

Date: April 22, 2014

To: Academic Senate Committees

From: Professor Linda Bisson, co- Chair  
Professor Jonathan Eisen, co-Chair  
ADVANCE Policy and Practices Review Initiative Committee

Dear Colleagues:

We are sending you a draft version of the current list of recommendations of the ADVANCE Policy and Practices Review Initiative Committee (PPRI). This is a true draft. We will submit the draft to the campus community for comment and as an initial ideas sweep to see if there are additional recommendations that should be included in this document. The final version of this document will be included in the third-year review materials for the NSF external review to be conducted next academic year. Once this initial sweep for comments and further ideas has been completed PPRI will re-issue this report for formal consultation, so you will see it a second time.

Because this report includes suggested processes for going forward, some committees are specifically named for development and implementation of select recommendations. I hope to get this up on the ASIS site for review by the senate and federation communities mid-May. We are sending it to you now for a quick review of your suggested roles so that if you are opposed to being charged with a task, or wish it to be a modified task, you can let us know. When we post the document for campus comments we will make it clear that this remains a draft and that these remain suggested paths forward should we as a campus adopt the recommendations. Thus this first review is to let you know this is coming and to see if you have objections to the roles suggested for your respective committees, knowing that these roles can be changed at a later date.

Thank you!

ADVANCE Policy and Practices Review Initiative Committee  
 Draft Report on Recruitment  
 Draft Recommendations Submittal for Academic Senate Review

Number	Recommendation	Committees Impacted
1	Modify UCAMP 210-1 section c4 (see ADVANCE Policy and Practices Review Initiative Overview)	Committee on Academic Personnel (CAP)
2	Make extensions to the clock for parental leave automatic (see ADVANCE Policy and Practices Review Initiative Overview)	Committee on Planning and Budget (CPB)
3	Develop a core set of best practices for optimizing demographic parity and inclusiveness in application pools for each college and school	Affirmative Action & Diversity (AA&D)
4	Evaluate and enhance college/school level review of position descriptions and search plans	College Faculty Executive Committees
5	Create a UC-wide website for the advertisement of faculty positions	CPB
6	Generate an institutional statement on commitment to diversity and inclusiveness to be included on all position announcements	AA&D
7	Create a new targeted program "Faculty Targets of Transformation" that could serve to recruit faculty at all ranks with the potential to strongly enhance efforts to build an inclusive campus environment.	AA&D, Graduate (GC) and Undergraduate Council (UGC), CAP & CPB
8	Enable commitment to diversity statements to be included in the application process	AA&D
9	Evaluate current pipeline programs	AA&D and GC
10	Require STEAD training of all search committee and department chairs and ultimately of all faculty engaged in faculty review and evaluation.	
11	<i>Eliminate valuation of gendered traits in pool evaluation</i>	AA&D
12	<i>Include awareness of pedigree bias in training of chairs and faculty.</i>	AA&D
13	<i>Review campus policies and practices with respect to search waivers for faculty positions</i>	AA&D, CAP, CPB
14	<i>Reward mentorship activity in advancement</i>	AA&D, CAP, Committee on Faculty Welfare (CFW)
15	<i>Institute standards for equitable resource availability</i>	College Faculty Executive Committees and CPB
16	<i>Establish clear policies on expectations of early career workload balance</i>	College Faculty Executive Committees and CPB
17	<i>Align University and federal family friendly policies via lobbying of federal granting agencies to institute automatic cost- extensions for parental leave</i>	CPB
18	<i>Review of department climates</i>	UGC
19	<i>Enhance childcare and caregiver support programs</i>	CFW
20	<i>Convert our POP program to a dual career program</i>	CFW, CPB
21	<i>Create a Program of Faculty Career Peer Advisors</i>	Faculty Privilege & Academic Personnel Advisers/Executive Council
22	<i>Create an online comprehensive faculty handbook</i>	Committee on Emeriti

## **DRAFT April 2014**

### **ADVANCE Policy and Practices Review Initiative Committee Draft Report on Recruitment**

Both cultural and structural factors pose barriers for diversification of the professoriate and can have career-long impacts as described in the Policy and Practices Initiative Overview report. The committee will present a set of recommendations for addressing these issues at each stage of the career ladder. These recommendations are draft recommendations as are the suggested paths forward should we elect to embrace the recommendation as a campus. We invite comment from the campus community on these draft recommendations, including suggestions for modification, additional recommendations or deletion and replacement of specific recommendations. The entities, Senate Committee or administrative unit are asked to review and comment upon their proposed roles in going forward and we also welcome suggestions for modification of the proposed plans for development and implementation of the suggested recommendations. Once the comment period has ended we will prepare a finalized formal report to be sent back to the appropriate committees and the campus community for further review, to be followed by delineating the processes of implementation for each of our final campus recommendations.

This report will focus on the issue of recruitment and what is and can be done to enhance the diversity of faculty appointments. We have divided recruitment into three specific aspects: pool composition, pool evaluation and creating and sustaining an attractive working environment.

#### ***POOL COMPOSITION***

Achieving a diverse faculty requires achieving a diverse pool of applicants. Recent analyses of potential pool composition (PhD degrees awarded) suggests that many of the early barriers that forced leaks in the pipeline have been addressed but there still remains a greater attrition of women and URM from the pipeline towards a tenure-track position. However there is a greater diversity in pipeline pools than at any time in our history. It is important therefore to capture that pipeline diversity within our applicant pools. Our goal is to have the demographic parity of the applicant pool match that of the pipeline for a given STEM discipline. There are several approaches that have been taken by other institutions that have enhanced the diversity of their faculty and we make the following recommendations to attain enhanced diversity of our own applicant pools. Most common among these other institutions is the establishment of detailed review of position request assuring that they meet key institutional goals. Although extensive over-sight and micromanagement can be effective in some circumstances we favor instead establishment of practices that are able to be easily followed by all hiring units avoiding a lengthy review process, but review of both position requests and search plans needs to be addressed.

*Recommendation 3: Develop a core set of best practices for optimizing demographic parity and inclusiveness in application pools for each college and school*

Pool composition is impacted by a variety of practices. Given the scarcity of URM applicants in the pipeline, narrowly written or constricted job descriptions may lead to small applicant pools and potentially send the message that the position was written with specific applicants in mind. Other institutions have found that unconstricted searches lead to a more diverse the pool and a more diverse faculty. “Unconstricted” in this case means job descriptions that span a discipline or even span multiple disciplines. On the other hand if an unconstricted search leads to several hundred applicants practices employed to narrow down the number of applicants to a manageable level for review, such as requiring a high number of first author publications or publications as first author in top tier outlets, can severely reduce diversity even below that of a more constricted search. The UC Recruit application software has the potential of enabling a comparison of pool composition and level of constriction of the job description. When possible, searches should be as unconstricted as possible to optimize pool diversity. However, often faculty positions must be targeted to specific gaps or needs in disciplinary expertise within the department. Balancing disciplinary expertise within a department can be vital to the continued success of the unit and enable a higher quality undergraduate and graduate education experience. This is particularly acute in STEM where successful candidates require a significant investment in laboratory space, equipment and start-up packages in order to establish successful research programs. Thus departmental resources might not in fact be flexible if academic programs and stature are to be maintained.

Current campus practice is that an Interim Recruitment Report (IRR) is required for all ladder rank faculty searches that must be approved by the Dean of the College or School. In 2013 the Vice Provost for Academic Affairs (VPAA) began requiring deans to provide the IRR to her office for evaluation, comment, and (where indicated) a recommendation to the Dean to either approve or disapprove the IRR (see: [https://academicaffairs.ucdavis.edu/local\\_resources/docs/VP\\_advisories/AA2013-08\\_Advisory\\_to\\_Deans\\_-\\_New\\_Process\\_for\\_Approval\\_of\\_Interim\\_Recruitment\\_Reports\\_091313.pdf](https://academicaffairs.ucdavis.edu/local_resources/docs/VP_advisories/AA2013-08_Advisory_to_Deans_-_New_Process_for_Approval_of_Interim_Recruitment_Reports_091313.pdf)). Evaluation of diversity (gender, ethnicity) is part of this process. The office of the VPAA will compare the diversity of the applicant pool, the “strongly considered” pool, and the “short list” pool against workforce-specific diversity information available through RECRUIT. According to Everett Wilson, who is the Data Coordinator in Academic Affairs, the workforce-specific diversity information comes from:

“...The figures are based on an annual survey of PhD's performed by the National Organization Research Center (NORC): <http://www.norc.org/Research/Projects/Pages/survey-of-earned-doctorates-%28sed%29.aspx> on behalf of six federal agencies (National Science Foundation, National Institute of Health, Department of Education, National Endowment for the Humanities, United States Department of Agriculture, and NASA). The latest year's results can be viewed on the NORC web site, a good overview can be found on the NSF web site at: <http://www.nsf.gov/statistics/srvydoctorates/#sd>. The survey includes gender/ethnicity and the field of service most closely associated with a given PhD's

study.

The office of the VPAA also examines the hiring goals for the department and college/school. If the diversity of the pool (and especially the short list pool) is below that expected by the workforce-specific figures provided, or will not potentially contribute to the hiring goals, then the VPAA can recommend that the Dean not approve the IRR and either fail the search or keep the search underway until the diversity of the pools increases. However it is important to make sure that Departments are aware of the best practices for obtaining a diverse pool and that we as a campus understand the factors at play when the pool does not attain demographic parity.

We recommend that the campus use the data on assessments of pool demographics for all fields not just those associated with STEM to create knowledge base of best practices for attaining an applicant pool that matches the demographics of the pipeline pool. An annual report comparing the effectiveness of specific practices with the attained demographics of the pipeline pool should be generated. In those cases where the demographics are not in sync an in-depth post mortem analysis should be done to determine the point at which pool demographics changed significantly. We acknowledge that such an effort would be more effective if done systemwide which would enhance our ability to both experiment with and assess differences in practices across the ten campus system and more quickly arrive at a consensus of most effective practices for each academic field.

*Suggested Process for Moving Forward:*

We suggest that Office of the Vice Provost for Academic Affairs (VPAA) prepare an annual report based on recruit categorizing searches by the levels of demographic parity in the applicant, the strongly considered and short list pools of the searches conducted in that year that have been completed. This report should be given to the Committee on Affirmative Action and Diversity (AA&D) for review. The AA&D committee can then compare cases of successful attainment and persistence of demographic diversity to the actual processes followed in the search in order to identify those practices that generally lead to demographic parity. Practices detracting from demographic parity can also be identified. Their findings should be reported on a regular basis to the campus community so that the adoption of best practices will become widespread. We further suggest that the annual reporting to the academic community described above be done at the systemwide level to take full advantage of the efforts in obtaining greater demographic parity that are currently being conducted on all ten campuses and AA&D may wish to take this idea forward to the systemwide Affirmative Action and Diversity Committee.

*Recommendation 4: Evaluate and enhance college/school level review of position descriptions and search plans*

Aligned with recommendation 3 is the recommendation to then apply the best practices information obtained to searches within a unit and across units. This recommendation has two components. The first component is the assessment of current college and school practices for review of both position requests and search plans employed by Dean's offices. The point of generation of the IRR may be too late in the process to make the meaningful changes that will have an impact on pool demographics. We assume all review practices are in compliance with the APM, but perhaps that is not the case. If practices are not in compliance with the APM they should be altered to become compliant.

