can be used to diversify what counts as merit. Decision makers may encourage alternative interpretations of common criteria, for example, or use admissions and hiring to encourage new perspectives or collective goals.

Preferences Imply Aversions

Decision theory offers a second perspective from which to interpret this study's findings. Although research and public debate about admissions often revolve around the justice and necessity of specific evaluation criteria, a contribution of this research is to show that decision-making processes are also implicated in ongoing reliance on pedigree and GRE scores. Specifically, participants' preference for a *process* that reduces uncertainties and preserves collegiality revealed their shared aversions to ambiguity, risk, and conflict.

Ambiguity aversion. The complexity of comparing hundreds of applicants who come from different national, institutional, and personal backgrounds makes faculty *ambiguity averse*. Faced with a "cloud of random applicants," professors hunger for clarity and conviction about which ones are likely to thrive as scholars and bring vitality to their departments and disciplines. From this standpoint, reliance upon test scores is not exclusively about what they signal, but instead about how they make it easy to compare applicants. Participants associated GRE scores with convenience because their apparent standardization, precision, and clarity seem to cut through all of the apples-to-oranges comparison that evaluating students' letters of recommendation, essays, and writing samples requires.

Professors in the humanities were both more thoughtful about and less averse to the subjectivity inherent in judging and comparing their applicants. Yet they too ultimately quantified their judgments to bring closure to deliberations, and they were especially likely to do so for the sake of expediency. Though fallible, "the numbers" provided something on which humanities faculty felt they could stand. By contrast, in the physical sciences, greater trust in numbers engendered trust in the validity of distributions of test scores and average ratings. In both cases, quantifying differences in perceived admissibility was a hallmark of deliberative bureaucracy, one that increased efficiency by masking perceived ambiguities. For busy professors, the GRE's appeal was in no small part the possibility that it would simplify their work.

Risk aversion. There is also abundant evidence that faculty in these prestigious programs engaged with admissions work from a posture of risk aversion. This tendency, paired with the specific assumptions about risk that they frequently made help clarify why they favored criteria that undermine increased racial diversity. Many respondents associated less selective undergraduate institutions and lower test scores and grades with less prepara-

tion and intelligence, and they associated less preparation and intelligence with a higher risk of academic struggle or failure. This risk aversion script pervaded ratings and deliberations at *all* stages of review. Risk aversion also provided a bottom-line basis for rejecting well-pedigreed individuals who had personality red flags, research interests that aligned only marginally with faculty expertise, or letters of recommendation that raised eyebrows.

Organizational theorist James March has found that a propensity to risk aversion is common in high-status organizations' decision-making processes.[20] In this case, faculty rationalized risk aversion as an organizational luxury, a matter of financial prudence, and a foundation of responsible decision making. They knew they could *afford* to be risk averse because their highly ranked programs attracted so many qualified candidates. Further, they convinced themselves they *should* be risk averse in light of the financial investment involved and what admitted students represent—namely, the program, discipline, and university's quality, status, and future.[21]

High-status organizations are also more prone to fundamental attribution error—the tendency to attribute the outcomes of their members to intention and skill rather than to the conditions in the environment. Under these conditions, high-status organizations often simply exclude prospects perceived as risky rather than measuring the risk and making a decision on that basis.[22] This point highlights a critical detail about the risk aversion I observed—that it was almost always grounded in perceptions of risk rather than generalizable evidence or calculations of it.

Availability bias offers a third way of understanding risk aversion in terms of cognitive biases. Trusting the reliability and validity of the limited information that is currently available to decision makers can lead them to overestimate the likelihood of rare but memorable events. Availability bias facilitates group-based stereotyping, and I observed it to be a process by which racism was subtly institutionalized in the admissions process. For example, a number of faculty regarded African American, Latino, or Native American applicants from less selective colleges or non-native English speakers from East Asian countries as admissions risks because their programs did not enroll many of these students and recollections of one or two students who had enrolled and struggled were hard to shake from their memories. Several participants admitted that they felt "spooked"—three using this particular word—when reading applications from students whose profiles reminded them of individuals who did not graduate.

