AC 2009-644: RECOGNIZING, CHARACTERIZING, AND “UNSETTLING" UNINTENDED BIAS IN THE FACULTY SEARCH PROCESS IN ENGINEERING

Kristen Constant, Iowa State University
Kristen Constant is an Associate Professor in Materials Science and Engineering at Iowa State University.

Sharon Bird, Iowa State University
Sharon Bird is an Associate Professor in Sociology at Iowa State University
Recognizing, Characterizing and ‘Unsettling’ Unintended Bias in the Faculty Search Process in Engineering – A Case Study
Abstract:

Faculty searches are one of the most critical activities undertaken with regards to shaping the future of a department, college or university. Despite the importance of this activity, most search committees are comprised of faculty who have very little time to commit to this task, and little if any training on ‘best practices’ or policies for conducting searches. National recognition of the need in faculty search processes for decision-makers to understand the prevalence of unintended biases and how to combat them has grown precipitously in recent years. Cognitive scientists have shown that many of the selection and evaluation processes we undertake on a daily basis are alarmingly contaminated, despite our good intentions. The term ‘cognitive errors’ has been coined to describe such errors in judgment. The objectivity of these evaluations is further compromised when we are overburdened and distracted. Common beliefs about how women and men, people of color and white people, act, feel, and express themselves often unintentionally enter into our decision-making processes. This is true even among people who eschew sexism and racism—because beliefs about gender and about race are part of the social norms of our society. Few of us stop to consider how taken-for-granted beliefs and expectations might be affecting our decisions. It is in this manner that gender and race biases commonly impact faculty search processes. In turn, engineering departments continue to struggle with hiring highly qualified women faculty. However, there are strategies to minimize the impact of unintended bias on faculty search results. In this paper, we discuss two examples of faculty search processes in which biases were demonstrated both in conversations and evaluations within search committees and in letters of recommendation for candidates. The extensive research on cognitive errors is used to provide a systematic analysis of these examples and approaches to minimizing the impact of unintended bias through promoting awareness are presented.

Background:

Despite that most universities have stated goals of diversity and non-discrimination, it is well established that gender inequities remain among the university faculty, especially within engineering. Numerous studies have investigated the barriers encountered by women with aspirations of university careers, many referred to in “Beyond Bias and Barriers”, a 2006 report by the National Academy of Sciences. These studies seek to provide a deeper understanding of various issues including those pertaining to the career pipeline, and faculty recruitment, retention and advancement. Recognizing the critical need for full participation of women in the sciences, the National Science Foundation has supported for the last 7 years efforts to study and improve recruitment and retention of women faculty in the sciences through the ADVANCE Institutional Transformation program. Work within this program has allowed researchers to study multi-faceted approaches to increasing participation of women faculty at many institutions, including our institution. In this paper, we will examine the faculty search process with respect to the barriers confronting women. First, the context of the faculty searches will be presented, then the concept of cognitive errors introduced. (See also empirical studies in this area.) Examples of such cognitive errors will be analyzed and approaches to minimizing unintended bias discussed. Most of these approaches are designed to ‘unsettle’ (in the minds of faculty) that which has become accepted knowledge about the normative practices of the university. In other words, we are attempting to encourage faculty to
consciously evaluate the assumptions they make and the practices they engage in when evaluating faculty candidates.

Iowa State University of Science and Technology is a land grant institution with a 150-year history of strength in science and engineering. The university, with about 27,000 students and 1,700 faculty, has 8 colleges, the second largest of which is the college of engineering with a faculty of 190 and a student population of 5,600. National faculty data show that the fraction of women in engineering faculty positions has increased in the last 6 years from 9 to over 11%, though that remains dominated by assistant professors (almost 20% of all faculty), and associate professors (13%), rather than full professors (6%). However, in the College of Engineering at ISU, there is a smaller fraction (8.6%) of women faculty, recovering this year slightly after a 5 year decline. Additionally, the attrition rate for ISU women faculty in STEM fields (Science, Technology, Engineering and Mathematics) is significantly higher for women as compared to men (especially in the first three years). While it is important to monitor statistics on women faculty and their rates of retention and advancement to indicate success of women already on the faculty, it is equally as, if not more, important to carefully examine the search and hiring process so that we gain a clearer understanding of the actual practices that might contribute to or detract from the hiring of women faculty.

Many universities have handbooks or guides for search committees, as does ours, but existence does not ensure use. Few universities require formal training and even fewer have training that focuses on broadening the applicant pool and combating unintended bias. However, recently, a number of universities (many of them NSF ADVANCE institutions) have developed resources and implemented training for search committees.

