Assessing Institutionalized Bias

Linda F. Bisson, Philip H. Kass, Kyaw Tha Paw U, and Laura Grindstaff

Abstract Ideally, higher education systems are meritocracies in which advancement or promotion is based on demonstrated accomplishment and scholarly impact. “Merit” is believed to be associated with innate intellectual ability, dedication to learning and knowledge generation, mastery of a field of study, and recognition by others of comparable training and academic standing. Evaluations of accomplishment are dutifully (and often wishfully) believed to be wholly objective despite an abundance of evidence to the contrary. Unfortunately, implicit bias and other barriers to inclusion are pervasive within meritocracies. For members of marginalized groups, their social identity may diminish how their accomplishments are perceived and valued; conversely, the accomplishments of those with privileged identities may be over-valued. Moreover, what counts as “valuable” is itself not objective or neutral but rather reflects socially-constructed and culturally-specific priorities. Because academic merit and reward systems, as well as local cultures, can intentionally as well as unintentionally reinforce and hence perpetuate bias and barriers to inclusion, one of our UC Davis ADVANCE initiatives centered on review of all policies and practices affecting faculty advancement. We appraised the potential for bias in hiring, promotion, progression, and retention of faculty. We also evaluated the importance of culture in replicating barriers to inclusion.

L. F. Bisson (✉)

Department of Viticulture and Enology, University of California, Davis, Davis, CA 95161, USA
e-mail: lfbisson@ucdavis.edu

P. H. Kass

School of Veterinary Medicine, University of California, Davis, Davis, CA 95161, USA
e-mail: phkass@ucdavis.edu

K. T. Paw U

Department of Land, Air and Water Resources, University of California, Davis, Davis, CA 95161, USA
e-mail: ktpawu@ucdavis.edu

L. Grindstaff

Department of Sociology, University of California, Davis, Davis, CA 95161, USA
e-mail: lagrindstaff@ucdavis.edu

© The Author(s) 2022

L. F. Bisson et al. (eds.), *Uprooting Bias in the Academy*,
https://doi.org/10.1007/978-3-030-85668-7_4

Keywords Meritocracy · Higher education · Bias · Institutional bias · Hiring · Review of faculty

1 Assessing the Potential for Institutional Bias

Any evaluative process can reflect systemic bias and inequity. This is particularly true for assessments of accomplishment and merit. Perceptions of “quality” are often subjective. Belonging to a privileged group can lead to the unconscious phenomena of the “halo” and “Matthew” effects. First coined in 1920 by Thorndike, the “halo” effect refers to the tendency to associate a positive initial impression with favorable assessment of other unrelated traits in the evaluation of merit or accomplishment—the notion that the “good get better.” For example, Thorndike (1920) noticed the strong trend to link physical attractiveness to positive assessments for other qualities like intelligence and leadership, even when the latter were not supported by evidence. A related phenomena occurs when first impressions influence “objective” evaluation of merit (Liao, 2020) or the awarding of disproportional credit (Merton, 1968). This is known as the Matthew effect—“the rich get richer.” In the sciences, this is seen in the evaluation and rewarding of grants (Liao, 2020), and in the assessment of relative merit in collaborative research and other activities (Merton, 1968). The Matthew effect leads to a cumulative advantage as academic capital begets academic capital (Merton, 1988). These twin phenomena come into play during seemingly objective assessments. In practical terms, people from historically underrepresented groups may be seen as playing a secondary role in their own creative activities.

One of the programs of the UC Davis ADVANCE grant was the formation of the Policy and Practices Review Initiative (PPRI). The PPRI’s goal was a thorough assessment of practices surrounding faculty hiring and promotion to determine the points at which systemic bias, inequity, or inequality might be operating. Processes for faculty appointment and advancement are described in the Academic Personnel Manual (APM), which memorializes systemwide policies used across the University of California’s ten campuses. Individual campuses also have additional local procedures described in policy, (for example, APM UCD-210 and 220). Changes to either local or systemwide policies require review by the UC Davis Division of the Academic Senate. This codification provides transparency of overall expectations of faculty. In addition, many colleges and departments on any given campus have their own practices or procedures that vary slightly, but must remain consistent with both the systemwide and local policies. At UC Davis, changes to either local or systemwide policies require review by the UC Davis Division of the Academic Senate. This codification makes transparent overall expectations of faculty.

Before undertaking their work, members of the PPRI conducted an extensive review of practices and policies at other comparative institutions. We also participated in campus implicit bias training, because interpretation and application of policies can themselves introduce bias. We collectively read and discussed the National Science Foundation report on bias, and reviewed other publications regarding bias

in academia (Stewart & Valian, 2018; National Academy of Sciences, National Academy of Engineering & Institute of Medicine, 2007; Valian, 1998). The more we learned about the nature of barriers to inclusion, and especially about the neurobiology of bias, the more we realized that it is impossible to create policies that absolutely preclude bias. Rather, the subjective dimensions of seemingly objective assessments need to be acknowledged, understood, and limited in influence.

To assess the potential for bias, we created a comprehensive master grid that categorized policies by type of action—recruitment, advancement, or retention, plus workplace climate. We evaluated policies with an eye to determining whether they presented a barrier to inclusion. We also evaluated policies for the potential insurgence of implicit bias and other unconscious forms of prejudice that might negatively affect faculty advancement.