Availability bias is especially dangerous when combined with fundamental attribution error. The combination of these two biases locks in faulty assumptions about who is successful and why, and then uses those faulty assumptions as a basis for distributing future opportunities. First, attribution

errors blind academic departments to how the climate and culture that they create may contribute to students' successes and struggles. Then, availability bias encourages extrapolation of a few students' struggles (which resulted in part from organizational conditions the department could change) to a larger category of students as a basis for evaluating their future potential, and thus their deservingness for opportunities. Taking steps to actively check the natural tendency toward these biases can prevent the experiences of a few individuals from becoming the basis for judging whole categories of future applicants as risks. It also highlights the importance of encouraging faculty mindfulness about their own role in shaping student success.[23]

Conflict aversion. Quantification also suppresses conflict among faculty. Conflict aversion is the flip side of the deeply rooted professional norm of faculty collegiality. Admissions committee members felt that quantifying their judgment—whether through test scores, individual ratings, or average ratings—provided more uniform interpretation than the many understandings that the committee might generate about individuals and their relative admissibility. With a few keystrokes, numbers could be sorted into a single ranked list, obviating the need for difficult negotiations about who should be deemed more or less admissible. Most departments viewed some deliberation as prerequisite to legitimate admissions decisions. However, committee members preferred not to argue with each other, especially about matters that were rooted in deeply held epistemological differences, such as the validity of statistical inferences about GRE scores. As such, two key elements to the deliberative bureaucracy decision making model that I observed included careful committee appointments and substituting potentially controversial discussion of applicants with relatively uncontroversial discussion of admissions procedure. The formal task of admissions may be to identify applicants with the most impressive achievements and greatest potential, but the process of collective selection was one of political compromise and avoiding uncomfortable conversations.

Resistance to Change, Ambivalence about Diversity

By combining elements from the first two theoretical perspectives—inconsistencies across context in how faculty weigh diversity and the important role of aversions in shaping faculty decision making—we can also interpret deference to entrenched evaluation criteria as a matter of resistance to change. An entire scholarly literature has developed around

resistance and ambivalence to organizational change, with Louis Menand going so far as to argue that twenty-first-century American professors are trying their best to maintain a nineteenth-century system. They want to make a better world through their research and teaching, but often eschew making change in their own ranks to help bring it about. Among the admissions committee chairs I interviewed, outright resistance to change did not seem to be the problem. Indeed, three of them welcomed me to observe their admissions committees to learn how their systems could be improved. However, their colleagues were in many cases more ambivalent—about changing admissions, generally, and about rethinking their applicant pools and diversity, specifically.

Why is it that faculty might stand by equity and diversity in principle, but feel ambivalent about practical efforts to achieve them? For one thing, many thought of these as obligations dictated by social norms and for their pragmatic benefits, not as conditions that demonstrably improved their work or their community. Faculty typically build their careers on the visions of academic excellence they inherit from their academic forebears, and for most, diversity is not yet integral to that vision.

Thus, in political science, linguistics, astrophysics, and philosophy, faculty supported their *colleagues'* advocacy for a small number of borderline applicants with nontraditional profiles, but they were reticent to take on such students as their *own* advisees. They expressed concerns about the additional mentoring investment that might be required. The self-trained philosopher from the Deep South, the astrophysicist born in the Himalayas, and the political scientist with unusually low GRE scores each received admissions offers, for example, but only after extended debate and assurance that someone else would take responsibility for their supervision.[24] Most faculty could be persuaded to admit specific individuals in the name of diversity, but I observed little evidence that diversity was either understood to be a collective responsibility or enacted as a shared value in the day-to-day of department life.

Postracial discourse in American society may also contribute to ambivalence about reform. For one, claims that racism is a problem of America's past undermine attention to continued structural inequalities and everyday biases.[25] Constraints on affirmative action reflect postracial sensibilities, and I learned that regardless of whether or not a program was subject to specific legal constrains, that bans, rumors of bans, and worries about lawsuits all made faculty reticent to discuss race at all with respect to recruitment and admissions.

However, like the other factors I have discussed, uncertainty about options for equity-based reform must be read as an explanation, not an excuse.