The second component of this recommendation is to develop a faculty position (an individual to be appointed by the local Faculty Executive Committee) to advise both the Dean and departments on best practices for conducting searches and who will review position descriptions and search plans for best practices. We do not suggest creating cumbersome time-consuming practices that would delay replacement of vital faculty. The process developed should enable assessment of the appropriateness of the level of constriction of the search and to make sure best practices in advertising and pool evaluation are being followed. These plans should also include a detailed statement of how the position will be advertised that is also reviewed. Each college should also develop a plan for conducting cross-departmental searches to fill positions when appropriate to broadening a search. These processes should partner with, engage and respect the important role of departments in these searches particularly in an environment of limited local resources

#### *Suggested Process for Moving Forward:*

We suggest each unit (college or school) appoint a joint senate-administration committee to document current practices in position and search plan review for their unit. The charge to these joint senate administration committees should also include development of plans for cross-departmental searches, and cross college/school searches. These plans would be submitted to the Senate committees on Affirmative Action and Diversity, Committee on Academic Personnel, Faculty Personnel Committees of the unit and to the Committee on Planning and Budget and to the Vice Provost for Academic Affairs for review. We also suggest the Committee on Affirmative Action and Diversity develop guidelines for the appointment and role of the unit Search Best Practices Advisor to be appointed by the unit Committee on Committees with the responsibility of reviewing search plans for and advising departments on best practices for achieving diversity within the ranks of the applicant pool.

To facilitate development of a practice for searches that will span colleges and support our efforts to create an inclusive campus faculty, we suggest that each local committee send one member to a campuswide planning committee that would develop a campus plan for cross college and school searches. One such possibility might be using CAMPOS (Center for the Advancement of Multicultural Perspectives in Science) to define faculty positions that will have a fractional responsibility for participation in

CAMPOS initiatives and other campus initiatives devoted to mentorship and creating an inclusive environment.

*Recommendation 5: Create a UC-wide website for the advertisement of faculty positions*

Pool diversity improves when positions are advertised broadly. Departments are given some leeway in selection of outlets for position advertising at UCD and there are budgetary constraints that may force less than optimal outreach for recruitment of applicants. Departments also use targeted emails to colleagues and departments that produce talented postdoctoral fellows and graduate students in our areas of interest. Other departments advertise more strongly at national conferences and candidates interested in positions often attend these recruitment sessions at those conferences. Thus it would be difficult to develop one core set of best practices for advertisement of all positions as disciplines vary in terms of optimal outlets for reaching candidates. We propose bringing the candidates to us by creating a high-profile single easy to search and navigate site for all UC positions. The existence of the site then could be broadly advertised and would quickly be a direct source of information for potential applicants. Drawing applicants to us across disciplines in a coordinated fashion would reduce the need for individualized advertising plans requiring oversight and review. This site could also be linked to regional dual career hiring networks and enable UC to take a lead in recruiting other academic institutions and corporations to these dual career networks. This portal could be easily linked to positions within Recruit and the Recruit application process should include a query of how the candidate first heard of the position. How candidates hear of the position could then be used to formulate best practices for attaining a diverse applicant pool.

*Suggested Process for Moving Forward:*

We suggest that the Committee on Planning and Budget in conjunction with the Vice Provost for Academic Affairs take the proposal for a systemwide website for position advertisement forward to the systemwide Committee on Planning and Budget and the Office of the President, respectively to both make the case and obtain endorsement for the creation of the position advertising website.

*Recommendation 6: Generate an institutional statement on commitment to diversity and inclusiveness to be included on all position announcements*

Many campuses are creating positions that require evidence of a commitment to diversity on the part of the successful candidate as a means to both addressing remaining issues in the pipeline, providing diverse role models, and achieving demographic parity. We recommend that language be developed that will clearly state UCD's commitment to and valuation of diversity. Currently the minimalist statement required at UCD for all applications (APM UCD500 exhibit D) is: "The University of California is an affirmative action/equal opportunity employer." This statement is inconsistent with the UC-wide minimum statement "The University of California is an

Equal Opportunity/Affirmative Action Employer. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability, age or protected veteran status.” This statement falls far short of stating a firm commitment to inclusiveness.

We therefore suggest UCD develop a clear institutional statement documenting our commitment to inclusion and that this statement be required on all position announcements. We offer as a suggestion the following or similar language be included in all position announcements, not just those limited to STEM:

**“UCD is committed to mentorship of students, staff and faculty including those from underrepresented and underserved populations and creating a vibrant and inclusive educational environment. We especially seek applications from candidates who share this vision.”**

*Suggested Process for Moving Forward:*

We suggest that the Committee on Affirmative Action and Diversity develop suitable language for an institutional statement of commitment to inclusiveness or alternatively suggest a much more expanded statement such as that of Cornell University (<http://diversity.cornell.edu/commitment-to-inclusion>) to propose to the campus for implementation.

*Recommendation 7: Create a new targeted program “Faculty Targets of Transformation” that could serve to recruit faculty at all ranks with the potential to strongly enhance efforts to build an inclusive campus environment.*

Currently the only targeted faculty programs on campus eligible for a search waiver are the Targets of Excellence and Partner Opportunity Program supporting dual career hires. Targets of Excellence are generally limited to senior faculty with strong records of research accomplishment, faculty who are already “stars” within their discipline. A Targets of Transformation program would allow Departments to identify and propose targeted hiring of faculty with solid research programs but that also have even stronger records in teaching and mentoring that would equally enhance the campus climate of excellence in areas other than research. An alternative approach would be to broaden the current operational definition of the TOE program. We do not favor this latter approach because of the review process that might be required. Clearly the experts within a discipline are best able to identify stars within that discipline and know what strategic hires would best enhance that discipline on campus. The TOT hires would carry non-disciplinary expertise in teaching, mentoring and advising and therefore might be best reviewed by a committee containing representation from Affirmative Action and Diversity and Undergraduate and Graduate Councils in addition to review by CAP and the Committee on Planning and Budget.

*Suggested Process for Moving Forward:*

We suggest that the Committee on Affirmative Action and Diversity consider best mechanisms for attracting “stars” in teaching and mentorship to our campus, individuals that will be truly transformative and carry strong records in commitment to diversity and inclusion.

*Recommendation 8: Enable commitment to diversity statements to be included in the application process*

If recommendation #6 is adopted, a clear campus commitment to diversity will be stated in each position announcement. As a companion to this recommendation we also propose the inclusion of a diversity statement similar to the one used in advancement actions of faculty in the application process via Recruit. We will need to provide examples of the types of commitments that are currently being made by existing faculty as guidance for the type of commitment we are seeking in applicants and provide guidance to departments and review committees on how to interpret such statements. UCSD has just started such a program and uses these statements as part of the evaluation of candidates

*Suggested Process for Moving Forward:*

We suggest that the VPAA make the changes necessary to the Recruit site to enable inclusion of diversity statements as part of the application process. We also suggest that the Committee on Affirmative Action and Diversity develop along with the VPAA examples of current faculty diversity efforts which will both serve as a tribute to the work those faculty are engaged in as well as provide a guideline on the types of service that we enable on campus that applicants may join once they arrive on campus. AA&D should also develop a guide on how to interpret diversity statements and use them in conjunction with other metrics used in identification of the final successful candidate.

*Recommendation 9: Evaluate current pipeline programs*

Potential faculty elect to leave the pipeline often because of concerns regarding the incompatibility of a career in science and family formation and caregiver responsibilities as well as conflicts between cultural schemas and equitable advancement practices. The availability of flexibly timed positions in the pipeline can be beneficial to potential faculty who wish to postpone starting a tenure track position until primary caregiver responsibilities have lessened. This can be critical for dual career couples. Campus policies that place a hard five year maximal time limit on postdoctoral positions may reinforce the structural problems associated with STEM careers and worklife balance. Policy for postdoctoral fellow employment also stipulates a review of any request for a reduction from 100% time. Although this is supposedly automatically granted for caregiver responsibilities the fact that it must be applied for and approved can be viewed by postdoctoral fellows as family unfriendly. In addition we also recommend exploring development of a reentry postdoc program similar to that in existence at UCB

to enable return to the pipeline for reestablishment of a record of productivity needed for successful attainment of a faculty position.

*Suggested Process for Moving Forward:*

We suggest that Graduate Council and the Committee on Affirmative Action and Diversity conduct a review of our current practices for postdoctoral appointments with a goal of addressing the structural issues impacting pipeline demographics and make recommendations for change if needed.

**POOL EVALUATION**

Equally important as advertising broadly and addressing pool composition issues is the development of best practices for evaluation and review of that pool. Schemas or implicit bias have been repeatedly shown to negatively impact equitable review of women and URM faculty. From a policy perspective, our policies do not reinforce implicit bias but nor can they correct for its existence. We view two components to bias that can negatively impact career advancement: the existence of implicit bias or schematic thinking and bias introduced by career trajectory or pedigree.

*Schema Abatement*

*Recommendation 10: Require STEAD training of all search committee and department chairs and ultimately of all faculty engaged in faculty review and evaluation.*

The formation of schemas or expectations of behavior of an individual based on their identification as a member of a group are implicit and therefore difficult to ignore. Other institutions have addressed this problem through the targeted training of search and review committees. Indeed a component of this ADVANCE grant is to develop similar training for UCD and the ADVANCE Committee for an Inclusive Campus Climate has developed a training program called STEAD (Strength Through Equity And Diversity) that will be invaluable in the training of search committees in understanding implicit bias in its multiple forms and how it impacts review of applicants.

*Suggested Process for Moving Forward:*

We support the plan to require all search committee members undergo STEAD training in order to recognize the many forms that implicit bias might take in record review. However we think implicit bias is pervasive throughout the academy and all of its disciplines. Therefore we recommend that we develop mechanisms for the training of all faculty, senate and non-senate, in implicit bias so that we may all serve as agents of change to external bodies to increase more globally awareness of the cultural aspects of disadvantage in our respective disciplines.

*Recommendation 11: Eliminate valuation of gendered traits in pool evaluation*

The concept of the ideal STEM scholar is often imbued with gendered traits such as excelling at self-promotion, aggressively competitive, stingy with both time and credit given to others, a tough negotiator and dedicated to career as evidenced by a lack of worklife balance. Often these personality traits when expressed by those groups not normally considered to possess them can be viewed as negative. Similarly if those traits are not expressed by a member of the group thought to possess them that individual can be equally disadvantaged. Evaluation committee training should include recognition of these traits as gendered and not correlated with true impact of scientific discoveries as outlined in the 2007 NAS report “Beyond Bias and Barriers”.

Over the past three decades there have been numerous attempts to encourage all scientists to develop these gendered traits in order to be successful along with the companion efforts to convince review committees to apply valuation of these traits equitably and without gender or ethnic bias. The fact we still have the problem in spite of these decades of effort suggests that a different approach might be needed. Also it is not clear if there is in fact a correlation between possession and expression of these gendered traits and the innovation and impact of creative activity and research programs and the “inherent nature of science”.

*Suggested Process for Moving Forward:*

We suggest that the Committee on Affirmative Action and Diversity develop a report on the impact of gendered traits and implicit bias on both candidate pool evaluation and advancement through the faculty ranks that can serve to inform our evaluation process and be used in the training of Department Chairs and faculty review committees. More recommendations on this particular topic will be included in our report on our reward system.

*Pedigree Bias*

*Recommendation 12: Include awareness of pedigree bias in training of chairs and faculty.*

A second form of bias that can become a factor in both evaluation of candidates and in reduction of pool size is the placement of a high value on a specific scientific pedigree during the review process. URM faculty in particular often have non-traditional career paths such as starting at a community college rather than a research 1 institution or having gotten advanced degrees at universities not considered to be top tier even when their record of accomplishment matches that of traditional path candidates. For some positions particularly in narrowly defined sub-specialties the pool may be narrowed by a focus on targeted recruitment from a small set of Departments considered to be top in the field. Both of these practices can narrow pool composition unnecessarily.

*Suggested Process for Moving Forward:*

We suggest that the Committee on Affirmative Action and Diversity in their report on the impact of gendered traits and implicit bias generated suggested in recommendation 11 include an assessment of the bias in hiring imposed by seeking faculty with a specific pedigree or that follow a more traditional path to the PhD degree. This report can also serve to inform our evaluation processes and be used in the training of Department Chairs and faculty review committees.