In addition to assessing existing policy, we evaluated faculty perceptions of hiring and advancement processes by holding a variety of meetings, including workshops, interactive roundtables, focus groups, public symposia, and one-on-one consultations. We also evaluated existing faculty surveys to further assess perceptions of bias in faculty advancement: the Collaborative on Academic Careers in Higher Education (COACHE) Faculty Job Satisfaction Survey, the UC Campus Climate Survey, and the UC Davis New Faculty Survey. We shared findings and held discussions with the appropriate Academic Senate Committees, and we hosted systemwide roundtables to facilitate discussion of findings and observations.

Armed with this array of information, the PPRI created a draft set of formal recommendations. These recommendations underscore how bias operates even in institutions perceived to be objectively meritocratic. Bias is learned but often unconscious, which makes understanding the processes that enable it all the more important. Bias can arise at any stage of evaluating (and consequently rewarding) merit.

2 Challenges in Assessing Policy

Determining whether a specific policy or its wording sustains bias or reinforces institutional exclusion is difficult if not impossible—the very nature of implicit or unconscious bias makes it difficult to guard against in policy formulations. How do you foreclose what is unconscious? Compounding the challenge are the criteria guiding faculty hiring and advancement, which can be vague, subjective, and open-ended. Consider “excellence”—the foundation of academic success. What is being valued under the rubric of excellence? How do you quantify/measure it? Who decides which achievements constitute excellence? When is an “excellent” performance excellent enough to warrant reward? The concept of excellence varies by field and is influenced by both local and societal-level definitions of achievement; moreover, recognition of excellence can be modified by the perceptions of others. Methods for quantifying excellence, such as citation-based measures, publication records, and external letters

of evaluation are potential vehicles for activating implicit bias. It is important, therefore, to understand the social contexts in which the concept of excellence is invoked and how it is operationalized. Transparency and consistency are essential here.

Regents special orders SP1 and SP2 (passed in 1995 and rescinded six years later) required that state university admissions and hiring processes depend heavily on quantitative assessments of merit—for example, test scores and grade point averages—under the assumption that such metrics were more equitable because more objective and less prone to bias. However, decreased diversity in the wake of these measures indicates the problems associated with trying to codify complex social issues. “Objective” measures that rely on access to resources (and therefore reproduce privilege) can in fact increase bias and barriers to inclusion by ensuring that privilege is a criterion for achievement.

2.1 *Is It Bias or Choice?*

The belief that merit assessment is objective posed a challenge during our review of institutional policies and practices. This, in combination with the unconscious nature of implicit bias, makes it difficult for evaluators to recognize, much less acknowledge, the subjective quality of their own assessments. Data clearly show the under-representation of specific identity categories throughout academia—for example, women and men of color. But is that under-representation due to marginalization and exclusion and/or being dissuaded from pursuing academic careers, or is it a matter of choice? Are members of these groups simply more likely to choose different, preferred career paths?

Such questions have plagued academia for more than half a century, often with regard to gender (Bernard, 1964; Rossi, 1965). For example, do women with children avoid academic careers simply because they prefer other work or because a family-unfriendly academic culture subscribes to an “ideal worker norm”? (As discussed in the Chapter, ‘[Barriers to Inclusion: Social Roots and Current Concerns](#)’, this norm assumes an employee unburdened by family obligations and completely devoted to paid work—a norm with a clear gender bias, given that women are disproportionately responsible for domestic labor and rarely have the luxury of being so single-mindedly focused). Arguably, the gender bias of the ideal worker norm is more explicit than implicit; what is implicit is the unconscious or subconscious accompanying belief that men are more dedicated to work than women because dedication is a masculine quality.

Teaching faculty about implicit bias and how it works, relying on data-driven research about bias, is critical for challenging the assumption that academic policies and practices are necessarily objective and meritocratic. As the faculty we consulted learned more about barriers to inclusion (including implicit bias) and their impact on hiring and advancement decisions, they increasingly recognized the limitations of seeing racial and ethnic homogeneity in the workplace as matter of choice. This is consistent with the findings and recommendations of Stewart and Valian (2018).

We therefore cannot emphasize enough the importance of data-driven training in addressing implicit bias, and the necessity for the consistent use and improvement of institutional data gathering so this kind of research can be conducted.

2.2 Policy Interpretation

Policy implementation has four components. First, policies have to be created and written down. Second, the intent of that wording has to be interpreted. Third, the policy has to be put into practice or implemented. Fourth, it has to be received/experienced by real people in real settings. In academia, often the various tasks associated with policy creation, interpretation, implementation, and reception are undertaken by different individuals representing different constituencies. As a result, the policy may not work well or consistently or as intended. For example, in a discussion we had with postdoctoral scholars about work-life integration and career choice, participants pointed out that institutional policy discourages child-bearing during the postdoctoral years. We were surprised to hear this, as we weren't aware any policy with that intent. As evidence, they cited the policy that discouraged part-time appointments for postdoctoral fellows:

Postdoctoral scholar may request appointments for less than 100% time (total of all appointments) for one of three reasons: health, family responsibilities or external employment. The postdoctoral scholar and the PI (Principal Investigator) must complete a memorandum of understanding regarding the responsibilities and duties to be expected.