Research by Roger Worthington and colleagues found that "colorblind-ness" and personal privilege predispose those within higher education to an overly rosy view of the status quo on their campuses.[26] Most faculty in the ivory tower—especially in elite institutions—are distanced and well sheltered from the struggles with exclusion, marginalization, and discrimination that have historically compelled equity-minded reform efforts in higher education. I found that, with the exception of some sociologists and a handful of others, most participants in this study did not see equity or diversity as *their* issues, or the students admitted under these banners as *their* students. I did not collect data to measure implicit bias against women and students of color, but we must not cast aside the evidence of this pattern uncovered in prior research. Faculty judgments of many applicants from China were consistent with model minority stereotypes, providing a window into the likely presence of other identity-based biases of which faculty may have been unaware or more careful to conceal.

Finally, organizational change tests collegiality, especially when it occurs through democratic processes. It can be costly in time and effort, two scarce resources for faculty operating in a reward system that privileges research productivity over student development or improving departmental struc-tures. Were faculty to receive rewards or incentives for time spent strength-ening admissions, not to mention student mentoring, maybe the gap between expressed and enacted commitments would not be quite so wide.

With myriad social forces acting upon faculty as they select and discuss applicants, we need to think systemically when we think about improving admissions. Faculty members' ambivalence about admissions reform has several dimensions, and must, itself, be understood as one among several explanations for continued inequality. Shared aversions to ambiguity, risk, and conflict shape the work of decision making in powerful ways, and pres-tige often trumps other organizational interests when faculty seek shared values to which they can collectively defer. In addition, the perceived costs of admitting more diverse cohorts, a lack of awareness about compelling alternatives to affirmative action, satisfaction with the way things are, priv-ilege and personal distancing from equity-related issues, and a changing sociopolitical context each stand as barriers to change.

From this angle, the slow pace of change is hardly a surprise, because change is needed on many fronts. There are no silver bullets, easy answers, or quick fixes to systemic problems. Faculty may be so reluctant to wrestle with the complexity of diversity and inequality that they avoid altogether the conversations required to get the process started. Comments from Jonathan, a white male associate professor, summarize these points perfectly:

It's uncomfortable to be reminded that I'm part of the problem. What I don't want are things that will feel unnatural, take 300 percent more time, or that my colleagues will fight. I want to keep my sanity as a professional, and I don't want admissions work to consume me. What I want are straightforward ideas that help me be part of the solution.

Implications for Practice

With this idea and Jonathan's quote in mind, I turn now to discuss the implications of this research for graduate admissions practice. "Admissions decisions are actionable choices," as higher education scholars Rachelle Winkle-Wagner and Angela Locks wrote, and although professors may inherit their colleagues' approach to admissions, they also have autonomy in most universities to change the way things are done. My research was not designed to encourage policy prescription, but integrating its findings with those of other recent research studies is useful for sketching a framework for structural changes needed in many fields and programs. Within this framework, graduate programs can craft and evaluate practices that suit their own unique needs.

Revisit Admissions Routines and Make Them Explicit

All but two of the graduate programs in my sample would have benefited from taking a fresh look at their routines for recruitment, admissions, and awarding financial aid. For example, developing a thorough list of the many steps in admissions that involve subjective judgments may aid faculty in checking themselves for implicit biases. Which email inquiries a professor chooses to follow up on with a phone conversation, which colleges a graduate program chooses to visit for recruitment, who should be awarded financial aid packages as a recruitment tool—each of these procedural decisions involves a degree of subjectivity.

Also, considering the important role of initial reviewer ratings, many graduate programs can improve admissions generally by improving the review of individual files. Like selection criteria, approaches to evaluating files come with benefits and drawbacks, and each program must decide for itself what the optimal review process is, in light of what they are trying to accomplish. I will describe here some benefits and drawbacks of two models for evaluating files.

Open, holistic review incorporates into evaluation all available information about students. One strength of a holistic approach is that students

Table 10   Example Rubric

	Academic Preparation	Potential for Scholarship	Contributions to Dept., Cohort, Discipline	[Insert additional evaluation standards]	Basis for Ratings
Rating	Weak-Acceptable-Strong	Weak-Acceptable-Strong	Weak-Acceptable-Strong	Weak-Acceptable-Strong	
Applicant 1					
Applicant 2					
Applicant 3					
... and so forth					

are easily assessed in the context of their individual trajectories and opportunities. Open-ended review can lack transparency, however, putting the applicant in the position of playing a game whose rules are unknowable. This opacity maximizes institutional discretion, but it can also allow unseemly preferences and biases to enter the review process.