*Recommendation 13: Review campus policies and practices with respect to search waivers for faculty positions*

Current policy is to allow departmental initiated requests for search waivers only in cases of TOE appointments (<https://academicaffairs.ucdavis.edu/programs/target-excel/index.html>). However, UCD APM 500 exhibit b lists the conditions under which academic appointments may be exempted from search requirements (<http://manuals.ucdavis.edu/apm/500b.pdf>). This includes discussion of acting or temporary faculty appointments, POP's, emergency appointments needed to meet critical teaching or clinical needs and a vague statement on exceptions for recruiting faculty with unique experiences and accomplishments that matches the language used for selection of the President's Postdoctoral Fellows. 500b also includes information on search waivers for grant funded programs when it is clear that individual hired is so uniquely qualified that they would have been hired anyway should a search have been conducted. Search waivers are to be approached with extreme caution as the absence of a pool of candidates may negatively impact efforts at attaining demographic diversity. The origin and intent of some of these exceptions to policy in hiring should be reviewed and their effectiveness assessed. In cases of emergency fast-tracking a pool-based search may be more advantageous than pre-selection of a qualified individual.

*Suggested Process for Moving Forward:*

We recommend that the Committee on Affirmative Action and Diversity review 500b and its application on campus in conjunction with CAP and the Committee on Planning and Budget and determine how search waivers are and have been used on campus and for what purposes and report their findings as well as any recommendations for change to the senate and campus administration.

**CREATING AND SUSTAINING AN ATTRACTIVE WORKING ENVIRONMENT**

The final task in development of a diverse faculty workforce is the creation of an enabling work environment that will be attractive to faculty candidates and create the inclusive, vibrant community that we seek. Two other ADVANCE committees, the Committee for an Inclusive Campus Climate and the Committee for Mentorship and Networking are along with the CAMPOS initiative are developing programs that will be critical to the creation of a highly attractive work environment for URM STEM faculty. We suggest that those programs being created and reports generated by those initiative

committees undergo full Senate review. From the policy and practices perspective we offer the following recommendations.

*Recommendation 14: Reward mentorship activity in advancement*

APM210 includes language specifying the importance of mentorship “mentoring and advising of students or new faculty are to be encouraged and given recognition in the teaching or service categories of academic personnel actions” and “opinions of graduates who have achieved notable professional success since leaving the university” is given as an example of appropriate documentation of mentorship ability. However often such letters are considered as not “arm’s length” by reviewing bodies and are discounted in the review process. We recommend that inclusion of such letters as well as other evidence of high impact mentorship in faculty review. The recent COACHE survey revealed a high valuation of mentoring by the campus academic community as well as indicating that all demographic subgroups think mentoring is undervalued and could greatly be improved. The ADVANCE Mentoring and Networking Initiative is developing mentorship programs and guidelines to address these and other concerns regarding the quality of and rewards for excellence in mentoring on this campus.

*Suggested Process for Moving Forward:*

We suggest that the Committees on Faculty Welfare and Affirmative Action and Diversity, along with CAP, review these proposed mentorship programs and develop clearer metrics for the evaluation of mentorship in our review process and that we publicize our efforts so that our commitment to mentorship is clear to prospective faculty. We will have further recommendations in this area in the report on our reward structure.

*Recommendation 15: Institute standards for equitable resource availability*

Obtaining the resources needed for establishment of a research program in STEM often relies exclusively on the negotiating talents of the candidate. As described above negotiation talent is often viewed as a gendered trait and thus reliance on this trait as an assessment of candidate quality can disadvantage women and URM faculty. Women and URM faculty report colleagues viewing them as greedy when the attempt to negotiate start up packages that match or exceed those of white male colleagues whereas white male colleagues are considered to be simply placing the correct evaluation on their needs. To address the cultural or schema impacts of expression of the negotiation trait we recommend a more formulaic approach to start-up package funding, one that strives for equitable treatment of all faculty and that is not reliant on variable personality traits. While efforts to train women and URM faculty in the art of negotiation are laudable, expecting someone to embody a personality trait that may be perceived as negative is counterproductive. A side benefit of developing such a program might be to stem the rapidly increasing magnitude of start-up packages by divorcing the magnitude of the package from the implied quality of the candidate.

*Suggested Process for Moving Forward:*

We suggest that the same college and school committees created under recommendation #3 be also charged with assessment of typical start up packages and the creation of standardized by discipline start up packages that reduce the reliance on negotiating skill in obtaining the funds needed to launch a successful career. The campus can develop competitive equitable methods for application for extra funds when well justified by the candidate's needs that would be put forward by the department Chair and not the candidate with full explanation. These recommended start up packages once developed by each unit should be reviewed by the Committee on Planning and Budget and the VPAA for review prior to implementation at the unit level.

*Recommendation 16: Establish clear policies on expectations of early career workload balance*

In concert with recommendation #15 in some cases reduced junior faculty workloads are part of the negotiation process and there is a tendency for these workloads to not be equitable across junior faculty. Workload reductions often take the form of reduced service and teaching loads for an initial period of time to allow the faculty member the time needed to develop their area of scholarship. The time demands of establishing a research laboratory and securing funding can indeed be severe. We recommend a uniform policy of workload reduction be developed by the campus and equitably applied to all new junior faculty. We realize it may be easier in some departments to offer reduced workloads without impacting the workload of other faculty and this also needs to be taken into account in the development and implementation of these policies.

*Suggested Process for Moving Forward:*

As in Recommendation #15, clear unit policies for junior faculty workloads should be developed by each local committee created under recommendation #3 and be reviewed by college/unit senate faculty personnel and executive committees. Policies should then go forward to the Committee on Planning and Budget and the VPAA for review prior to implementation at the unit level.

*Recommendation 17: Align University and federal family friendly policies via lobbying of federal granting agencies to institute automatic cost- extensions for parental leave*

The inflexibility of timelines of federal grants can negate institutional family friendly policies in those STEM fields where obtention of such grants is critical to establishment of a record of accomplishment. Providing no-cost extensions in cases of parental leave ignores the fact that at UCD the majority of the funding is being used to support the research team not the faculty member and a hiatus in funding will have a negative impact on the individuals with salaries or stipends covered on the grant. It is simply not possible to shut down the entire research enterprise during a parental leave.

### *Suggested Process for Moving Forward:*

We urge the UC administration to work with federal granting agencies to address this issue and develop policies for the fully-funded automatic extension of the grant for the period of the leave. Although family leave policies are largely mandated by state law the federal government should partner in compliance with these policies for federally funded research. Rather than an extension of individual grants which may pose a nightmare of documentation and excessive paperwork, an alternative would be to develop a mechanism by which the funding of this program could be included in overhead requests with the funds going to a specific pool to be used to augment grants for parental and other approved leaves mutually agreed to by both the University and the granting agency. We suggest the Committee on Planning and Budget along with the VPAA develop a strategy for aligning federal and local leave policy and bring the request forward to the systemwide Committee on Planning and Budget and to the Office of the President respectively for endorsement and review, then for systemwide review as this would need to be a systemwide policy applicable to all individuals funded from research grants.

### *Recommendation 18: Review of department climates*

The Committee for an Inclusive Campus Climate is developing tools for assessment of departmental climates and to provide assistance in creating and maintain an inclusive climate. Under current practices issues within a department or unit climate are only identified via a “whistleblower” mechanism, the lodging of a complaint against individuals within a department. Non-inclusive climates may be a cultural phenomenon of the entire department and not map on a specific individual or subset of individuals that could be named in a complaint. We therefore suggest that we develop a mechanism of climate assessment based on the report to come from the Inclusive Climate Initiative Committee that enables review of local climates and recommendations for change to those climates as needed.

### *Suggested Process for Moving Forward:*

We recommend that the Campus Climate Initiative Committee prepare a report on best types of climates as well as best practices for creation of an inclusive climate and for correction of one that does not appear to be achieving inclusiveness and that this report receive Senate review. We further recommend that the Academic Senate institutionalize include climate assessment by including it as a factor in review of undergraduate programs. We suggest that Undergraduate Council consider and propose mechanisms by which an assessment of departmental climate can occur during the regular review of undergraduate programs. However we note that the COACHE survey data for the campus suggests that campus climates are thought of more positively than at comparison institutions but since this is a comparison and not an absolute scale we should develop processes to assure all units enjoy positive and inclusive climates.

### *Recommendation 19: Enhance childcare and caregiver support programs*

One factor often raised by faculty with young children as needing attention and improvement is the availability on campus of quality childcare at affordable costs. In fact investment in child care is associated with a positive campus climate and commitment to equitable faculty advancement (Trower, 2013 *Success on the Tenure Track: Five Keys to Faculty Job Satisfaction*, Chapter 4). While we do have some programs in this area we recommend a review of their effectiveness and availability for faculty in light of Trower's recommendations in this area. She cites the importance of child care facilities locally located (within the school or college) that would be amenable to work schedules of faculty, meaning the potential for extended hours to capture faculty that may need on occasion to leave early to make a meeting off-campus or who may need to arrive late in order to participate in key networking opportunities such as dinners for invited speakers.

*Suggested Process for Moving Forward:*

We suggest that the Committee on Faculty Welfare review the recommendations in Trower's book and make recommendations specific to our campus based on their findings and the specifics of our local culture for the provision of child care on campus.

*Recommendation 20: Convert our POP program to a dual career program:*

Our current Partner Opportunity Program is important in the recruitment of dual career couples or individuals of a dual career couple and the importance of such programs was highlighted in Trower's *Success on the Tenure Track: Five Keys to Faculty Job Satisfaction* (2013). However by its very name the POP program reinforces that one spouse is the "leading spouse" career-wise and the other the "trailing spouse". In contrast campuses with dual career programs do not make valuations of the relative importance of the two careers.

*Suggested Process for Moving Forward:*

We recommend that the Committee on Faculty Welfare and Committee on Planning and Budget review our current POP program and consider suggesting changes in the program, at least at a minimum to rename the program, to a dual career program and review the information on successful dual career programs as described by Trower.

*Recommendation 21: Create a Program of Faculty Career Peer Advisors*

Mentoring is a critical component of both career success and career satisfaction and is highly valued by faculty on the UCD campus as evidenced in our COACHE survey results. Matching faculty with good mentors is a goal of the Mentoring and Networking Initiative but faculty may have questions and issues that they would feel more comfortable bringing up with a knowledgeable individual not part of their normal networks or departments. The creation of Faculty Career Advisors that would function similarly to the current Privilege and Academic Personnel Advisors is suggested.

Alternately the charge to the Privilege and Academic Personnel Advisors could be expanded from their current role to include being familiar with mentoring programs available on campuses and other resources needed by faculty.

The committee shall include members both with experience in the privilege and tenure process and in the academic personnel process, and every member shall be experienced in at least one of the processes. Current members of the Committee on Academic Personnel, Faculty Personnel Committees, or the Committee on Privilege and Tenure are ineligible to serve on this committee. Appointments are for one year and may be renewed.

*Suggested Process for Moving Forward:*

Executive Council should consider the formation of Faculty Career Advisors or expanding the role of the Privilege and Academic Personnel Advisors. In addition the existence of the Privilege and Academic Personnel Advisors should be advertised better and be better known to the faculty that this expertise is available.

*Recommendation 22: Create an online comprehensive faculty handbook*

As noted in Do Babies Matter (2013) surveys indicate that a majority of faculty are unaware of institutional family friendly policies and the availability of these programs. Trower also cites in Success on the Tenure Track: Five Keys to Faculty Job Satisfaction (2013) the importance of having a one-stop clearly presented web site with easily searchable and findable policies and practices for every aspect of faculty life. We recommend that resources be provided to the academic senate to create such an on-line faculty handbook that would be easy to navigate present information clearly, and be comprehensive. We suggest this handbook could be loosely patterned after the one at Stanford (<http://facultyhandbook.stanford.edu/>) that includes both policy statements akin to our APM but also information on programs available to faculty.

*Suggested Process for Moving Forward:*

We suggest the Senate in conjunction with the Emeriti Committee identify a group of Emeriti willing to devote time to creation of a comprehensive on-line faculty handbook. This would benefit current faculty as well as those considering applying to a position at our institution but acknowledges that our current practice of scattering such information across administrative web sites requires a depth of knowledge of “who does what” not common among new applicants or even among current faculty.