—<https://grad.ucdavis.edu/postdoctoral/appointment-promotion>

This policy was created to ensure that part-time appointments did not have the same expectations as full-time appointments. However, postdoctoral scholars who wanted to start families saw the need to secure formal administrative approval to shift to part-time status as problematic. They felt the policy served to document a lack of commitment to their career, which was never its intent. A challenge in any policy assessment is to fully understand the ways in which a policy's wording is interpreted, which may or may not match the intended meaning. A given interpretation may be "incorrect" from the point of view of the policy's makers, but even if the "true" intent is made known to those targeted by the policy, the policy can have negative effects or be perceived as having negative effects.

2.3 Influence of Local Culture

Another challenge with policy assessment arises from differences in interpretation and application by the entities charged with enacting the policy. In academia, departments are responsible for faculty review in hiring and advancement. Departments

are organized around the delivery of a curriculum and are generally limited to sub-disciplines. Departments develop their own local culture—the synthesis of values of the institution and the discipline, along with the shared principles and goals of the department faculty, students, and staff. With respect to policy implementation, local cultures may differ from one another: in interpretation, in the level of transparency of review and criteria for advancement, in the degree of faculty consensus before voting on appointment or advancement, and in the nature of the department’s overall climate.

2.3.1 Interpretation

The local culture of a unit affects how policies are interpreted and implemented. Policies, particularly those associated with faculty evaluation and advancement, deliberately allow for flexibility in interpretation to enable the wide diversity of academic disciplines in a research-intensive university such as UCD. The language for assessing intellectual merit for the purposes of hiring and promoting faculty is as follows:

Superior intellectual attainment, as evidenced both in teaching and in research or other creative achievement, is an indispensable qualification for appointment or promotion to tenure positions.

What constitutes “superior intellectual attainment” will vary by field, and thus is open to interpretation by local academic units. How is “superior” defined? Who has ultimate authority for determining if superior attainment has been achieved? This is a qualitative assessment in which phenomena like the halo and Matthew effects may play a role. In some fields, intellectual attainment is assessed solely in the context of the individual’s accomplishments, and collaborative efforts are not highly valued, while in other fields, collaborative achievements are recognized as equally valuable. The required use of external evaluators in promotion actions can serve to anchor local unit interpretations within that of the field, but external evaluation may involve individuals who have not gone through the same implicit bias training as UCD faculty. Review of actions by centralized or campus-wide faculty peer committees can also ensure that local interpretations of intellectual attainment are in line with those of the general campus. The absence of such practices can lead to divergence across units in what constitutes “merit.” University policy also allows flexibility in valuation of the four areas of faculty work: (1) teaching, (2) research and other creative work, (3) professional competence and activity, and (4) university and public service, with a leitmotif of reward for diversifying activities in these areas. Local units can disagree, for example, on whether one’s service and diversifying activities should be recognized as “superior intellectual attainment.” In our discussion with groups of faculty, several individuals pointed out they deliberately limited their discussion of service in advancement materials. They cited activities such as mentoring of junior colleagues and support of student activities and groups outside the classroom

as counting against them in evaluation. For them, intellectual attainment appeared easier to quantify and demonstrate in the other three areas of evaluation.

2.3.2 Transparency

Local cultures may also differ in the level of transparency about judgments of “superior intellectual attainment.” Academic Senate Bylaw 55 (<https://senate.universityofcalifornia.edu/bylaws-regulations/bylaws/blpart1.html>) allows departments to limit participation in advancement actions to individuals at the same rank or above that of the faculty being evaluated. In practice, this means that assistant professors may or may not see what a successful case for promotional advancement looks like, as this would entail seeing the records of their more senior colleagues and participating in the discussion of what aspects are valued, as well as how policies are interpreted. This can be especially challenging with regard to service and diversity activities because the criteria for advancement in these areas are less clear than for research. Changing department rules on voting procedures in such cases can only be modified by a two-thirds vote of a department, thereby preferentially empowering the more senior faculty, potentially at the expense of more junior faculty. Faculty surveys show divergent views about the preferred level of transparency surrounding advancement actions within their own units.

2.3.3 Climate

The climate of a local culture, meaning the collective values, perspectives, and inclinations of the members of the unit, can also affect policy application. For example, UC Davis, like many other institutions of higher education throughout the country, has a policy of allowing extensions to the tenure clock for the birth or adoption of children. When we started our work, existing policy required new parents to request an extension of the tenure clock via their department chairs, which was then approved by the Vice Provost for Academic Affairs. Our data indicated, however, that many eligible faculty members made no extension request. At a meeting of department chairs we asked how many encouraged their faculty to take advantage of extending-the-clock policies. Roughly half said they either provided no advice or actively discouraged taking advantage of the extension.

Initially we thought these departments were not supportive of work/life balance policies, however, chairs who discouraged their faculty from extending the clock pointed out that, in their disciplines, the competition for external funding is so intense that any slowdown in one’s publication record reduces competitiveness and leads to long-term negative effects on career advancement. Recognizing the need to maintain competitiveness while also rearing young children, these chairs described department cultures that focused on integrating work and life without recourse to the formal extend-the-clock policy. In follow-up discussions, faculty in those units described a suite of practices, attitudes, and accommodations that enabled work and family

integration and thus that sustained productivity in all four areas of assessment. For example, faculty highlighted the following: more flexible work schedules, the selection of child-friendly spaces for meetings, child-care during meetings, the option of online meetings, flexibility in online versus in-person teaching, being able to care for infants in an office, and avoiding scheduling meetings or receptions during times critical to picking up children from daycare. In effect, such units developed a work-around to the university-wide policy that accomplished much the same goal. This is not to suggest the original policy was unnecessary or ineffective, but that local field- or department-specific cultures affected whether and how it was used.