More structured approaches to review define in advance the criteria on which every application will be assessed. The National Science Foundation's ADVANCE program has developed rubrics and other evaluation tools that codify and define selection criteria and ask reviewers to provide a rationale for their ratings. Table 10 presents a very basic skeleton of a rubric that departments could customize. Rubric-based evaluation raises transparency and does not preclude reviewers from contextualizing applicant characteristics or attending to underlying rationales. Rubrics can also be designed to note desirable qualities that come along so rarely that it makes little sense to include them in evaluations of everyone.[27] Using a rubric increases *equity,* by comparing all applicants using the same criteria, and *efficiency,* by focusing reviewer attention on key information. Rubrics do raise the stakes associated with each criterion, so they need to be chosen with care.[28]

Strengthen Recruitment and Align It with Admissions

The need to strengthen outreach and recruitment early in the admissions cycle was an almost universal pattern across the ten programs. Two programs had experimental outreach efforts in place to build the diversity of their applicant pools, but weak coordination with admissions committees undercut those initiatives' efficacy. The problem in the other eight programs, however, was that early recruitment strategies were virtually nonexistent, consisting of little more than email responses to inquiries from prospective applicants. Departments looking to build a more diverse applicant pool have several options, including developing relationships with sister departments in institutions whose undergraduate alumni of color frequently go on to earn doctoral degrees. A recent study examined the baccalaureate origins of African American, Latino, and Asian/Pacific Islander doctoral degree recipients, and assembled lists of the colleges and universities that produce the most doctorates of color for each racial/ethnic group. Table 11 lists these institutions, many of which are minority-serving institutions (MSIs).[29] Graduate school administrators and diversity officers could assist academic programs in developing discipline-specific lists of this sort, which would enable faculty to recruit where students are, rather than waiting for students to approach them.

Table 11   Top 30 Institutions Producing African American, Latina/o, and Asian/Pacific
Islander PhDs

	Top 10 African American	Top 10 Latina/o	Top 10 Asian/Pacific Islander
1	Howard University	University of Puerto Rico–Piedras	University of California–Berkeley
2	Spelman College	University of Puerto Rico–Mayaguez	University of California–Los Angeles
3	Florida A&M University	University of California–Los Angeles	Massachusetts Institute of Technology
4	Hampton University	University of Texas–Austin	Harvard University
5	Southern University A&M College	Florida International University	University of California–San Diego
6	Jackson State University	University of Texas–El Paso	Cornell University
7	Morehouse College	Harvard University	Stanford University
8	University of Michigan–Ann Arbor	University of Florida	University of Hawaii–Manoa
9	North Carolina A&T University	University of New Mexico	University of Michigan–Ann Arbor
10	University of California–Berkeley	University of Arizona	University of California–Davis

Sources: Lundy-Wagner, Vultaggio, and Gasman 2013, 158, citing Survey of Earned Doctorates 2009.

Informal recruitment relationships can become institutionalized over time, such as the Fisk-Vanderbilt Masters-to-PhD Bridge Program. This respected multidisciplinary program identifies promising undergraduates from Fisk University who would like additional preparation or research experience before beginning a PhD. Students enroll in a master's degree program at Vanderbilt, which subsequently offers them "fast-track admission" to participating doctoral programs there.[30] Bridge programs are growing in popularity as a means of broadening access to doctoral education, especially in STEM fields, and the National Science Foundation regularly fields calls for proposals for institutions that would like to develop them.

Undergraduate research programs also develop students who are interested in and qualified for doctoral study. Faculty-supervised research experience has long been promoted as a nonremedial strategy to encourage the success of undergraduates from underrepresented backgrounds. Among the studies that find positive outcomes of undergraduate research participation, scholars at UCLA's Higher Education Research Institute recently con-

cluded through a rigorous quantitative methodology that it increases black, Latino, and Native American students' participation in graduate education.[31]

Other research, using qualitative methods, indicates that student research promotes continuation to graduate education through mutually reinforcing processes: (1) attracting students by solidifying aspirations and providing early socialization, and (2) making applicants, especially with less conventional profiles, more attractive to faculty in the admissions process. Through extended engagement with research, students cultivate forms of cultural capital deemed valuable in doctoral education and the academy, including fluency with disciplinary jargon, ease in discussing technical details and the significance of one's research, and national presentations or publications.[32] I found in this study that professors make critical inferences about prospective students from their research experiences, such as how they might contribute to the future of the discipline. As with many admissions criteria that involve some subjectivity, decision makers may want to set norms before they review files about how they will evaluate research experience, rather than allowing themselves to be impressed by students with the most prominent research experiences.