## ADVANCE Policy and Practices Review Initiative Overview April 2014

The UC Davis ADVANCE program aims to increase the participation of URM (Underrepresented minorities) women in STEM fields (science, technology, engineering and mathematics). The overall campus goal is to create an enabling environment and to “build a vibrant, welcoming and diverse STEM research community”. A second objective is to establish an institution-wide, inclusive STEM climate that values and embraces diversity. ADVANCE also plans to develop and provide programs and resources that will empower individuals for a STEM academic career and promote equitable advancement. The effectiveness of programs, practices and policies developed to meet these goals will be assessed in an ongoing basis.

The specific overarching objective of the UCD ADVANCE program as stated in the NSF grant application is to understand barriers and catalysts for Latinas in STEM fields as a route to enhancing the diversity of STEM on campus and to empower STEM fields to match the demographics of the populations that we serve. The very foundation of STEM disciplines is to create and transmit new knowledge and promote the advancement of human society by tackling critical issues via research and creative activity. The major issues facing society today are intricate and complex often requiring a multidisciplinary collaboration. Equally key to accomplishment in these research endeavors is the inclusion of culturally diverse perspectives and approaches. STEM must migrate from a culture of being open to diversity in its ranks to one that acknowledges that its very success and future impact is dependent upon that diversity in all of its forms.

The charge to Committee for the ADVANCE Policy and Practices Review Initiative is to conduct a systematic review of policies and procedures that impact recruitment, retention and career progress of faculty and, when deemed necessary, propose changes to those policies and practices in order to eradicate implicit bias, promote diversity and eliminate inconsistency in implementation. Our recommendations will be presented to the Davis Division of the Academic Senate for consideration and modification and ultimately for dissemination, review, implementation and to obtain feedback from the systemwide UC Academic Senate as appropriate. We are also charged with developing online spaces for protected faculty feedback on policies and implementation which will be done in conjunction with the Davis Division of the Academic Senate and the UC Davis Federation. We have divided our charge into four sections: recruitment, faculty advancement and the reward system, career development and retention. Four separate reports will be generated each focusing on one of these topics.

### *The case for change:*

The barriers to creating a fully inclusive environment and attaining demographic parity are both structural and cultural. Our campus objective is to achieve demographic parity throughout the ranks of the STEM faculty. Key to success in achieving this goal is the

recruitment, advancement and retention of a diverse workforce. Much attention has been given to the pipeline that feeds academe and its leaky nature for women and URM faculty as the primary barrier to creating a diverse academy. A focus on the pipeline was launched by the AAAS presidents lecture by Sheila Widnall "Voices from the Pipeline" published in Science in 1988. This lecture served as a call to action to address on-going and chronic pipeline issues. Today evidence demonstrates fixing of many of the barriers that forced leaks in the pipeline (Ceci and Williams, 2009) and suggests instead that the very nature of science leads to self-removal of women and URM from the pipeline (Blickenstaff, 2005). Other data reinforce the continuing impact of implicit bias on the career trajectories of women and URM faculty (Valian, 1999). Comparative studies show that other previously male-dominated professional fields that continue to embrace gendered criteria of success such as medicine, veterinary medicine and law show better demographic parity than STEM in academia. Other university disciplines, social sciences and humanities, likewise display better demographic parity than STEM. It then becomes important to understand what makes STEM differ from other academic disciplines and from other career paths within a STEM discipline. We differ slightly with the view of Blickenstaff that the core of the problem in both self-determined exit from the pipeline and in implicit bias is rooted in the very nature of science. Instead we believe it is rooted in the current value system and metrics used to judge the quality of a scientist and a scientific career. These quality metrics impose a critical barrier to attainment of demographic parity. Our goal in this and our remaining three reports is to launch a campus-wide critical assessment of our value system and metrics used in order to eliminate those that engender bias and encourage highly talented and qualified segments of our scientific workforce to exit the pipeline.

*Structural barriers to demographic parity:*

Academic STEM is characterized by a very rigid career structure. The inflexibility of this structure is generally incompatible with family formation and caregiving responsibilities and indeed often equates attention to those responsibilities with a lack of dedication to one's research and career. This perceived lack of dedication negatively impacts career development and progression even after the time demands of caregiving have lessened. A major reason talented individuals self-exit the pipeline as students and postdoctoral fellows is the realization of the enormity of the difficulty in balancing family life with success in their discipline (Mason et al, 2013). The years leading up to the tenure decision are often the most critical in family formation and our lack of acknowledgement of this in our reward structure leads to an exodus from the pipeline which is particularly acute for potential women and URM faculty. The major factor imposing rigidity on STEM careers is the pervasive use of and reliance on clocks in metrics of career assessment and the highly competitive nature of obtaining funding for the research enterprise that also runs exclusively on a fixed clock. Although we have made some progress in academia in developing and successfully implementing stop the clock policies the fact that funding agencies have not done so creates a conflict in adoption and use of those policies. No cost extensions, while helpful, ignore the fact that although the PI may be on parental leave the rest of the research team is not and funding will need to be expended to cover costs of salaries during the absence of the PI.

In addition a senior person able to oversee research during parental leave may be necessary to the continued success of the research endeavor. We cannot hope to fix the problem in STEM by focusing on institutional policies alone.

There are rigid timelines for normative progress throughout STEM academic careers beginning with graduate degree programs. Rigid timelines in degree programs are often justified by limited physical space in STEM laboratories and the need for turnover to allow new individuals the opportunities to become trained as scientists. Other disciplines that do not have such firm individual space or resource needs and limitations have been able to develop more flexible or open timelines than the STEM fields. Following graduate school postdoctoral appointments are also on a clock, generally imposed by granting agencies, with clear progress expected within a specific time frame. Although time in the postdoctoral ranks can be somewhat flexible allowing for family formation the impact on research productivity in the form of peer-reviewed publications in and of itself negatively impacts career metrics and may lead to a loss of competitiveness in obtaining research funding in the future. In addition to timeline norms and expectations for stages of an academic career as a metric of quality an accompanying metric is the level of productivity per year in each of those stages. There is no point at which starting a family and caregiver responsibilities can be accommodated without a negative impact to the career of the primary caregiver (Mason et al, 2013).

Upon appointment to the ranks of faculty the tenure clock starts and even given the prevalence of family friendly policies to stop the clock the need to remain competitive for grants re-imposes the clock and, depending upon the competitiveness of the discipline, may completely negate the value of institutional family friendly policies. Publications per year are also continually used as a metric for advancement. This is an outcome of the practice of the “60 hour work week” for faculty, which often in STEM disciplines given the nature of the work particularly if biological materials or animals are involved can be in reality 24/7. Under these circumstances policies that offer part-time employment as a means to achieve worklife balance often are not effective or applicable. It is also challenging to review files from individuals on part time status given the expected norm of open ended time investment for full time appointments. The clocked nature of the normative advancement poses not just a worklife imbalance but is in fact antagonistic to attaining a balance.

In spite of federal agency adoption of family friendly awareness programs for grant review data continue to show family status is strongly predictive of grant receipt (Mason et al, 2013. Chapter 6). Tenure track women with children are 26% less likely to have federal grant support than tenure track women without children. In addition the need to meet grant goals within a rigid timeline in order to maintain competitiveness places faculty in an untenable position with respect to the need for individuals within the lab supported by those grants to take parental leave. PIs clearly would prefer to support members of their research team in taking parental leaves but know that the research progress will suffer and competitiveness for the next grant may be in jeopardy. Postdocs and graduate students are also aware of this imbalance in institutional family friendly

policies and the realities of grant competitiveness and this conflict often leads to an exit from the pipeline (Mason, et al, 2013).

*Cultural barriers to demographic parity:*

In addition to these structural issues cultural issues can also pose a barrier to equitable treatment within academia. Metrics of success used in evaluation for both appointment and career advancement display inherent or implicit bias. Also known as schemas, implicit bias refers to the nonconscious expectations of individuals based on gender or ethnicity developed in childhood that in the professional world can give an advantage to one group over another. In a more colloquial sense bias is taken to mean attitudes, perspectives and prejudices of one group that confer an advantage to that group over another but that are not shared by both groups. In a similar sense implicit bias is also a set of attitudes, perspectives and prejudices that confer an advantage to one group over another but that are often shared by both groups in essence culturally driven. The breadth and depth of the negative impact of implicit bias on faculty recruitment and advancement was cogently detailed in Virginia Valian's book "Why so slow? The advancement of women" (1999) and reinforced in the NAS report "Beyond Barriers and Bias: Fulfilling the Potential of Women in Academic Science and Engineering (2007)". Valian clearly outlines the role of gender and racial schemas in evaluation of faculty and documents the negative impact of schemas on career advancement. An example of gender schemas in action can be found in the type of terms used to express the role of women versus men in science (Madera et al, 2009). Women are often described in terms of collaboration, consensus-building and nurturing while men are more often viewed in terms of independence and leadership. This differing vocabulary can appear in outside letters of recommendation but also may be used by the candidate in their own statement for advancement. If evidence of independence and leadership is crucial for advancement the reward system may lead to bias favoring the group for which that vocabulary is more commonly used.

A second type of example of intrinsic bias in the reward system concerns the perceptions of the appropriateness of expression of traits associated with success within our culture. One common such trait in evaluation of academic success is self-promotion. The trait of self-promotion is on a negative trajectory meaning that when carried to an extreme it is associated with the negative traits of arrogance, pride and egotism. The point at which expression of this trait crosses into the negative perception of that trait differs by cultural subgroup. Women and URM faculty often experience a negative reaction to levels of self-promotion that are acceptable for white males. But women and URM faculty are just as likely as white males to find that level of self-promotion to be negative. Over the past few decades there have been many calls for training and accepting self-promotion across all cultural sectors, but none of these have been successful. The cultural advantage of not displaying a negative trait can outweigh the career disadvantage of doing so. This is evidenced by looking at our own UCD culture. The campus is often criticized for being too humble, too nice and insufficiently self-promoting. As a consequence the campus contributions are often undervalued by peer institutions. Over the years we have launched several initiatives and adopted

practices aimed at increasing our level of self-promotion and we are well aware of the institutional advantages of doing so, yet our campus culture remains the same. In addition if individuals ascribe to the same cultural schema they may be fine with others displaying a negative trait and being equitably rewarded for it but unable to embrace that level of expression of that trait themselves as to do so would compromise their own self-identity.

It is imperative that we understand the breadth of the role played by schemas or intrinsic bias in our own reward system and develop schema abatement practices and policies in order to obtain a truly equitable review system for all faculty. Indeed this is the goal of the STEAD (Strength Through Equity And Diversity) training being developed by the ADVANCE Inclusive Campus Climate Initiative. It may also be essential to purge our reward system of weighting traits like self-promotion that may simply be incapable of equitable application across all segments of our culture. Is the ability to self-promote truly correlated with the quality and impact of the creative work being done? A reward system that values what can be perceived as a negative trait uncorrelated with the true worth of accomplishments may be a key factor in encouraging differential exit from the academic pipeline.

*The path forward:*

Both structural and cultural factors impose barriers to obtaining demographic parity among the ranks of the faculty. Although some of these barriers are more acute factors for STEM fields others impact all fields of academic inquiry. These barriers both disadvantage those on an academic career track as well as encourage exit from the academic career pipeline. The campus needs to collectively evaluate the impact of these factors on our reward system and advancement at all stages of an academic career and develop mechanisms to mitigate the effect of these barriers if not remove them altogether. The Policy and Practices Review Initiative Committee will issue four reports each focusing on a different critical area. Each of these reports will be presented to the campus academic community in draft form for review of the proposed recommendations and to enable the community to modify, reject or pose additional recommendations for going forward. A final report will be prepared for formal consultation after these comments and suggestions have been obtained so that these recommendations and guidelines will truly reflect the views of the campus academic community as a whole. The combination of these reports in final form will create a clear campus roadmap for going forward and attaining our goal of demographic parity within our ranks.