2.4 External Drivers of Policy Interpretation

If winning outside funding to maintain competitiveness is indeed an expected metric of “superior intellectual attainment,” this raises the issue of external factors affecting policy interpretation. Even in fields where extramural support is not the norm, external pressure for creative output—articles, books, performances and exhibitions, etc.—to maintain competitiveness can be intense. Are the requirements for advancement consistent with the resources already provided or with those that may be available? Teaching is another potential factor; teaching loads are driven by student enrollment numbers, and the relative value of teaching versus research may be affected by time dedicated to teaching. In our case, this was particularly evident when we evaluated letters written by external reviewers that downgraded the level of research or creative activity if teaching loads were not similarly high at their own institution. In other words, if the teaching load of the faculty member being evaluated was lighter than that of the colleague writing the letter, the latter expected comparatively higher research/creative output. We also read letters that faulted a candidate for “too much teaching” and, oddly, what was called “too much effort in teaching.” Moreover, the local culture of the external reviewer may influence how they perceive the candidate’s accomplishments, as can potential implicit and explicit bias on the part of the reviewer.

2.5 Practices Outside of Policy

The existence of practices outside of policy and approved procedures poses another challenge in policy assessment. For example, the evaluation procedures for faculty advancement are clearly described in both the system-wide and campus-specific Academic Personnel Manual (APM) documents, but in the late 1990s, faculty complained that the UC Davis Committee on Academic Personnel (CAP) was including evaluative criteria outside of policy. A task force was then formed to review these complaints. It learned that a small number of departments had developed a practice of encouraging individual faculty to communicate directly with members

of CAP to influence the advancement of a colleague. Although the practice was not widespread and although it was not clear that such communications impacted CAP's recommendations, policies governing individual communication with CAP were subsequently implemented.

The task force identified an additional problem: review committees were sometimes including in their evaluation process external information outside of the official dossier for some faculty and not others. This occurred at different levels of review—department, college, and CAP—and is problematic on two fronts. First, using information beyond the official dossiers is a violation of policy; second, inconsistent application of any practice (whether sanctioned or not) is a likely vehicle for introducing or amplifying bias because inconsistencies by definition create an uneven playing field, potentially helping (or hindering) some individuals over others. If the inconsistency is consistently applied—for example, if the admission of information outside the dossier routinely benefits or disadvantages one group over others—then it becomes discriminatory.

3 Bias in Assessment of Merit

The concept of “superior intellectual attainment” is not amenable to clear, transparent metrics; “superior” is a qualitative judgment subject to the influence of bias. Moody (2012) outlines a number of cognitive errors operating in assessments of faculty that, she argues, compromise/contaminate academic evaluations. They are not so much “errors” as outcomes of normal brain function; nevertheless, they impact evaluation. Because they are implicit (that is, subconscious or unconscious rather than conscious), we prefer to characterize errors of the sort identified by Moody as “cognitive bad habits.” They include the halo and Matthew effects, among other phenomena. The impact of bad habits leads to bias in evaluation and, in turn, can lead to inequity in assessment of merit (Fig. 1).

Among us, we have had extensive experience on faculty search committees and review committees, across multiple levels of the academic hierarchy. Inspired by Moody, below we summarize ten key cognitive habits that we observed on our own campus and that we believe have considerable potential to affect faculty evaluation.

Cloning: Cloning occurs when evaluators, in considering a pool of candidates, conflate “excellence” with the qualities they themselves possess. In other words, they unconsciously see excellence when they see a version of themselves. Operationally, cloning may be manifest as a preference on the part of evaluators for candidates trained at the same prior institution, conducting similar research, or possessing a shared social location or identity. A corollary of this cognitive habit is “provincialism”—devaluing those with whom the evaluator does not identify. We see this often in “pathway bias,” meaning devaluing faculty candidates that may have started academic careers at community colleges rather than four-year research institutions, or

Fig. 1 Inequity in assessment of merit. Illustration by Chastine Leora Madla

at “regular” (including public) research institutions rather than Ivy League universities. Here, evaluators can conclude erroneously that such markers indicate a candidate is unqualified simply because he or she followed a different path to the Ph.D.

First Impressions: Evaluators may reach conclusions rapidly about candidates based on insufficient evidence. This can happen when external markers are taken as proxies for quality, as in the halo effect (Thorndike, 1920). Familiarity with the candidate’s institution, the reputation of a dissertation advisor, or reputations of letter writers may lead to evaluators to conclude “superior intellectual attainment” regardless of the actual record of research or creative activity.

God Fit/Bad Fit: Judgments of “fit” with local culture are often subjective. Unconscious bias can play a strong role here, particularly when assessing whether someone from an underrepresented and/or marginalized group will “fit in” with the rest of the department. In this case, the “bad fit” may also be a function of the assumption that the person lacks shared values and attributes. Alternately a candidate may be viewed as being a “good fit” if a shared identity or set of values is perceived. “Fit” often cannot be determined from reading an application, yet, as Heilman et al. (2015) also found, we have seen this concept used as a criterion for selecting candidates to interview. Of course, certain dimensions of “fit” are partially articulated a priori in position descriptions. Even so, departments occasionally recommend an applicant because

the faculty feel he (or less often she) will make a “better” colleague than others in the pool (that is, he resembles the majority identity of the department) despite the fact that the person’s expertise diverges significantly from the position description. Here we see one aspect of “fit” (local departmental culture) outweighing another aspect of “fit” (the actual position description).