Still, more often than not, when faculty search committees are named, a chair is appointed and the process begins with little or no preparation on the part of committee members for the task. Additionally, most often the task of serving on a search committee is assigned as an overload without relief from other responsibilities.

Context:

This report is derived from case studies of two separate search committees that were tasked with identifying and evaluating candidates for engineering faculty positions. One was a department-level search for an assistant professor position and the other was a college level search for multiple, targeted, interdisciplinary positions of unspecified rank. In each case, approximately 250 applicants were reviewed. The department committee comprised 6 faculty members over all three ranks, tenure track assistant through full professor, 5 men, and 1 woman. The college level committee was 11 faculty, all tenured, 4 associate and 7 full professors, 10 men, 1 woman.

Introduction to the Concept of ‘Cognitive Errors’:

Cognitive scientists demonstrate that many of the selection and evaluation processes we undertake on a daily basis are alarmingly “contaminated,” despite our good intentions. In her work, JoAnn Moody, a national diversity consultant who works with university campuses, looks at a "number of cognitive errors and shortcuts routinely and unwittingly made by individual evaluators,” and then casts light on common dysfunctions within academic organizations that can
and usually do intensify the severity of the cognitive errors. When those involved in evaluation are not coached and given the opportunity to be thorough and deliberate (as is the case when a large number of candidates are reviewed in a short period of time) then errors will surely result. The issues discussed here will be discussed in the context of some of the most commonly witnessed cognitive errors according to previous research.3-7

It is difficult to suggest to faculty, who usually have served on numerous search committees, that the process they’ve engaged in, some for decades, is flawed. Cognitive errors are particularly difficult to discuss (and thereby avoid) because people are often insulted if they’re asked to take extra care in evaluating the applications of underrepresented minorities. One faculty member claimed “I don’t see gender or race in people. Really, I don’t. It doesn’t matter to me if they’re a man, woman, black or white – I’m just looking for excellence.” The reality is that everyone makes assumptions based on gender and race. We live in a society where culture, media and daily life is nowhere near gender or race-neutral – and we are affected by it. Volumes of research has shown this to be true, even in the specific case of evaluating faculty candidates.13,15 It therefore is naïve to assume that academic search committees among all others are above such influences. Additionally, these errors are not only committed by the majority. A study of randomly chosen academic psychologists showed that both men and women were more likely to vote to hire a male applicant over a female applicant with identical records, giving more weight to teaching, research, and service experience of the male applicant.13 In other words, these cognitive errors are errors that we make that are below the threshold of our self-awareness. However, with commitment and training, they can be reduced.

Examples of Cognitive Errors in this Case Study:

Negative (or Positive) Stereotypes: While not apparent explicitly, it is common for search committees to require a more stringent level of “proof” of competence from people that are underrepresented among their colleagues, and less proof in those that are familiar (either because of shared background, cultures, institutions, etc.). An example of this is the absence of the phrase “qualified man or qualified white person” in our vocabulary, whereas a “qualified woman” or “qualified minority” is a very common (and in most people’s minds a legitimate) phrase. This was apparent in some cases where questions were asked or doubts were raised about the technical competence of some candidates (very often women or minority candidates), but not others. There was more than once articulated the hope that a minority candidate would demonstrate extraordinary qualifications (presumably so that no one would suspect that they were hired solely because of their minority status). Also, it was suggested that “apparent” past successes of some candidates be dismissed by citing special conditions that were presumably not related to the candidate’s competence. Specifically, in one case it was suggested that a candidate secured multiple research grants only because she was given special consideration by the funding agency. (This is related to the ‘ignoring evidence’ - discussed below).14

This factor also creeps in to letters of recommendation. As has been documented by research, letters of recommendation for women and minorities are very different.15 Women’s personalities (kind, cooperative, friendly, etc.) and work habits (hard working, diligent, persistent) are referenced by letter writers far more often than those for men. Letters of recommendation for men more often issue praise for specific career achievements and note research brilliance and creativity. Women’s family situations are more frequently referenced in letters. In these case
studies there were several letters of recommendation that showed how praise for women is often qualified by disclaimers, including: “even though she had a baby, she’s been productive in her work”, “she followed her husband to a new job and changed fields, but has learned quickly”, “she is gentle and soft-spoken”, “she is cooperative and considerate”, and “she works very, very hard”. These are meant to be compliments, but belie some unintentional negative bias. The result of such bias in letters of recommendation is demonstrated by the following comments, “her letters are not nearly as strong as his”, and “she didn’t seem to pick the right person to give her a recommendation, this guy doesn’t really talk about her research”, and even more damning, “I suppose we need to find out what her husband does”.