The first report will focus on recruitment and cover the sub topics of pool composition, pool evaluation and the creation of an attractive campus climate. This report will be issued for informal comments in the spring of 2014. The second report will concentrate on our reward system and include not just our criteria for advancement but also a broader assessment of our overall incentives for faculty use of time and will be issued at the end of fall quarter in 2014. The third report will cover the topic of faculty development programs at all stages of the career ladder and will be issued for informal

comment in the spring of 2015. The final report will address retention on campus to be issued for informal comment in the fall of 2015. We expect to have final reports for each of these within 3 to 6 months of obtaining initial comments and suggestions depending upon the magnitude of those suggestions and need to explore them in depth. Each recommendation of each report will be numbered in sequence and each will include a section proposing a draft strategy for development and implementation. The campus community will be invited to critique both the recommendation and the proposed strategy for going forward.

*Immediate steps:*

Although the reward system is the subject of our second report we wish to make two recommendations immediately to address a longstanding concern of members of the Policy and Practices Review Initiative Committee. The APM in an effort to negate the impact of additions to the tenure clock states (UCAPM 210-1, section c4, Assessment of evidence (emphasis is ours)):

*“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.”*

The use of “if,” and implication that the work has not been completed within a usual or normal time frame greatly reduces the family friendly intent of this policy. These accommodations must be considered the norm and not an exception to be viewed without prejudice. Although it may be challenging to change this wording in the systemwide APM we recommend that the Davis Division of the Academic Senate and Chancellor make efforts to do so.

*Recommendation 1: Modify UCAMP 210-1 section c4*

We offer as a suggestion to modify UCAPM 210-1 section c4:

**The University of California allows flexibility in the timing of tenure review as defined in APM-760. It is our policy that the overall record of productivity and scholarly attainment form the basis of evaluation. Time since appointment is not a consideration in review of the candidate’s record of accomplishment.**

*Suggested Process for Moving Forward:* We suggest that the Committee on Academic Personnel consider proposing such a change to the APM with the new language to be modified as they see fit and to take this suggestion systemwide via UCAP.

*Recommendation 2: Make extensions to the clock for parental leave automatic*

We further recommend that extensions to the clock for parental leaves be automatic and not require notification that a faculty member has elected to take advantage of the clock extension by the Chair or Dean. The faculty member may then elect to request promotion at any time within their allotted clock. We repeatedly hear from faculty that Chairs and Deans have advised them against using this program so as to not negatively impact future advancement and ability to be competitive for grants. Given the baby penalty that has been well documented recently (Mason et al 2013. *Do Babies Matter?*) this advice unfortunately is not at all inaccurate. Under current policy and practice a request for an extension to the clock is automatically approved but the process of requesting it seems to provide an opportunity for others to caution against it and getting such an extension should be an entitlement automatically generated.

*Suggested Process for Moving Forward:* We suggest that the Committee on Planning and Budget work with the administration in identifying a process by which such an extension can be automatic, either by eliminating the need to appeal for an extension after the fact if it was not requested immediately or by working with HR to identify faculty who have added a dependent to their health insurance so that the extension of the tenure clock can be made absent any action on the part of the individual.

#### References:

Blickenstaff, JC. 2005. Women and science careers: Leaky pipeline or gender filter? *Gender Educ.* 17:369-386.

Ceci, SJ, Williams, WM. 2009. Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci., USA* 108:3157-62.

Institute of Medicine, National Academy of Sciences, and National Academy of Engineering. 2007. *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. The National Academies Press. Washington, DC ([http://www.nap.edu/catalog.php?record\\_id=11741](http://www.nap.edu/catalog.php?record_id=11741) )

Madera, JM, Hebl, MR, Martin, RC. 2009. Gender and letters of recommendation for academia: Agentic and communal differences. *J. Appl. Psych.* 94:1591-1599.

Mason, MA, Wolfinger, NH, Goulden M. 2013. *Do Babies Matter? Gender and Family in the Ivory Tower*. Rutgers University Press. New Brunswick.

Valian, V. 1999. *Why So Slow? The Advancement of Women*. The MIT Press, Cambridge, MA.

Trower, CA. 2012. *Success on the Tenure Track: Five Keys to Faculty Job Satisfaction*. The Johns Hopkins University Press, Baltimore.

## Gender and Letters of Recommendation for Academia: Agentic and Communal Differences

Juan M. Madera  
University of Houston

Michelle R. Hebl and Randi C. Martin  
Rice University

In 2 studies that draw from the social role theory of sex differences (A. H. Eagly, W. Wood, & A. B. Diekmann, 2000), the authors investigated differences in agentic and communal characteristics in letters of recommendation for men and women for academic positions and whether such differences influenced selection decisions in academia. The results supported the hypotheses, indicating (a) that women were described as more communal and less agentic than men (Study 1) and (b) that communal characteristics have a negative relationship with hiring decisions in academia that are based on letters of recommendation (Study 2). Such results are particularly important because letters of recommendation continue to be heavily weighted and commonly used selection tools (R. D. Arvey & T. E. Campion, 1982; R. M. Guion, 1998), particularly in academia (E. P. Sheehan, T. M. McDevitt, & H. C. Ross, 1998).

*Keywords:* gender stereotypes, letters of recommendation, academia, social role theory

The problem of pipeline shrinkage for women in academia is a well-known and researched phenomenon (Bellas & Toutkoushian, 1999; Camp, 1997; Olsen, Maple, & Stage, 1995; Taylor, 2007; Windall, 1988). This phenomenon refers to the fact that women enter graduate school at about the same rate as do men, but women are less likely to enter and succeed in academia at the same rate as their male counterparts, particularly in science and engineering disciplines. In fact, the National Science Foundation (2008) has reported that women comprise about 29% of science and engineering faculty at 4-year colleges and universities and comprise only 18% of full professors. One contributing factor to this gender disparity may be gender differences in letters of recommendation. In particular, there is little research that addresses whether letters of recommendation for academia are written differently for men and women and whether potential differences influence selection decisions in academia. The present study addresses this issue.

The focus on letters of recommendation is justified because they are an important and commonly used selection tool that provides information on applicants' past performance and qualifications (McCarthy & Goffin, 2001), confirms or supplements information

provided by applicants (Brem, Lampman, & Johnson, 1995), and describes applicants' motivation (Tommasi, Williams, & Nordstrom, 1998). In fact, Cascio and Aguinis (2004) stated, "the fact is, decisions are made on the basis of letters of recommendations" (p. 278). In particular, they have been found to be among the most important criterion used to screen and evaluate applicants for internships (Lopez, Oehlert, & Moberly, 1996), graduate programs (Landrum, Jeglum, & Cashin, 1994), medical schools (Johnson et al., 1998), military training programs (McCarthy & Goffin, 2001), and psychology faculty positions (Sheehan, McDevitt, & Ross, 1998).

Regardless of the reasons for using letters of recommendation, research has shown that letters of recommendation can be written differently for women than for men (McCarthy & Goffin, 2001; Trix & Psenka, 2003). In qualitative studies, researchers have reported that letters of recommendation for college (LaCroix, 1985) and graduate school (Watson, 1987) contained stereotypical gender-related words and phrases, describing female applicants as feminine and male applicants as masculine. Using discourse analysis, Trix and Psenka (2003) analyzed over 300 letters of recommendation for doctors applying for medical faculty positions and found that letters were longer for male than female applicants. In addition, letters for men contained more standout adjectives, such as *superb*, *outstanding*, and *remarkable*, and contained more research-related descriptors than did letters for women. In a study that replicated Trix and Psenka's (2003) study, Schmader, Whitehead, and Wysocki (2007) examined letters of recommendation for science faculty positions and found that letters for male compared to female applicants also contained more standout adjectives. However, there were no statistical differences in length, positive and negative language, and research- and teaching-related words for male and female applicants.

There are, however, important limitations to the previously conducted research studies. First, many of these findings were based on descriptive rather than on inferential statistics. Second, many studies used subjective rather than objective means for

---

Juan M. Madera, Conrad N. Hilton College of Hotel and Restaurant Management, University of Houston; Michelle R. Hebl and Randi C. Martin, Department of Psychology, Rice University.

This research was supported by National Science Foundation Grant SBE-0317739 to Rice University, State University of New York at Stony Brook, and Pennsylvania State University, and by National Science Foundation Grant HRD-0542562 to Rice University. We thank Suparna Rajaram and Judith Kroll for many discussions of the factors influencing the evaluation of women that led to the present research. We also thank Laura Martin for her help in collecting and analyzing the data. Study 1 was presented at the 22nd annual meeting for the Society for Industrial and Organizational Psychology in New York.

Correspondence concerning this article should be addressed to Juan M. Madera, Conrad N. Hilton College of Hotel and Restaurant Management, University of Houston, 229 C. N. Hilton Hotel and College, Houston, TX 77204-3028. E-mail: jmmadera@uh.edu

scoring gender differences, with the authors of the study carrying out the scoring rather than relying on more objective alternatives (i.e., the Linguistic Inquiry and Word Count [LIWC] program; Pennebaker, Francis, & Booth, 2001). Third, many studies did not use statistical procedures (e.g., hierarchical linear modeling) that address the fact that letters of recommendation are nested within applicants. Fourth, there were important variables that were not controlled for in these previous studies that might have affected the results. Fifth, these prior studies did not examine whether gender differences in letters actually affected judgments about hireability.

Thus, to improve on the methodology of these earlier studies, we examined gender differences in letters of recommendation with objective methods (i.e., language content analysis; Pennebaker et al., 2001), with statistical procedures appropriate for nested data, including indicators of productivity as control variables (e.g., publications, teaching experience, and honors), and by assessing the effects of gender differences on judgments of hireability. We also situated our examination of potential differences within contemporary theorizing about gender bias, specifically drawing on the social role theory of sex differences (Eagly, Wood, & Diekmann, 2000).

### Gender Stereotypes: Agentic and Communal Characteristics

According to social role theory, behavioral sex differences arise from the division of labor—the differential social roles inhabited by women and men (Eagly et al., 2000). Historically, men have been more likely to engage in tasks that require speed, strength, and the ability to be away from home for expanded periods of time, whereas women were more likely to stay home and engage in family tasks, such as child rearing. Accordingly, men are perceived and expected to be agentic, and women are perceived and expected to be communal. Agency includes descriptions of aggressiveness, assertiveness, independence, and self-confidence (Eagly & Johannesen-Schmidt, 2001). Agentic behaviors at work include speaking assertively, influencing others, and initiating tasks. Communal behaviors at work include being concerned with the welfare of others (i.e., descriptions of kindness, sympathy, sensitivity, and nurturance), helping others, accepting others' direction, and maintaining relationships (Eagly & Johannesen-Schmidt, 2001).

Central to understanding gender stereotypes is that they are both descriptive and prescriptive (Burgess & Borgida, 1999; Rudman & Glick, 2001). Stereotypes suggest not only how men and women do behave (i.e., descriptive), but also how men and women should behave (i.e., prescriptive). The prescriptive nature of gender stereotypes also specifies what women should not do—often leading to penalties for women who do not conform to their respective norm. As a result, women are expected to engage in a feminine gender role that reflects communal qualities but not agentic ones (Wood & Eagly, 2002). The descriptive and prescriptive nature of these stereotypes can affect women's entrance and mobility in certain jobs. For example, agency has been found to be associated with roles of leadership (Eagly & Karau, 2002; Heilman, Block, & Martell, 1995). Managerial and executive level jobs are usually considered to be a masculine role—thought to require agentic qualities, such as ambition, aggressiveness, and achievement. Therefore, attitudes are often less positive toward female than male leaders, and it is more difficult for women to become and succeed

as leaders (Eagly & Karau, 2002). This occurs because women are perceived and expected to be communal, but leaders are perceived and expected to be agentic. It is this *lack of fit*, or mismatch, between attributes of gender or their prescriptive components of shoulds and work roles that can affect women in the workplace (Heilman, 2001). Thus, it is important to examine whether women are being described in letters of recommendation as less agentic and more communal than men because agency is related to higher status and success in the workplace.