Stereotyping: The extent to which identities are reduced to stereotypes can affect seemingly objective judgments. The stereotype effect may be either positive or negative, and thus, for example, the candidate may be presumed competent or incompetent, a team player or a trouble-maker, a leader or a non-leader. Assumptions of incompetence especially plague women of color faculty (Gutierrez et al., 2012), and can appear in evaluations when similar accomplishments or work conducted receive differential rewards or levels of credit. The need to avoid stereotyping is critical when evaluating collaborative or team-based efforts. Here, competence and leadership may be entwined: if the individual (e.g., a Latina physicist) belongs to identity categories perceived as incompetent or unlikely to produce a leader, an evaluator may inappropriately attribute credit for an accomplishment to others on the team. By contrast, if the person (e.g., a white male engineer) belongs to identity categories that are “presumed competent,” he may be given undue credit for collaborative works even if he himself describes a minor role. In effect, the stereotype of (in)competence reinforces a double standard.

Specific actions—speaking with a non-English accent, wearing “ethnic” attire—can also trigger stereotypical thinking, which may, in turn, affect evaluation of merit despite the fact that diction and appearance are unrelated to research. When stereotypes shape thinking and behavior, they may manifest unconsciously as implicit bias or more overtly as microaggressions (discussed in the Chapters, ‘[Barriers to Inclusion: Social Roots and Current Concerns](#)’, and ‘[Making Visible the Invisible: Studying Latina STEM Scholars](#)’).

Elitism: Elitism reflects the belief that people who are traditionally found at the top of a hierarchy (e.g., white men in STEM) are occupying their rightful place, and that, by implication, those below them are also where they belong in the hierarchy. One consequence is that the bar for gaining entry at the top may be higher for perceived “outsiders” such women and men of color. Members of these identity categories may have to work harder and perform better just to get admitted into the tier and, once there, they are more heavily scrutinized, particularly if, as discussed in the Chapter ‘[Barriers to Inclusion: Social Roots and Current Concerns](#)’, they are token representatives of “their” group.

Raising the Bar: Related to elitism, raising the bar means creating a higher standard for perceived outsiders in order to reconcile their inclusion into the group. To raise the bar is to enact a double standard. Let’s say the desired group is “tenured faculty in the Philosophy Department” and 80% of the department faculty are white men. During evaluation, if a woman of color candidate excels in all the listed criteria, there may be an unconscious tendency to downplay her strengths and subtly “raise the bar” to diminish the competitiveness or the impact of her accomplishment. This

cognitive habit is also manifest when evaluative criteria are changed or realigned in ways that favor some candidates over others. Moody (2012) refers to this type of cognitive error as “relying on pretext” to favor or disfavor a candidate.

Bias Validation: Unconscious bias is sometimes rationalized in conscious decision-making. For example, bias may be validated when members of dominant groups are evaluated on the basis of “potential” and members of marginalized groups are assessed solely on track records or accomplishments. The assignment of personal “potential” can serve to elevate one’s ranking. “Star status” is another type of non-quantifiable judgement validation in assessing merit, and is often associated with insider knowledge on the part of an evaluator who “knows” the candidate’s lab, institution, or advisor, or who has directly contacted colleagues who confirm that the candidate is in fact better than his or her record. This cognitive bad habit expresses the Matthew effect (Merton, 1968) and is what Moody (2012) characterizes as prioritizing rhetoric over evidence.

Accounting for “Work Not Done”: Bias can also appear in the form of evaluating work not done; by this we mean situations in which a candidate’s performance is negatively assessed because actual productivity is being compared to a hypothetical *possible* productivity based on implicit beliefs about dedication to work. As already discussed in the Chapter, ‘[Barriers to Inclusion: Social Roots and Current Concerns](#)’, dedication to work as an index of success arose during the era of single-career couples in which, in the typical case of heterosexual marriage, husbands were breadwinners and wives were homemakers. Dedicated breadwinners put career ahead of family; sacrificing family time demonstrated strong commitment to one’s profession. Dual-career couples are common in academia today and partners often share responsibility for breadwinning and homemaking (albeit not necessarily equally). Implicit adherence to the concept of dedication to work as a measure of achievement negatively impacts individuals who display commitment to parenting if evaluators use “work not done” as a criterion for advancement. For example, in one promotion case, faculty who voted against the promotion cited as justification the fact that the candidate coached his own daughter’s soccer team, and that if he had devoted the time instead to his research/scholarship he would have had a “better” publication record. Because dedication to work is a masculinized concept and commitment to family is a feminized one, men who commit to family may be penalized more than their female counterparts just as women’s dedication to work may be less rewarded than men’s.