Finally, I want to offer a few comments about a common recruitment dilemma. Graduate programs striving to increase their enrollment of a particular group often find themselves stuck in a negative feedback loop: For admitted students from the underrepresented group debating whether or not to enroll, an absence of individuals who share their identity can raise red flags about what the quality of their experience will be. This uncertainty may contribute to their decision to matriculate elsewhere, which perpetuates the group's underrepresentation. Breaking this cycle requires strong leadership and evidence that diversity is, for the program, more than a platitude. Those in recruitment roles should be ready to initiate honest conversations about how and why diversity matters, what it means for their scholarly work, what students from underrepresented backgrounds will gain from and offer to the program, and what the climate is like for students from that group in the program, campus, and broader community. Especially for people from privileged backgrounds, it can be hard to speak frankly about these issues because they are always works in progress. However, developing the proficiency to do so in recruitment, admissions, mentoring, and instructional contexts is part of the change process itself.

Examine Assumptions about Merit

Perhaps the most fundamental implication of this study is the need to reframe relationships among fundamental principles like excellence, merit,

and diversity. Diversity was one dimension of merit when faculty assessed applicants on the short list, but many spoke about "diversity with excellence" as well—as if these are independent entities accompanying one another. The Association of American Colleges and Universities recommends educational institutions work toward a culture of inclusive excellence, which

> integrate[s] diversity, equity, and educational quality efforts into their missions and institutional operations . . . The action of making excellence inclusive requires that we uncover inequities in student success, identify effective educational practices, and build such practices organically for sustained institutional change.[33]

In admissions, inclusive excellence means broadening recruitment efforts beyond the relatively small network of colleges and universities through which opportunities tend to flow, reassessing assumptions about what it means to be qualified (and what the grounds are for those assessments), and tracking the equity implications of current selection practices and any proposed reforms.[34]

Making excellence inclusive also means recognizing that common performance metrics do not tell the full story about underrepresented students' potential.[35] Research is needed to refine measures and assess their predictive validity, but there is a growing movement to formally assess students' non-cognitive strengths as part of a holistic definition of what it means to be qualified for doctoral education.[36] For example, the Fisk-Vanderbilt Master's-to-PhD Bridge Program added an interview to their selection process that assesses students' grit and resilience. It asks them to reflect upon their interest in science, how they persevered through challenges, and the resources and relationships to which they turn in struggles. Their outcomes have been impressive. Since 2004, 81 percent of its entrants have continued on to doctoral studies, which flips a national trend in which 80 percent of students of color with STEM bachelor's degrees do not continue to graduate school.[37] Raising the profile of promising alternatives to current educational and selection models has value when a dearth of alternatives is often presented in defense of the status quo.

Stop Misusing Standardized Test Scores

For some faculty, meeting a standard of fairness in admissions will require that they learn more about the proper interpretation and limits of standardized test scores (GRE, TOEFL, GMAT, others) and reassess their use in the admissions process. Documentation from ETS, for instance, indicates that

the GRE measures skills. It advises that raw scores and percentiles should not be used as sole or primary admissions considerations, and that score thresholds should not be applied.[38] However, decision makers in these programs routinely used scores as a primary consideration in the first round of review. About half thought of scores as signaling intelligence, and it was common practice to assess applicants against formal or informal score or percentile thresholds. Simply put, they misused the GRE.

The appeal of simple metrics is undeniable. Indeed, anything that makes this complicated review process more convenient is hard to ignore. However, the patterns I observed both run contrary to the psychometric properties of the test and put already underrepresented populations at a disadvantage. By failing to read scores in context, reviewers made overly simplistic comparisons about students' relative potential. Many had a general awareness that GRE scores have patterns associated with race, gender, and first-year graduate school GPA. Fewer knew that research studies have come to very different conclusions about the strength of the correlation between test scores and longer-term academic outcomes, or that stereotype threat helps explain group-level disparities.[39] Professors thus need better knowledge of ETS guidelines, of research about the full scope of factors that test performance can reflect, as well as its validity for different groups of students and for short- versus long-term outcomes. With this knowledge, admissions leaders can come to more thoughtful decisions about whether to require the test and, if they choose to do so, how to appropriately use and interpret its results.[40]