## Study 1

### Overview and Hypotheses

To examine whether social role theory might explain gender differences in letters of recommendation, we analyzed letters of recommendation written for applicants for faculty positions in a psychology department at a Research I university (as designated by the Carnegie Classification of Institutes of Higher Education). According to social role theory (Eagly et al., 2000), because men are usually ascribed agentic characteristics, we expected that men would be described in more agentic terms than would women in letters of recommendation and that women would be described in more communal terms than would men.

*Hypothesis 1.* Women are more likely than men to be described in communal terms in letters of recommendation.

*Hypothesis 2.* Men are more likely than women to be described in agentic terms in letters of recommendation.

### Method

#### Sample

The sample consisted of 624 letters of recommendation and 194 applicants for eight junior faculty positions from 1998 to 2006 at a Southern university in the United States. Of those whose sex could be identified, 46% ( $n = 89$ ) of the applicants were women, and 54% ( $n = 103$ ) were men; 30% ( $n = 193$ ) of the recommenders were women, and 70% ( $n = 477$ ) were men.<sup>1</sup> Applicants' ages ranged from 26 to 40 years, with a mean of 32 ( $SD = 3.69$ ). The mean number of letters per applicant was 3.23.

#### Procedure

This study used archival data obtained from the psychology department. After receiving institutional review board approval, we transcribed the original letters of recommendation to electronic form (Microsoft Word document) and used a computer text analysis program, the LIWC program (Pennebaker et al., 2001). The LIWC program analyzes text files and computes the percentage of words from a file that fall into each of 74 possible linguistic

<sup>1</sup> The faculty positions were in the follow areas: applied experimental ( $n = 2$ ), applied psychology ( $n = 49$ ), cognitive ( $n = 22$ ), cognitive/neuroscience ( $n = 43$ ), cognitive/neuroscience developmental ( $n = 6$ ), health ( $n = 32$ ), industrial/organizational ( $n = 37$ ), and social ( $n = 3$ ). Male recommenders wrote 262 letters for male applicants and 194 letters for female applicants; female recommenders wrote 78 letters for male applicants and 109 letters for female applicants.

categories, such as negative emotion, self-reflection, causation, and physical issues. The program dictionary is composed of 2,300 words and word stems and was developed with emotion rating scales (e.g., the Positive Affect Negative Affect Scale; Watson, Clark, & Tellegen, 1988). The LIWC dictionary was validated by having judges rate the content of hundreds of text files, comparing their results to those of the computer program (Pennebaker & Francis, 1996; also see Pennebaker & King, 1999).<sup>2</sup>

### Measures

**Communal adjectives.** We created a dictionary for communal adjectives derived from Eagly's work in communal and agentic characteristics (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Eagly et al., 2000; Wood & Eagly, 2002).<sup>3</sup> The final list included terms such as *affectionate, helpful, kind, sympathetic, sensitive, nurturing, agreeable, tactful, interpersonal, warm, caring, and tactful*. The average percentage of words for communal adjectives in each letter of recommendation was 0.69%.

**Social–communal orientation.** The social orientation index in LIWC counts the number of words that deal with other people. Psychologically, it reflects how much letter writers referred to other people when writing about the applicant. Words in this category include *husband, wife, kids, babies, brothers, children, colleagues, dad, family, they, him, and her*. The average percentage for social–communal words in each letter of recommendation was 8.22%.

**Agentic adjectives.** We created a dictionary for agentic adjectives. The list of words was also derived from Eagly's work on communal and agentic characteristics (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Eagly et al., 2000; Wood & Eagly, 2002). This list includes *assertive, confident, aggressive, ambitious, dominant, forceful, independent, daring, outspoken, and intellectual*. The average percentage for communal adjectives in each letter of recommendation was 0.81%.

**Agentic orientation.** The cognitive mechanism, motion, and achieve indexes in LIWC count the number of words that deal with other peoples' cognitive processes, achievements, and actions. Psychologically, these indexes reflect how much letter writers referred to the applicants as active, dynamic, and achievers. As such, we constructed a single category by using a composite score of the three indexes. Words in this category included *earn, gain, do, know, insight, and think*. The average percentage for agentic orientation words in the letters of recommendation was 7.42%.

**Gender.** Gender for both the applicants and the recommenders was coded female (1) or male (2).

**Control variables.** We used seven control variables. These were the number of years in graduate school, the number of total publications, the number of first author publications, the number of honors, the number of postdoctoral years of education, the position applied for, and the number of courses taught.

### Results

Given that the letters of recommendation were nested within applicants, we used the HLM 6 program (Raudenbush, Bryk, Cheong, & Congdon, 2004) to analyze the data by conducting hierarchical linear modeling. We used full maximum likelihood estimation procedures and included random effects. Hypothesis

testing involved three steps: (a) the control variables were entered into the model, (b) the main effects variables (applicant gender and letter writer gender) were entered in the equation, and (c) the interaction variable (Applicant Gender  $\times$  Letter Writer Gender) was entered into the model. Thus, for the analyses, the intercepts of the Level 1 variables, communal adjectives, social–communal index, agentic adjectives, and agentic orientation index, were predicted by the Level 2 variable, gender of the applicant.<sup>4</sup> For exploratory reasons, we also included the gender of the letter writer. Descriptive statistics and intercorrelations for the control, independent, and dependent variables are reported in Table 1.

### Hypothesis 1

As shown in Table 2, gender of applicant significantly predicted both communal adjectives ( $\beta = -.16, p < .05$ ) and the social–communal index ( $\beta = -.12, p < .05$ ). More specifically, the results show that women were described by more communal terms than were men. In addition, letters written for women mentioned more social–communal terms than letters for men. Thus, the results supported Hypothesis 1. The gender of the letter writer significantly predicted the social–communal index ( $\beta = .11, p < .05$ ), but not the communal adjectives ( $\beta = .08, p < .05$ ). The interaction between gender of the applicant and gender of the letter

<sup>2</sup> Despite the published evidence for the validity of the LIWC program (e.g., Pennebaker & Francis, 1996; Pennebaker et al., 2001; Pennebaker & King, 1999), we further validated our LIWC measures with three coders (who were unaware of the hypotheses) who rated the extent to which letter writers described the applicants as communal and agentic. More specifically, the coders used a 9-point scale (1 = *not at all* to 9 = *very much*) to respond to the following two questions: "To what extent was the applicant described as agentic (assertive, independent, aggressive)?" and "To what extent was the applicant described as communal (kind, nurturing, caring)?" The coders were provided with the definition and examples of agency and communal descriptors. The coders' ratings of agentic description were significantly related to agentic adjectives ( $r = .17, p < .05$ ) and agentic orientation ( $r = .12, p < .05$ ). Similarly, the coders' ratings of communal description were significantly related to communal adjectives ( $r = .19, p < .05$ ) and social–communal orientation ( $r = .12, p < .05$ ).

<sup>3</sup> We also used the program WordNet (<http://wordnet.princeton.edu/>) to search for synonyms of the word in the list. WordNet is an online lexical reference system, designed with psycholinguistic theories of human lexical memory, which organizes English nouns, verbs, adjectives, and adverbs into synonym sets, each representing one underlying lexical concept. We used it for both the Communal Adjectives and Agentic Adjectives measures.

<sup>4</sup> Analysis for each outcome variable (i.e., the communal and agentic measures) begins with fitting an unconditional model (i.e., with no predictors) to estimate or partition the variance components. This essentially tests the within-applicant and between-applicant variance. If no within-applicant variance exists in the dependent variables, then hierarchical linear modeling is not appropriate because there is only one level of analysis (i.e., no within-applicant variance). The results of the unconditional (null) models indicated that there was significant between-applicant variance in each dependent variable ( $p < .01$  for all variables) and that a substantial proportion,  $P = r^2/(r^2 + t_{00})$ , of the total variance in these dependent variables was within applicants. That is, 26% of communal adjectives, 23% of social–communal, 20% of agentic adjectives, and 10% of agentic orientation variance was within applicant. Thus, there is substantial variance between and within applicants that warrants the use of hierarchical linear modeling to examine Level 1 and Level 2 dependent and independent variables.

Table 1  
Means, Standard Deviations, and Intercorrelations for Study 1

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Gender of applicant	1.51	0.50	—												
2. Gender of writer	1.71	0.28	.23*	—											
3. Communal adjectives	0.0069	0.30	-.16*	-.13	—										
4. Social–communal	0.082	1.32	-.28*	-.29*	.27*	—									
5. Agentic adjectives	0.0081	0.29	.11	-.07	.04	.03	—								
6. Agentic orientation	0.074	1.04	.11	.19*	.01	-.05	.23*	—							
7. Years in graduate school	4.17	2.0	.10	.11	-.10	-.11	-.11	-.06	—						
8. No. of publications	4.24	3.5	.09	.10	-.11	-.33*	-.11	-.01	.06	—					
9. No. of first-author publications	1.93	2.2	.13	.10	-.05	-.35*	-.11	-.01	.01	.75*	—				
10. No. of honors	0.91	1.4	-.10	-.10	-.02	.05	.01	.15*	-.01	.05	.06	—			
11. Postdoctoral years	1.10	1.5	.15*	-.08	-.11	-.24*	-.15*	-.01	.05	.39*	.44*	.18*	—		
12. No. of courses taught	5.45	3.3	-.09	-.01	.11	.18*	-.01	-.02	.12	-.10	-.11	-.03	-.02	—	
13. Position applied	3.38	1.7	.01	.09	.10	.02	.02	.16*	-.01	.07	.03	.09	.09	.08	—

Note. Gender was coded as female = 1, male = 2. Level 1 variables were aggregated: communal adjectives, social–communal, agentic orientation, and agentic adjectives.  
\*  $p < .05$ .

writer was not significant for communal adjectives or social–communal orientation.

Hypothesis 2

As shown in Table 2, the results revealed that gender of applicant significantly predicted agentic adjectives ( $\beta = .16, p < .05$ ); men were described by more agentic terms than were women. However, gender of applicant was not related to the agentic orientation index ( $\beta = .04, p > .05$ ). Thus, Hypothesis 2 was partially supported. The gender of the letter writer was not a significant predictor of agentic adjectives and agentic orientation. As depicted in Figure 1, there was a significant interaction of gender of the applicant and gender of the writer on agentic orientation ( $\beta = .12, p < .05$ ). Analyses revealed that for female applicants, male writers used more agentic orientation terms than did female writers ( $\beta = 3.02, p < .05$ ), whereas the pattern for male applicants was the opposite but was not significant ( $\beta = .63, p > .05$ ).

Discussion

With respect to social role theory, our results confirmed our hypotheses, demonstrating that female applicants were more likely to be described with communal terms (e.g., affectionate, warm,

kind, and nurturing) than male applicants. Letters of recommendation for female applicants also mentioned more social–communal terms, such as student(s), child, relative, and mother. In contrast, male applicants were more likely to be described in agentic terms (e.g., ambitious, dominant, and self-confident) than were female applicants. Thus, we found support for our hypotheses. It is important to note that these differences were obtained even though we included objective measures of performance from applicants’ curriculum vitae. That is, unlike past research (e.g., LaCroix, 1985; Trix & Penska, 2003; Watson, 1987), the current study included productivity factors, such as the number of publications, teaching experience, postdoctoral years, and honors. Such factors could affect the quality of letters of recommendation and the use of agentic and communal descriptions.

The interaction between applicant gender and letter writer gender showed an interesting pattern; for female applicants, male writers used more agentic orientation terms than did female writers. Although we did not hypothesize an interaction, it could be the case that men are more likely to emphasize agency than women when writing letters of recommendation. An alternative explanation could be that women focus more on communality and deemphasize the agency of women. This idea supports the universality of gender norms and the prescriptive stereotypes of agency and

Table 2  
Hierarchical Linear Modeling Results With Applicant Gender, Writer Gender, and Their Interaction as Predictors

Dependent variable	Applicant gender			Writer gender			Interaction		
	Estimate	SE	t	Estimate	SE	t	Estimate	SE	t
Communal adjectives	<b>-.16</b>	.19	-3.02*	.08	.03	-1.44	.11	.03	1.71
Social–communal orientation	<b>-.12</b>	.05	-1.98*	<b>.11</b>	.09	-2.10*	.10	.09	1.16
Agentic adjectives	<b>.16</b>	.06	2.46*	.02	.03	-0.77	.06	.03	-0.15
Agentic orientation	.04	.18	0.64	.04	.08	0.67	<b>.12</b>	.08	-2.26*

Note. Applicant gender was coded as female = 1, male = 2. Results are after controlling for the number of years in graduate school, the number of total publications, the number of first-author publications, the number of honors, the number of postdoctoral years, the applied position, and the number of courses taught. Values in bold are statistically significant.  
\*  $p < .05$ .