Defining and Quantifying Merit: This is the elephant in the room—bias is embedded in the very definition and quantification of merit. Many concepts and measures of merit originate from within dominant social groups to reward forms of achievement these groups value; moreover, they are sometimes inconsistently applied. This means women and URM scholars may be disadvantaged when judged by measures developed in the context of historically exclusive (white, male, upper-class) academic institutions. How can such measures be expanded and improved? The variety of academic programs at UCD and other research universities further

complicates definitions and quantification of merit; obviously the number of publications may be a completely inadequate measure of merit for a musician, whereas the number and importance of the venues of performances may be an inappropriate measure for a chemist—but even beyond this, are publications or performances the best or most important metrics in their respective disciplines? Another degree of complexity emerges when evaluators use criteria based on individual accomplishment in an era where many fields are becoming strongly collaborative. Adherence to traditional notions of merit represents one potential institutional bias that is proving difficult to change in any substantive way.

Bias Beyond Merit Assessment: Our review of policy focused on the merit and advancement process. However, bias can affect review in multiple areas of faculty life. Consider the importance of “discovery” to the scientific enterprise and to one’s claim to be a “good” scientist. Who comes to mind when we think about scientific discovery? People from marginalized groups—particularly women of color—are not typically envisioned as discoverers. Their discoveries are often deemed provisional until confirmed by dominant-group member, and their discoveries/contributions may be ascribed to others. Lynn Conway (2018) refers to this phenomenon as “the disappearance of the others.” Problems with accurate external attribution of contribution negatively impact faculty review as well as publication and grant review. Implicit “discoverer bias” can lead to a higher bar for proving novelty or merit and demonstrating ability or potential for success. Such issues are cumulative over a career and challenging to address (Conway, 2018).

4 Lessons Learned

The recognition that academic institutions must re-evaluate their evaluative criteria is not new; long after problems with traditional metrics were first identified, they persist; progress has been slow and incremental. We learned several key lessons in the process of evaluating policy. We offer these in the form of advice for those considering a similar institutional self-assessment.

4.1 *The Importance of Organizational Learning*

Knowledge of the bases of bias and discrimination is an essential first step to creating an inclusive academy. The acquisition of knowledge has to be wide-spread; ideally, all members of an organization from top to bottom should learn about barriers to inclusion and how they can affect faculty advancement. Only by understanding the roots and nature of these barriers can we begin to understand their impact and how to effect institutional change, collectively and collaboratively. Although many scholars outside of STEM have studied the causes and consequences of social inequality

for decades (indeed, centuries), scientists tend to believe that their own disciplines are insulated from these forces. STEM faculty tend to view themselves as “identity-blind” (for example, color-blind, or gender-blind), but the unconscious brain both sees and processes color, gender, and other identity categories. Consequently, educating all members of an organization must start with acknowledging there is a problem: existing criteria for evaluating merit and achievement are not objective because bias is real and its negative effects are not equally distributed. Beyond this, we must acknowledge that what scientists choose to study, how they study it, and who they collaborate with all present barriers (or, conversely, opportunities) for cultivating diversity, equity, and inclusion in science.

Toward that end, we created training via the Strength Through Equity and Diversity (STEAD) Committee, drawing on resources provided by the University of Wisconsin and the University of Michigan. The STEAD Committee is comprised of UC Davis faculty members who, first, educated themselves about implicit bias and other barriers to inclusion and, second, widely presented workshops to other faculty and to administrators about best practices for achieving excellence, equity, and diversity in faculty recruitment. The committee chose recruitment as the focal point for education because recruitment processes make salient the ways in which problematic subjective assessments can compete with more equitable, objective ones. The STEAD trainings address the following issues, among others: the composition of the search committees; the wording of position announcements; resources for broadening the applicant pool; strategies for widespread advertising coupled with more targeted outreach aimed at organizations, associations, and conferences serving underrepresented faculty; discussions of the language used in letters of recommendation; checklists for reviewing and evaluating applicants that employ a consistent set of parameters; and, most important, sharing the results of data-driven research about the ways in which bias can creep into every stage of the recruitment process. For example, do letter writers refer to female candidates by their first names and male candidates by the last names? Are the letters for white men longer than for white women and people of color? Are personal characteristics and qualities discussed in letters for female and minority candidates but not for white male candidates? Are members of the search committee using the prestige of the degree-granted institution as a proxy for “promise” or “potential”? Surveys of participants indicate that the workshops are highly valued and have positively affected candidate review.

4.2 The Importance of Diversity in Policy Creation and Assessment

Bias associated with membership in an identity category favors dominant social groups. It is critical to recognize this fact. Some of us (those in the STEM fields) began our work thinking we needed policies that were “identity blind”—that is, policies where identity categories would not be seen; we now agree this is impossible. If

representatives of a historically-dominant group prevail in an institution, that group's priorities will likely become the priorities expected of all members of the group, to the detriment of "outsiders." In other words, because we craft, interpret, and apply policy through the lenses of our own identity and experience, policies created by members of dominant groups (with their own priorities in mind) may inadvertently establish barriers to inclusion everyone else. A policy may be discriminatory not because that was the original intent, but because it is the natural outcome of a process that lacks sufficient diversity of perspective. Moreover, under-represented scholars do not benefit from policies that fail to recognize their social identities, because that often means suppressing an aspect of the self in order to be seen as part of the broader collective. Policies thus need to be "identity independent"—neither favoring nor disfavoring any group—but not blind to an identity if the goal is genuine inclusion. Broad consultation as well as diversity of community membership are therefore vital for policy creation and review.