In these case studies, some gender stereotypes also crept into conversations about candidates. A women was described as an undesirable colleague because she was assertive and aggressive (what successful scientist isn’t?), or in another case a woman was thought of poorly because she was seen as self-promoting in her cover letter. Both instances suggest a double standard when it comes to evaluating women vs. men. While common characteristics of successful scientists, these are qualities that are generally in conflict with our gender schema for how we \textit{think} women ought to (stereotypically) act. Conversely, another woman was described as “timid”, and the question was raised “could she really hold her own?” in our aggressive field. The unintended disadvantages that these kinds of assessments of women candidates create include the expectation that a woman not be too assertive, on the one hand (lest she be labeled “pushy” or “bossy”), and that she not be modest (i.e., “timid”), on the other. This creates a very small “box” for women candidates to fit into.

Another stereotype that was voiced frequently is that of a woman being more often thought of as a helper or a contributor rather than a leader. In effect, a woman’s contribution to a group effort was often questioned, “was she just a joiner or did she have original, creative ideas?” Could she succeed on her own? This is not to say that the same question wasn’t asked of a few men as well but on average, this kind of question was asked of far fewer men proportionately than of women. This was particularly problematic in the college-level search where candidates were expected to be interdisciplinary (and therefore work with colleagues from multiple departments).

\textbf{First Impressions:} It is important to note that unintentional biases or “cognitive errors” can result in positive \textit{or} negative bias. Unintentional biases, on average, however, result in overly-positive evaluations of men and overly-negative evaluations of women. Physical appearance, clothing, etc. can (and does) result in assumptions based on our personal values and can bias the fair review of the science and a candidate’s potential to contribute. Sometimes these were positive, sometimes negative. One candidate was referred to as very young (although no younger or less experienced than two others) simply because she was very short and petite. It had to be pointed out that she, indeed, was NOT younger or less experienced than some others in the pool. Another male candidate was large and athletic with a booming voice, and it was suggested that he would have a great presence in the classroom and be an effective teacher.

\textbf{The Longing to Clone:} Candidates that come from an institution familiar to the evaluator will often get a closer look at their records and have their achievements overestimated compared to those who come from unfamiliar institutions. Candidates who have a professional background that is not similar to any of the evaluators tend to be underrated. This is related to “provincialism” – giving more weight to letters of reference that come from people that are known to the evaluator. There were several occasions that this unintended bias was evident.
Research has shown for years that this practice produces a more homogeneous workforce than results from evaluation practices that involve use of the same categories of criteria to evaluate all candidates. In other words, if informal recommendations (positive and negative) about any one candidate can be introduced as part of the evaluation criteria, the search committee should be required to seek out similar kind of information about other candidates in the pool.  

**Good Fit/Bad Fit:** It was often suggested that a candidate was a “good fit” or a “bad fit” for the department (or college), and when asked to elaborate, sometimes the evaluator could not refer to specific stated criteria. In other words, the evaluation was based on a highly subjective impression. This language can be used in response to the (unspoken) question “will I feel comfortable and culturally at ease with this new hire?” Of course, this is contrary to the goal of diversifying our faculty. If a candidate is considered to engage in research that does not complement the departments existing expertise, or does not have qualifications or potential in teaching that is required then that is how the deficiency should be described (not simply as a ‘bad fit’). It is legitimate to consider whether a candidate would be collegial and share responsibility in the department, but this is a separate issue from feeling immediately comfortable with someone.

**Distorting or Ignoring Evidence: Extraneous Myths/Assumptions:** Another common cognitive error involves psychoanalyzing candidates or making paternalistic assumptions. Some evaluators, in other words, express what they assume to be the candidate’s preferences—even though such preferences were never stated by the candidate herself. For example: “She wouldn’t feel comfortable here”, “I’m sure she’ll be taken by the time we make a decision”, and “She says she’s interested, but I can’t imagine why”. Speculation regarding whether any given candidate “might feel comfortable” is not a criterion for evaluating that candidate’s strengths or weaknesses. Acting on such speculation effectively removes the candidate’s right to decide for herself or himself whether she or he is interested in a given job.