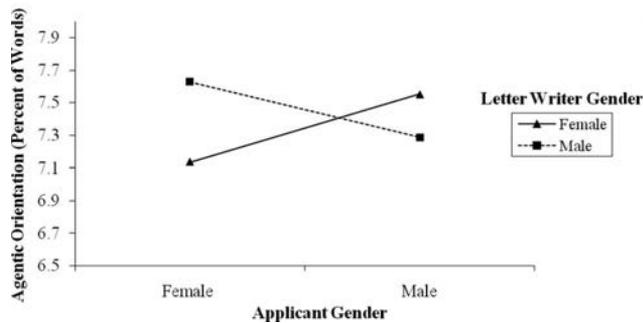


Figure 1. Applicant Gender × Letter Writer Gender interaction on agentic orientation from Study 1.

communion across both men and women, which is consistent with the broader literature on stigma (Crocker, Major, & Steele, 1998; Goffman, 1963; Hebl, Tickle, & Heatherton, 2000) and, more specifically, with the literature on sex differences (e.g., see Heilman & Okimoto, 2007; Heilman, Wallen, Fuchs, & Tamkins, 2004).

One question that emerges is whether the behavior of men and women differs along agentic and communal lines, as suggested by the frequency with which these terms were used, or whether the frequency of these terms reflects the writers' perception of the individuals colored by social role stereotyping or perhaps even by what they felt was appropriate for describing the individual. That is, a letter writer might perceive a male and female applicant to be equal in warmth, kindness, and helpfulness but may feel more comfortable about commenting on all of these things for the female than for the male applicant. Our data cannot speak directly to these issues.<sup>5</sup>

## Study 2

### Overview and Hypotheses

Study 1 showed that there are gender differences in communal and agentic characteristics in letters of recommendation. Therefore, it is important to examine whether differences in agentic and communal characteristics influence hiring decisions in academia. Central to understanding how agentic and communal characteristics are related to hiring decisions in academia is the gender typing of occupations as male or female on the basis of (a) job responsibilities believed to be gender linked or (b) the sex of the usual job holder (Heilman, 1995; Lyness & Heilman, 2006). Agency has been related to leadership and high-status occupations, such as academia, more than communal characteristics have (Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Heilman, 2001). Academia remains a male-dominated occupation (Bellas & Toutkoushian, 1999; Camp, 1997; Olsen et al., 1995; Windall, 1988). Thus, academic positions for research-oriented universities can be perceived to be masculine, and as a result, agentic characteristics might be positively related to hiring ratings. In contrast, communal characteristics might be negatively related to hiring ratings. As such, we predicted the following:

*Hypothesis 1.* Agentic characteristics included in letters of recommendation will be positively related to hiring ratings, but communal characteristics will be negatively related to hiring ratings.

Heilman's research (Heilman, 1995, 2001; Heilman & Okimoto, 2007; Heilman et al., 2004) demonstrated that communal characteristics can hinder women in the workplace, because women are expected to be communal. However, when women are perceived to be communal, they can be evaluated negatively when performing in occupations in which agency is perceived to be important. With reference to the present study, this body of work led to the prediction that communal characteristics in letters of recommendation would be negatively related to hiring ratings, and because women are more likely to be described with communal characteristics, women would be rated as less hireable than men for academic positions at a research university. Conversely, men would be rated as more hireable than women, because men are more likely than women to be described with agentic characteristics, which would be positively related to hiring ratings. According to this reasoning, gender differences in hiring ratings for academia would be mediated by gender differences in communal and agentic characteristics in letters of recommendation. More formally, we predicted the following:

*Hypothesis 2.* Men will be rated as more hireable than women for academic positions at a research university.

*Hypothesis 3.* Communal and agentic characteristics will mediate the relationship between gender and hireability in selection decisions for academia at a research university.

### Method

#### Participants and Procedure

Six psychology professors served as subject matter experts (SMEs). The SMEs were three professors in industrial-organizational psychology and three in cognitive psychology. After receiving institutional review board approval, the SMEs were provided with the letters of recommendation from Study 1 and were instructed to rate each applicant on hireability (i.e., hiring decisions in academia based on letters of recommendation). The SMEs were provided with a random sample of 100 letters of recommendation. They were provided with an additional 25 letters that were the same across raters to get an estimate of interrater reliability. In other words, SMEs reviewed 100 unique letters and 25 letters that all six SMEs reviewed. The letters of recommendation were modified by removing names of applicants and recommenders, names of schools, and the gender of the applicants and recommenders (i.e., gender identifying information was replaced by *he/she*, *his/her*, and *him/her*). The SMEs were instructed, "imagine that you are reviewing and rating applicants for a tenure-track assistant professor position in the psychology department" and asked to "rate the applicant based on this letter of recommendation."

<sup>5</sup> Given that we controlled for aspects of the curriculum vita that indicated research productivity, at least the agentic behaviors that underlie this productivity could not have been the source of the discrepancy in the use of the agentic terms. On the social-communal side, we had less in the way of objective measures that might have related to these terms; thus, we had fewer grounds for distinguishing among behavior, perception, or the letter writers' conformity to social norms as the basis of the differences.

## Measures

*Agency and communion.* The scales from Study 1 were used to measure agentic and communal characteristics: Communal Adjectives, Social–Communal Orientation, Agentic Adjectives, and Agentic Orientation. For the sake of simplicity, we developed composites of the agentic and communal measures by standardizing the scales and taking their respective means. Thus, for Study 2, we had an agentic composite (mean of the Agentic Adjectives and Agentic Orientation scales) and a communal composite (mean of the Communal Adjectives and Social–Communal Orientation scales).<sup>6</sup>

*Gender.* Participant gender was coded female (1) or male (2).

*Hireability.* Participants rated the hireability of the applicants with a 9-point scale (1 = *not at all*, 9 = *very much*) on four items (see the Appendix). The intraclass correlation coefficient (ICC1; interrater reliability) was .98 for the 25 overlapping letters. The intraclass correlation coefficient (ICC2; group mean reliability) was .97, and the alpha coefficient for the measure was .99. Thus, there was sufficient evidence for the reliability of the raters and the scale.

*Control variables.* The same control variables from Study 1 were used. We also included the gender of the letter writer as a control variable because the results from Study 1 showed some gender differences in letter writers.

## Results

Table 3 shows the means, standard deviations, and correlations for the control variables, applicant gender, letter writer gender, the communal composite, the agentic composite, and the outcome variable—hireability. To test Hypotheses 1 and 2, we conducted hierarchical linear modeling with the HLM 6 program (Raudenbush et al., 2004). We used full maximum likelihood estimation procedures and included random effects. For this hierarchical linear modeling analysis, the control variables were entered in the first step, and the communal composite, agentic composite, and applicant gender were entered in the second step.<sup>7</sup>

Table 4 shows the regression results. Supporting Hypothesis 1, communal characteristics were negatively related to hireability ( $\beta = -.28, p < .05$ ), indicating that a greater proportion of communal characteristics in the letters of recommendation was related to lower ratings of hireability. However, the proportion of agentic characteristics was not significantly related to hireability ( $\beta = .09, p > .05$ ). Thus, Hypothesis 1 was partially supported. Gender of the applicant was not related to ratings of hireability ( $\beta = .04, p > .05$ ), not supporting Hypothesis 2.

Because the regression analysis showed that gender and the hiring ratings were not related, we could not use the traditional Baron and Kenny (1986) mediation analysis to test Hypothesis 3. However, James, Mulaik, and Brett (2006) proposed structural equation modeling or path analysis as an alternative to test mediation. This approach does not require the distal variable to correlate with the outcome variable (i.e., gender and hireability in the current study). In fact, scholars have questioned whether it is necessary to provide evidence for a significant path from the distal variable to the outcome variable to establish mediation (Collins, Graham, & Flaherty, 1998; James et al., 2006; MacKinnon, 2000; MacKinnon, Krull, & Lockwood, 2000; Shrout & Bolger, 2002).

Rather, the simultaneous test of the significance of both the path from the distal variable to a mediator and the path from the mediator to the outcome variable (i.e., the structural equation modeling approach) provides, relative to other approaches (e.g., Baron & Kenny's, 1986, steps), the best balance of Type I error rates and statistical power (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). In light of this literature, we used the structural equation modeling/path analysis approach for testing the mediation hypothesis.

We used path analysis with Mplus (Muthen & Muthen, 1998) to test Hypothesis 3—that communal and agentic characteristics would mediate the relationship between gender and hireability in selection decisions for academia. To determine the adequacy of model fit, we used four fit indices: (a) chi-square and degrees of freedom, (b) the comparative fit index (CFI), (c) the incremental index of fit (IFI), and (d) the root mean square error of approximation (RMSEA). It is suggested that good fit indices for CFI and IFI are greater than .90 and that a good fit index for RMSEA is less than .08 (Byrne, 2001; Steiger, 1990; Vandenberg & Lance, 2000). The model we tested included gender as the distal predictor, with the communal and agentic composites as the mediators of the gender–hireability relationship and with the control variables in the model. This model demonstrated adequate fit,  $\chi^2(54) = 52.66, p > .05$ ; CFI = .97; IFI = .98; RMSEA = .044 (see Figure 2). Women were described as more communal than men,  $\beta = -.28, p < .05$ , and men were described as more agentic than women,  $\beta = .19, p < .05$ . The communal composite was negatively related to hireability ( $\beta = -.26, p < .05$ ). The agentic composite was not significantly related to hireability ( $\beta = .16, p > .05$ ). The variables in the model explained 27% of the variance in hireability ratings.<sup>8</sup> Thus, the results partially supported Hypothesis 3.

## Discussion

The results of Study 2 revealed that communal characteristics were negatively related to hireability ratings and that the communal ratings mediated the relationship between applicant gender and hireability ratings for a research-oriented university. The results for agency, however, were not as clear. Although we expected a positive relationship between hireability and agency, the results did not show a significant effect. It might be the case that agency is expected from applicants in academia, and therefore the agentic characteristics did not have an impact on hireability. It also might

<sup>6</sup> Bivariate correlations showed that the two composites had stronger relationships with the outcome variable—hireability—than the four scales, suggesting that the composites might be more appropriate than the individual scales.

<sup>7</sup> The results of the unconditional (null) models indicated that there was significant between-applicant variance in the hireability dependent variable ( $\chi^2 = 340.39, p < .01$ ) and that a substantial proportion,  $P = r^2/(r^2 + t_{00})$ , of the total variance in hireability was within applicants; that is, 22% of hireability variance was within applicant.

<sup>8</sup> Applicant gender explained 7% and 3% of the communal and agentic composites, respectively. As recommended by MacKinnon et al. (2002), we used the Sobel test to examine the significance of the indirect effect of gender and communal characteristics on hireability. The test indicated that gender had an indirect effect on hireability through its direct effect on the communal composite ( $z = 3.41, p < .05$ ).