4.3 Empowerment Through, Not Despite, Policy

The concept of unconscious bias is difficult for many people to accept, especially faculty in STEM fields, and so, unsurprisingly, it's difficult to design policies to recognize and prevent it. Moving up in a "meritocratic" academic hierarchy has been described by Lynn Conway as climbing a ladder to the past (personal communication). Old notions of achievement and dedication to work persist, and we continue to hold faculty accountable to them, despite empirical evidence they perpetuate inequality and do not reflect the reality of many if not most of our lives. Currently, the majority of academic STEM faculty are in dual-career relationships, and the expectation of overwork is a key reason why scholars peremptorily exit STEM fields—a phenomenon known as the "leaky pipeline" (Xie & Shauman, 2003). Although faculty as a whole recognize this problem, they have few tools to change the culture in ways that will enable alternative, more inclusive measures of "superior intellectual attainment," meaning measures that do not rely on outdated, gendered assumptions about work. One of our biggest challenges is to recognize the links between definitions of excellence, merit, organizational culture, and structures of social inequality so that when the world changes around us, we can change with it.

5 Key Recommendations

In evaluating merit and advancement actions, review committees at UC Davis, as at most institutions of higher learning, have been instructed to caution against bias. But until recently, no training in recognizing and avoiding bias was stipulated on our campus. Consequently, the Vice Provost for Academic Affairs initiated mandatory implicit bias training for all search committee chairs; this training was then replaced

by peer-to-peer learning under the Strength Through Equity and Diversity (STEAD) initiative of the UC Davis ADVANCE program. Surveys of faculty indicated that most agreed with our assessment that recruitment and advancement procedures and processes were clear and transparent, but not necessarily identity-independent as intended, because of implicit bias and other barriers to inclusion embedded in their local cultures. Our analysis of policy and practices thus yields several recommendations that focus on faculty development and support along with changing the climate of local cultures.

5.1 Checking Parental Bias

Although most policies as worded are seemingly identity-neutral, we concluded that the language associated with the parental extension to the tenure clock policy was potentially problematic.

UC APM 210-1, Section c.(4), Assessment of Evidence, includes the following statements (emphasis added):

If there is evidence of sufficient achievement in a time frame that is extended due to a family accommodation as defined in APM—760, *the evidence should be treated procedurally in the same manner as evidence in personnel reviews conducted at the usual intervals*. The file shall be evaluated without prejudice as if the work were done in the **normal period** of service and so stated in the department chair's letter.

This language implies that such leaves are both unusual and abnormal. Starting the first sentence with “if” also suggests that a more stringent review be done of the evidence presented. However, when we proposed changes to the language, not everyone agreed this was necessary; some felt the more serious issue was that the language was repeated in letters sent to external reviewers who might not share our same cultural context. As a result, the wording of those letters was changed:

UC Davis encourages its faculty members to consider extensions of the (pre-tenure/review) period under circumstances that could interfere significantly with development of the qualifications necessary for (tenure/advancement). Examples of such circumstances may include birth or adoption of a child, extended illness, care of an ill family member, significant alterations in appointment. Please note that under this policy the overall record of productivity and scholarly attainment forms the basis of your evaluation. Time since appointment is not a factor in this review.

A second issue was that the parental leave policy had to be requested and approved, rather than being automatically granted. Again, as soon as we suggested that extensions to the clock be automatic, the policy was in fact changed throughout the UC system. Making extensions automatic both implied they were normative (thus removing any potential stigma implied by being granted “special” dispensation) but did not prevent or penalize faculty from opting out.

5.2 Gender Bias in Negotiation Is Real

No set policies govern negotiations of start-up employment packages for new faculty. We discovered that the practices varied by units and that women applicants often felt caught between asking for what they needed and triggering negative stereotypes of professional women as “demanding.” Instead, many women asked for what they thought the negotiator would see as reasonable so as not to appear “greedy.” We concluded that it would be better for an intermediary to assume responsibility for negotiations, a practice that had been adopted in some departments. In most cases, the principal negotiator on behalf of the new faculty member has been the department chair, but occasionally other faculty serve informally as intermediaries by encouraging candidates to negotiate competitive start-up packages.

5.3 Self-Promotion Versus Bragging

The faculty merit process requires faculty to write a “candidate’s statement” that highlights their accomplishments over the review period and assess their own impact as scholars. There are two pitfalls to this process. First, external reviewers may view self-promotion differently depending on the candidate’s social identity category. In STEM fields, as pointed out by Lynn Conway (2018), white men are deemed, unconsciously, as natural “discoverers”. When others claim a discovery, the claim may be viewed with skepticism, and their contributions may be diminished as a consequence; moreover, evaluators can view their attempts to take credit negatively.

Second, and related to this, acts of self-promotion may be seen as “bragging.” This concern is not gender-neutral because of stereotypical assumptions that women should be caring, compassionate, and self-effacing, not self-promotional. Whether or not self-promotion is comfortable for a candidate may also be culturally-specific—Asian women, for example, may feel pressure to conform to cultural stereotypes that they are especially docile and accommodating. Concerns about “bragging” can work in the reverse, too. White male scholars sensitive to being stereotyped as “jerks” who are “naturally” inclined to brag may avoid self-promotion in an effort to counter the stereotype. One solution here is for department chairs and other faculty to clearly highlight the achievements of a candidate in ways that are consistent with but go beyond the candidate’s own self-narrative, so that the responsibility for promoting achievements are dispersed across multiple actors and levels of authority. Chairs and other colleagues can also draw attention to achievements and contributions that a candidate him or herself may gloss over, thereby strengthening a case. Online ballots with space for faculty commentary enable chairs to directly incorporate the promotional language of colleagues in order to further lessen the burden of promoting the case on the individual candidate. Of course, these interventions depend on chairs and their faculty understanding the existing social biases that distinguish promotion from bragging.