Drawing conclusions based on preconceived notions or biases: “His major professor is one of the very best in that field – and he always has the best students – I’ll bet this guy is really sharp.” But then “Her major professor is one of the very best in that field – I wonder if any ideas were really hers?” “Sure she has been awarded a number of large grants, but the work isn’t very creative”. “Sure he has a great publication and grant record, but I wonder if the work is really his?” It was also disturbing (but not uncommon) to see the revision of evaluations after the candidate visited and even sometimes a complete reversal after conversations with influential faculty. (This is related to another cognitive error – ‘raising the bar’ – effectively changing the stated criteria during, or even after, the search process.) In other words, sometimes, aspects of the candidate’s record that had not previously caused concerns, were later pointed to as significant flaws. (This could mean that the records were not reviewed carefully enough initially or that the records were viewed through a different lens after the visit or after conversations about the candidate). While it is true that a candidate’s visit can (and indeed, should) provide additional information, it is likely evidence of a cognitive error when a particular aspect of a candidate’s record that was previously considered positive is later considered a deficiency. (In this example, the evaluator is seizing a pretext, presumably to disguise the real concern.)

Another factor that contributes to biased evaluations is flaws or bias inherent in the scoring system. For example in our case study, the evaluation form had specific categories of ratings (e.g. establish and direct a successful funded program, effective communicator, collaboration
potential) but also included a less defined “overall impression”. While it is valid to attempt to capture that which is difficult to specify, it was disturbing that most of the time the “overall assessment” was more positive for majority candidates than the average of other factors and each time less positive for the minority candidates than the average of other factors. For example, when comparing two candidates, the numerical average for the four defined categories gave one (minority) candidate a 1.38 (lower is better), and the other a 1.43. However, the “overall assessment” score for the first candidate was 1.64 (a considerably worse rating than the numerical average) and the other a 1.36 (better than the numerical average). Note that the ‘best score’ reversed depending upon whether specified criteria or ‘overall impression’ was given more weight. For the one search in which numerical designations for particular criteria were requested in addition to an overall assessment value, the trend held true (to varying degrees) for all 6 candidates (of which 2 were from underrepresented groups and 4 were not). The academic equivalent of this is after assigning points for specific criteria on a paper, awarding one student an A and the other a B in conflict with the actual scores. This pattern leads one to wonder, by what criteria are evaluators adding value to majority candidates and subtracting value from minority candidates?

Premature Ranking/Digging In: “All too often, evaluators rush to give numerical preferences to the candidates or applicants they are considering.” This rush-to-ranking seems to relieve evaluators of the worry of biased evaluation and falsely assure them that they are being objective. In one search, rankings were requested and reported – before any formal conversations were held to discuss strengths and weaknesses. In some cases, the conversation turned to the relative rankings instead of the candidates’ attributes.

Factors that contribute to cognitive errors:

There are a number of factors inherent in the typical faculty search process that potentially contribute to errors. In other words, often the structure of search committees and their standard practices promote (or exacerbate) cognitive errors. Unless a search is very restrictive, most will attract dozens if not hundreds of applicants, each with curriculum vita, publications list, research plans and other supporting documents that result in a package that can be 15-50 pages long (or sometimes longer). This overload results in rushing – scanning documents to develop a short list. Additionally, few faculty are given release time to serve on these committees, and this work (however important) is seen as service work which does not contribute to promotion. As such, there is little accountability or reward for diligence in evaluation, much less engaging in efforts to broaden the applicant pool.

Very often in early stages of evaluation, if not at all stages, there are not clearly defined and agreed upon selection criteria to specify which credentials are desirable and which are required. It is not uncommon for scores and rankings for individual candidates to be assigned after cuts and selections have been made. There are even some who eschew the assigning of scores based on criteria claiming “I know excellence when I see it”.

Another factor that contributes to poor search committee functioning is the lack of training. Corporations devote enormous resources to ensure managers are well trained in searches. Most universities do not. Additionally, because there is little training, there is usually no one serving as a process monitor to point out cognitive errors or lapses in protocol.
Finally, it is rare that the committee reconvenes after the search is over to reflect on the process and the successes and failures of the search. For that reason, all the inefficiencies and errors as well as successes are lost but in the memory of the committee members – who may or may not use them to improve the next search. With regard to the searches discussed here, more than once it was stated, “How could the search process be considered faulty? We got a great new faculty member”. This faulty logic does not acknowledge the opportunities lost.

Minimizing the impact of cognitive errors and the factors that contribute to them:

A number of excellent web-based resources exist for improving the recruitment process that are summarized on the national ADVANCE Portal. Additionally, ISU’s ADVANCE team has developed a number of resources for search committees in the form of short (often 8-10 minutes) presentations. Presented here are just a few key steps to optimizing faculty search processes.