Awareness of the many problems with standardized test scores—that they are more complicated than they appear, that reliance on them can undermine equity efforts, that they do not reliably predict students' long-term success—has led some higher education institutions to reduce their reliance on these measures or eliminate their requirement entirely. As of summer 2014, the National Center for Fair and Open Testing reported that 103 selective colleges and universities had implemented test-optional admissions processes. At the graduate level, doctoral programs in at least thirty-two different fields have eliminated their GRE requirement, including ones at Columbia University, University of Pennsylvania, Johns Hopkins University, and the University of Michigan.[41] Research is needed about the effects of going test optional, both at the undergraduate and graduate levels. In one study, results indicated that making standardized test scores optional resulted in an increase in the enrollment of women, students of color, and international students, without changing the institution's mean GPA or graduation rate. But in another study, researchers found that going test-optional did not on its own increase diversity, but instead increased selectivity because only

the applicants with higher scores submitted theirs.[42] If increasing the enrollment of students of color is specifically a goal, making GRE scores optional or declining to review them are options that should be on the table. However, they will be more successful as part of a coordinated set of efforts.[43]

Seek Small Wins in Admissions as Part of a Multifaceted Change Process

Changing culture and confronting inequality can seem daunting prospects, and that may be part of the problem. According to organizational theorist Karl Weick's idea of "small wins," how we define long-term challenges affects the chances of making progress on them, because when social problems are conceived at large scales, people become immobilized from taking action and they perform more poorly when they try. "When the magnitude of problems is scaled upward in the interest of mobilizing action," he writes, "the quality of thought and action declines, because processes such as frustration, arousal, and helplessness are activated." Weick therefore recommends that organizations approach major challenges by "identify[ing] a series of controllable opportunities of modest size that produce visible results," particularly in how they interpret and classify problems, which shapes the scope of possible solutions.[44] Among the examples of small wins that Weick cites is Alcoholics Anonymous conceptualizing sobriety as a goal that is achieved one day—even one hour—at a time.

Improving admissions is a small win relative to systemic problems of inequality in academia and society. Yet to bring the scale down to something even more manageable, admissions itself holds possibilities for small wins because the work is replete with classification. The endgame, of course, is to classify a subset of the applicant pool as admitted, but along the way decision makers employ an elaborate mental classification system when they infer complex qualities like intelligence and potential from the application. Counterscripts—alternative interpretations of common criteria—serve as a small win by systematically shifting how decision makers mentally classify students. By focusing on engrained assumptions, small wins can build consensus and create sustainable change.[45] By shifting discourse they help shift conventional wisdom, motivating change as they signal that such change is already under way.[46]

The field of philosophy offers an excellent case of how small wins in admissions can be part of a broader strategy for equity-minded change, in their case, with respect to gender. In one of the two the committees I observed, female members and male allies actively challenged their colleagues on misinterpretation of GRE scores, and they offered counterscripts about

how to read undergraduate college prestige. On the latter, they introduced into the deliberations some of the reasons other than talent (such as financial barriers, family commitments, or opting out of the prestige chase) that students attend less-selective institutions. The support of a critical mass for these counterscripts shifted the dominant narratives about at least three borderline applicants. Instead of playing a determinative role, scores and affiliations counted as two among many considerations.

In addition to broadening the meanings attributed to GRE scores and college prestige, individual faculty, graduate students, and small groups in both of the philosophy departments I worked with are engaged in other departmental and national efforts to increase women's participation. At the department level, they are hosting coffee hours and dinners for aspiring female philosophers, and creating structures through which individuals can safely report experiences of discrimination or harassment. At the national level, female faculty in these departments are working through formal and informal collectives and campaigns.[47] There are active networks promoting equity and opportunity for women through the American Philosophical Association's Committee on the Status of Women (APACSW) and the Women in in Philosophy Task Force, for example. One recent APACSW campaign gathered and posted online information from master's and PhD programs about their enrollment, retention, and job placement rates by gender, as well as comments about each one's efforts and achievements to increase women's participation. Another popular campaign compiled photos of more than 200 women in philosophy into posters with the line "Philosophy. Got women?" There are also anonymous, cooperatively published websites (titled "Feminist Philosophers" and "What is it like to be a woman in philosophy?") that provide a public forum for speaking out and sharing experiences. As philosophers work to change the culture of philosophy, changing the assumptions that decision makers bring to admissions evaluations is just one part of their efforts.