5.4 Rethinking “Superior Intellectual Attainment”

Operationally, we found that evidence of “superior intellectual attainment” was defined differently by various local cultures. In some units it was limited exclusively to research. In those cases, the prevailing culture did not view teaching, professional activity and competence, or service as relevant to the review process; rewarding faculty for contributions to diversity and mentoring was even seen as incompatible with superior intellectual attainment. To encourage recognition of teaching and service, the UC Davis campus changed its existing process to a “Step Plus” system of academic advancement. Faculty were explicitly instructed to consider outstanding attainment in each of *four* categories—research, teaching, service, and diversity activities—during the process of evaluating candidates for advancement. Although relatively new, some faculty say this altered review process has indeed enabled greater reward for non-research activities. Prior to the implementation of the step system and helping to pave the way, policies in the APM made “contributions to diversity” a criterion for faculty advancement.

5.5 Accountability for Changing Local Culture

According to University of California policy, accountability for the local (that is, department) climate rests with department chairs. However, at the time we began our assessments of policy, scant resources were available for chairs to learn how to assess current levels of inclusivity and how to develop strategies to make departmental climates more inclusive. We recommended that chairs be equipped with the tools needed to be effectively accountable—workshops, new chair orientations, online reference material, suggestions for best practices, and experts available for consultation. Creating inclusive local climates is a concern not just for UC Davis but for all ten campuses within the UC system; consequently, a suite of tools, training guides, and support methods has been developed. When assessed for their performance, chairs are expected to report on issues related to climate, although not all departments use these issues in their evaluation of chairs.

The mandatory implicit bias training for search committees actively recruiting faculty has served to create local expertise in identifying this and other barriers to inclusion at the department level. Bias training enables local discussions of inclusive climate within departments. We therefore recommended that all faculty receive training in recognizing implicit bias, a recommendation closely linked to unimplemented mandatory training recommendations discussed decades ago by the UC Davis Senate Committee on Affirmative Action and Diversity. Sexual harassment training is now also mandatory for all faculty and is done online. Although we did not explicitly request that harassment training include information about spotting and avoiding implicit bias, the course now does so. The general availability of knowledge about

barriers to inclusion will, ideally, help department chairs ensure a more inclusive identity climate for all faculty.

Local culture also shapes practices associated with mentoring. The UC Davis COACHE (Collaborative on Academic Careers in Higher Education) Faculty Job Satisfaction Survey from 2013 (a component of our ADVANCE program) indicated that faculty believed mentoring for junior colleagues was deficient on our campus. Subsequently, the ADVANCE program leadership recommended that each recruitment search plan include a strategy for mentoring new faculty members.

6 Concluding Thoughts

The social origins of implicit bias and its unconscious application together make crafting policies to prevent it challenging, if not impossible. We therefore recommend a two-pronged approach—making sure that policies and reward structures are as identity-independent as possible, while at the same time providing broad training and extensive awareness about the nature of barriers to inclusion. Understanding the role of cognitive “bad habits” and the influence of phenomena like the halo and Matthew effects on assessment can diminish the impact of those types of bias in assessment. It is important that the group of evaluators also be diverse as this has the potential to dilute the impact of cognitive bad habits in objective assessment. We acknowledge that UC Davis as an institution may be more open to adopting inclusive strategies than some other institutions. However, there is still considerable room for improvement and much work to be done.

Acknowledgements We thank Chastine Leora Madla for drawing Fig. 1.

References

- Bernard, J. (1964). *Academic women*. The Pennsylvania State University Press.
- Conway, L. (2018). The disappeared: Beyond winning and losing. *Computer*. <https://doi.org/10.1109/MC.2018.3971344>
- Gutierrez, G. y Muhs, Nieman, Y., Gonzalez, C., & Harris, A. (2012). *Presumed incompetent: The intersections of race and class for women in academia*. University Press of Colorado.
- Heilman, M. E., Mauzi, F., & Braun, S. (2015). Presumed incompetent: perceived pack of fit and gender bias in recruitment and selection. In M. Broadbridge & S. L. Fielden (Eds.), *Handbook of gendered careers in management*. Edward Elgar Publishing. <https://doi.org/10.4337/9781782547709.00014>
- Liao, C. H. (2020). The Matthew effect and the halo effect in research funding. *Journal of Informatics*, 15, 101–108. <https://doi.org/10.1016/j.joi.2020.101108>
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159, 56–63.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623.
- Moody, J. (2012). *Faculty diversity: Removing the barriers* (2nd ed.). Routledge.

- National Academy of Sciences, National Academy of Engineering & Institute of Medicine. (2007). *Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/11741>
- Rossi, A. S. (1965). Women in science: Why so few? *Science*, 148(3674), 1196–1202.
- Stewart, A., & Valian, V. (2018). *An inclusive academy: Achieving diversity and excellence*. MIT Press.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4, 25–29. <https://doi.org/10.1037/h0071663>
- Valian, V. (1998). *Why so slow? The advancement of women*. MIT Press.
- Xie, & Shauman. (2003). *Women in Science: Career processes and outcomes*. Cambridge: Harvard University Press.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.