Motivating the effort: There are a number of strategies that can be employed to help people be more mindful of their own decision making processes, but it is first necessary to motivate them to make the effort. Training should first address two issues. The first is to validate the efforts towards diversification. One faculty member stated “It doesn’t make any sense to bend over backwards to recruit non-traditional candidates when we have so many great candidates with such very strong ‘normal’ backgrounds”. This can be addressed from a number of perspectives including developing a faculty that better mirrors the rapidly changing student population. However, perhaps a greater motivation in an era where most research institutions strive to improve their national rankings, is that a growing body of evidence suggests that diverse teams result in higher productivity and creativity. The second strategy to motivate change is to show evidence that cognitive errors are common and unintended bias does affect the outcome of the hiring process and advancement of women and minorities. For engineering faculty, it is particularly useful to present literature that quantifies the effects, for example “The Impact of Gender of the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study” and “Exploring the Color of Glass: Letters of Recommendation for Female and Male Medical Faculty”.

Discussions of Cognitive Errors and Unintended Bias: The search committee should discuss common cognitive errors in evaluation of faculty candidates to become more aware of how these can compromise the fairness of the search process. This is best done before any vitae are reviewed. Joanne Moody’s handbook developed for this purpose is an excellent resource.

Setting Ground Rules and Developing a Matrix for evaluation: Clarify what the expectations are from each/all search committee members. If a committee member cannot meet these expectations for a given search, that committee member should be substituted. Ideally, time should be allotted to perform the tasks associated with the search, but as that is often not possible, at minimum secure agreement from the committee member that they will participate fully in the process.

Have a documented procedure to follow before beginning the search. This includes; specifying the criteria for evaluating candidates (some may be common to all searches, some may be specific to a particular search), creating or reviewing existing evaluation forms for suitability for
the current search, discussing procedures and constructing a timeline, discussing methods of 
evaluating the quality of a candidate’s scientific contributions to be applied uniformly to all 
candidates and reminding reviewers to address specific criteria when discussing a candidate.

While a matrix for evaluating the candidates is very useful and can contribute to fair evaluation, 
avoid ranking candidates until after their strengths and weaknesses have been thoroughly 
discussed.

Consider a process monitor: A process monitor is a non-voting member of the committee who is 
familiar with the process and the reporting requirements and can refocus the committee’s 
attention if they get off track. This person can also ask “what evidence do you have for that 
assessment?” Restrain from assigning this task to a woman faculty member already on the 
committee to evaluate candidates.

Plan and execute a self-reflection meeting: After the search has concluded, schedule a meeting to 
evaluate the process. Include evaluation of the schedule, the advertising, personal recruitment, 
phone interviews, campus interviews, paper and/or e-mail communications with the candidate 
and references. Consider the strengths and weakness, unintended consequences and areas for 
improvement. After this meeting, it is important to document the results to provide guidance to 
the next search committee. Consistent application of this practice will result in continuous 
 improvement and increase both the efficiency and effectiveness of search efforts.

Summary:

This case study involving two searches demonstrates that cognitive errors and unintended bias 
can (and does) influence the faculty search process, which can result favoring traditional, 
majority candidates. However, with education and training, and improved search process 
procedures, faculty can learn to recognize, reduce and minimize the impact of bias in search 
committees. Such education and training can serve to ‘unsettle’ unexamined assumptions. This is 
the first step towards more equitable evaluations. While these case studies are of search 
committees, faculty members are evaluated throughout their careers. Evaluations for 
recognition, awards, tenure, and promotion define the trajectory of a faculty career. Developing 
a culture of mindful evaluation and awareness of bias both in oneself and in colleagues will 
contribute to better utilization of all of our nation’s workforce resources independent of gender 
or race.

References:

1 National Academy of Sciences, 2006, Beyond Bias and Barriers: Fulfilling the Potential of Women in 
Academic Science and Engineering National Academies Press


“The Number of Women Faculty in Engineering Colleges Is Increasing Rapidly - This Study Compares Fractions of Women Faculty in Engineering and Engineering Disciplines to Fractions of Doctoral Degrees Awarded Annually - Will the Growth Continue?” Engineering Trends – REPORT 0907D Sept. 2007


Eveline, Joan. 2004. Ivory Basement Leadership. Crawley, Western Australia: University of Western Australia Press


Iowa State University’s ADVANCE Website Resources for Faculty Searches http://www.advance.iastate.edu/resources/fac_search.shtml.