Conclusion

I have argued that judgments of the "best" graduate programs in the country and the "best" applicants in the pool come about only by negotiating across priorities associated with specific social contexts. Thus, there is wisdom in my colleague Paul Courant's admonition to "beware the tyranny of best practices." The best practices for admission to a department's doctoral program may be different from those used to admit master's-level students, who may have different professional goals. What is best for recruiting

graduate students to New York or Los Angeles may be different from what it takes to draw students to Madison or Boulder. From this perspective, a final implication of this study is not the need for a specific set of practices, but rather for faculty to approach their gatekeeping work with a different state of mind. Instead of proceeding by default or adopting a specific checklist of procedures developed for some other program's needs or goals, admissions decision makers need to approach gatekeeping with mindfulness of their own situation and needs. When evaluation and selection happen by default rather than intention, the social consequences of decision making may be misrecognized as the normal course of events—the way things *should* be rather than the way things ended up. And over time those outcomes can be thought of as natural rather than deliberate, for the more deeply held a belief is, the more likely it is to keep a person from seeing things any other way.

Today's faculty choose students on the basis of an array of perceptions that only sometimes have a strong evidentiary basis. In a process so competitive that the mere *presence* of doubt can seal an applicant's fate, perceptions often carry the weight of truth. They drive decision makers to act—to rate and advocate, to admit and reject—as if perception were reality. As the theorem made famous by sociologists William and Dorothy Thomas puts it: "If men define situations as real, they are real in their consequences."

Context and circumstance will always establish broad parameters within which we make choices. Yet social constructionism teaches us that the social world is as we perceive it to be because we act upon our perceptions.[48] Admissions judgments are socially situated; they result from inference upon inference, and from filters associated with organizational contexts and goals. The benefit of research like this study, which puts participant interpretations at the center of analysis, is in uncovering these taken-for-granted filters, for it is only by stepping outside of and analyzing them that we can understand their power. Like the prisoners in Plato's cave, misled to believe that shadows on the wall were reality, scholars may not even realize that the way they have been conditioned to see excellence in themselves and others is not natural, but constructed to serve specific ends. Yet unlike the prisoners, scholars can shape the social contexts in which their judgment is situated. In doing so, they refine the lenses through which their own understanding—and that of their colleagues and students—will be filtered.

Indeed, one of the great values of understanding an organization's gatekeeping systems is that it lays bare cultural values that drive policy, which are often so engrained as to be taken for granted. Too often our preferences and aversions go unspoken—even unrealized at a conscious level—so that

misguided assumptions and implicit biases are never challenged and continue to shape the outcomes of review. Honest dialogue about the values and norms that shape interpretations of scholars' records is therefore needed, not only in admissions, but across the disciplines and the range of gatekeeping processes the academy employs. These conversations are relevant to how we admit students; how we hire postdoctoral researchers and faculty; how we select individuals for awards, grants, and fellowships; how we advance students to candidacy; and how we assess faculty scholarship for tenure and promotion. Each of these evaluation processes is fundamentally an activity of defining and monitoring organizational boundaries, of determining the grounds for membership, belonging, and recognition.

The faculty I spoke with want their students to succeed, and they like the idea of increasing diversity of many sorts. They carry out gatekeeping activities consistent with field-level norms and with those of their own training. Yet having mostly graduated from selective doctoral programs themselves, and with limited exposure to other models, their inclination may be to so strongly associate their own training with "quality" training that they recreate its selection process and the educational gauntlet in the interest of preserving what they see to be the purity of the discipline and excellence of the program.

However, society is changing. The labor market for PhDs is changing. And higher education is changing, too, both intellectually and demographically. If doctoral education and our means of identifying talent do not change with them, it will be doctoral education and the professoriate that fail by falling behind, not our students. Re-creating academic programs and disciplines in our own image may be the natural tendency, but in a changing world—one that is more diverse, more collaborative, more interdisciplinary— stewardship of our disciplines' futures means being flexible to recognize what the world needs from our fields of study and adapting our ideals of excellence accordingly.