Table 3  
Means, Standard Deviations, and Intercorrelations for Study 2

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Years in graduate school	4.17	2.0	—											
2. No. of publications	4.24	3.5	.06	—										
3. No. of first-author publications	1.93	2.2	.01	.75*	—									
4. No. of honors	0.91	1.4	-.01	.05	.06	—								
5. Postdoctoral years	1.10	1.5	.05	.39*	.44*	.18*	—							
6. No. of courses taught	5.54	3.3	.12	-.10	-.11	-.03	-.02	—						
7. Position applied	3.38	1.7	-.01	.07	.03	.09	.09	.08	—					
8. Gender of letter writer	1.71	0.28	.11	.10	.10	-.10	-.08	-.01	.09	—				
9. Gender of applicant	1.51	0.50	.10	.09	.13	-.10	.15*	-.09	.01	.01	—			
10. Agentic composite	-0.003	0.77	-.11	-.07	-.09	.11	-.10	-.02	.11	.09	.14*	—		
11. Communal composite	-0.001	0.75	-.14	-.28*	-.26*	.02	-.24*	.19*	.19*	-.27*	-.27*	.03	—	
12. Hireability	5.85	1.4	-.07	.46*	.38*	.07	.10	-.02	.18*	.10	.09	.02	-.29*	—

Note. Gender was coded as female = 1, male = 2. Level 1 variables were aggregated: communal adjectives, social-communal, agentic orientation, letter length, grindstone, research terms, teaching terms, doubt raisers, negative intensifiers, and positive intensifiers.  
\*  $p < .05$ .

be the case that the number of publications is a better indicator than the adjectives and descriptions provided by the letter writers. Correlations showed that the number of publications was significantly related to hireability. Although the raters did not have access to curriculum vitae, letter writers might have mentioned the total number of publications or discussed papers that might be under review or other papers being prepared for publications. If so, then the raters may have relied on this more direct information about productivity than on agency descriptors, thus accounting for the lack of an effect of agency on hireability decisions. The communal characteristics used by letter writers might be a better indicator of communality than other possible indicators in letters (i.e., the control variables from Study 1 and Study 2), because communion infers interpersonal information about exhibiting nurturing and socially sensitive attributes, which is information not necessarily conveyed in a vita.

Table 4  
Hierarchical Linear Modeling Predicting Hireability

Variable	Model 1		Model 2		Pseudo $R^2$
	Estimate	SE	Estimate	SE	
Years in graduate school	-.07	.04	-.10	.05	.07
No. of publications	.25*	.03	.21*	.03	
No. of first-author publications	.03	.07	.10	.07	
No. of honors	.05	.03	.03	.07	
Postdoctoral years	-.03	.04	-.09	.08	
No. of courses taught	.05	.01	.09	.03	
Gender of letter writer	.03	.10	-.05	.14	
Position applied (block)					
Gender of applicant			.04	.12	
Agentic attribution composite			.09	.11	
Communal attribution composite			-.28*	.12	

Note. Gender was coded as female = 1, male = 2. Standardized estimates and the standard errors are reported.  
\*  $p < .05$ .

General Discussion

The studies presented in the current article replicate and extend past research by showing (a) that there are gender differences in letters of recommendation—women are described as more communal and less agentic than are men (Study 1)—and (b) that communal characteristics have a negative relationship with hiring decisions in academia (Study 2). These results can be understood within the social role theory framework (Eagly et al., 2000). The data suggest that female applicants are described in accordance with communal gender norms, which are both descriptive and prescriptive (Eagly et al., 2000; Heilman et al., 1995). In addition, the results suggest that there is a lack of fit between the attributes of communality and the work role of academia. Such findings are particularly important because letters of recommendation are important and commonly used selection tools (Cascio & Aguinis, 2004; Sheehan et al., 1998).

This research not only has important implications for women in academia but also for women in management and leadership roles. A large body of research suggests that communality is not perceived to be congruent with leadership and managerial jobs (e.g., Eagly & Johannesen-Schmidt, 2001; Eagly & Karau, 2002; Heilman, 2001). In particular, Heilman’s (2001) lack of fit model, suggests that “fit-derived performance expectations, whether positive or negative, can profoundly affect evaluation processes” (p. 660). Thus, for occupations in which agency is linked to success or perceived as more important than communality, the perception of

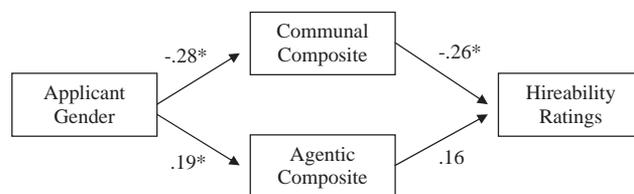


Figure 2. Communal and agentic composites as the mediators of the gender-hireability relationship from Study 2. Standardized estimates are shown. Results are with control variables in the model. \*  $p < .05$ .

lack of fit between a female applicant and the job requirements can arise as a result of women being described as more communal and less agentic than men. It is important to take caution, however, because letters of recommendation are not heavily weighted in some organizations and occupations.

As with most research, there are limitations to the current research. Although we used archival data and not hypothetical letters of recommendation in Study 1 (which is a strength), we cannot rule out the possibility that the differences in communal and agentic descriptions in the letters were based on real gender differences. Future research might try to disentangle true differences from perceived differences in agency and communion. Another potential limitation is that the hireability measure from Study 2 was based on perceived intentions to hire and not on actual hiring; therefore, we take caution with claims about possible career development for women in academia. However, the participants in Study 2 were faculty members from universities and therefore have experience in reading letters of recommendation and making selection decisions.

In addition, the percentages of communal and agentic adjectives in the letters of recommendation were low, such that one or two words could make a difference. However, the gender differences were large enough to be statistically significant, and although the effect was small, research by Martel, Lane, and Willis (1996) has shown that seemingly small gender differences may have enormous impact when compounded over time. Thus, small differences or what seem like molehills of disparity can become mountains of disparity over time and experiences (see Valian, 2000). Though quantifying the content of the letters of recommendation is an objective method and strength, the use of LIWC was also a potential limitation. In particular, LIWC does not take into account the context of the meaning of words. For example, a letter writer who mentioned that the applicant worked with a conscientious research assistant was scored the same as a letter writer who mentioned that the applicant was conscientious.

Despite these limitations, the potential implications of the current research for the use of letters of recommendation are important. The importance of letters of recommendation for academia is well established (Arvey & Campion, 1982; Guion, 1998; Sheehan et al., 1998), and the current research demonstrates differences in how men and women are described in letters according to gender norms of communality and agency (Eagly et al., 2000; Eagly & Johannesen-Schmidt, 2001; Wood & Eagly, 2002). This research showed that communal characteristics mediate the relationship between gender and hiring decisions in academia, suggesting that gender norm stereotypes—and not necessarily the sex of applicants—can influence hireability ratings of applicants.

## References

- Arvey, R. D., & Campion, T. E. (1982). The employment interview: A summary and review of recent research. *Personnel Psychology, 35*, 281–322.
- Baron, R. M., & Kenny, D. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.
- Bellas, M. L., & Toutkoushian, R. K. (1999). Faculty time allocation and research productivity: Gender, race, and family effects. *Review of Higher Education, 22*, 367–390.
- Brem, C., Lampman, C., & Johnson, M. (1995). Preparation of applications for academic positions in psychology. *American Psychologist, 50*, 533–537.
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy and Law, 5*, 665–692.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- Camp, T. (1997). The incredible shrinking pipeline. *Communications of the ACM, 40*, 103–110.
- Cascio, W. F., & Aguinis, H. (2004). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Collins, L. M., Graham, J. W., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research, 33*, 295–312.
- Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, 4th ed., pp. 504–553). New York: McGraw-Hill.
- Eagly, A. H., & Johannesen-Schmidt, M. C. (2001). The leadership styles of women and men. *Journal of Social Issues, 57*, 781–797.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598.
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Traunter (Eds.), *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Erlbaum.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York: Prentice Hall.
- Guion, R. M. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: Erlbaum.
- Hebl, M. R., Tickle, J., & Heatherton, T. F. (2000). Awkward moments in interactions between nonstigmatized and stigmatized individuals. In T. Heatherton, R. Kleck, M. Hebl, & J. Hull (Eds.), *The social psychology of stigma* (pp. 275–306). New York: Guilford Press.
- Heilman, M. E. (1995). Sex stereotypes and their effects in the workplace: What we know and what we don't know. *Journal of Social Behavior and Personality, 10*, 3–26.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*, 657–674.
- Heilman, M. E., Block, C. J., & Martell, R. (1995). Sex stereotypes: Do they influence perceptions of managers? *Journal of Personality and Social Psychology, 10*, 237–252.
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology, 92*, 81–92.
- Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male tasks. *Journal of Applied Psychology, 89*, 416–427.
- James, L. R., Mulaik, S. A., & Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods, 9*, 233–244.
- Johnson, M., Elam, C., Edwards, J., Taylor, D., Heldberg, C., Hinkley, R., & Comeau, R. (1998). Medical school admission committee members' evaluations of and impressions from recommendation letters. *Academic Medicine, 73*, S41–S43.
- LaCroix, P. P. (1985, fall). Sex in recs. *Journal of College Admission, 24–26*.
- Landrum, R., Jeglum, E., & Cashin, J. (1994). The decision-making process of graduate admissions committees in psychology. *Journal of Social Behavior and Personality, 9*, 239–248.
- Lopez, S., Oehlert, M., & Moberly, R. (1996). Selecting criteria for American Psychological Association-accredited internship programs: A survey of training directors. *Professional Psychology: Research and Practice, 27*, 518–520.

- Lyness, K. S., & Heilman, M. E. (2006). When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology, 77*, 777–785.
- MacKinnon, D. P. (2000). Contrasts in multiple mediator models. In J. S. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds.), *Multivariate applications in substance use research: New methods for new questions* (pp. 141–160). Mahwah, NJ: Erlbaum.
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effects. *Prevention Science, 1*, 173–181.
- MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*, 83–104.
- Martel, R., Lane, D. M., & Willis, C. (1996). Male–female differences: A computer simulation. *American Psychologist, 51*, 157–158.
- McCarthy, J. M., & Goffin, R. D. (2001). Improving the validity of letters of recommendation: An investigation of three standardized reference forms. *Military Psychology, 13*, 199–222.
- Muthen, L. K., & Muthen, B. O. (1998). *Mplus user's guide*. Los Angeles: Muthen & Muthen.
- National Science Foundation. (2008). *ADVANCE: Increasing the participation and advancement of women in academic science and engineering careers* (NSF 07–582). Retrieved June 23, 2008, from [http://www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf07582](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf07582)
- Olsen, D., Maple, S. A., & Stage, F. K. (1995). Women and minority faculty job satisfaction: Professional role interests, professional satisfaction, and institutional fit. *Journal of Higher Education, 66*, 267–293.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10*, 601–626.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC 2001): A computerized text analysis program*. Mahwah, NJ: Erlbaum.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*, 1296–1312.
- Raudenbush, S. W., Bryk, A., Cheong, Y. F., & Congdon, R. (2004). *HLM6: Hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.
- Rudman, L. A., & Glick, P. (2001). Perspective gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743–762.
- Schmader, T., Whitehead, J., & Wysocki, V. H. (2007). A linguistic comparison of letter of recommendation for male and female chemistry and biochemistry job applicants. *Sex Roles, 57*, 509–514.
- Sheehan, E. P., McDevitt, T. M., & Ross, H. C. (1998). Looking for a job as a psychology professor? Factors affecting applicant success. *Teaching of Psychology, 25*, 8–11.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180.
- Taylor, D. (2007). Employment preferences and salary expectations of students in science and engineering. *Professional Biologist, 57*, 175–185.
- Tommasi, G. W., Williams, K. B., & Nordstrom, C. R. (1998). Letters of recommendation: What information captures HR professionals' attention? *Journal of Business and Psychology, 13*, 5–18.
- Trix, F., & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse and Society, 14*, 191–220.
- U.S. Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43*, 38290–38315.
- Valian, V. (1998). *Why so slow? The advancement of women*. Cambridge, MA: MIT Press.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–69.
- Watson, C. (1987). Sex-linked differences in letters of recommendation. *Women and Language, 10*, 26–28.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.
- Windall, S. E. (1988). AAAS presidential lecture: Voices from the pipeline. *Science, 241*, 1740–1745.
- Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin, 128*, 699–727.

## Appendix

### Hireability Scale Items

1. How likely would you be willing to hire this candidate?
2. To what extent is this a “top-notch” candidate?
3. Is it likely that this candidate will make an effective academician?
4. How “excellent” is this candidate based on this letter?

Received August 10, 2008

Revision received March 11, 2009

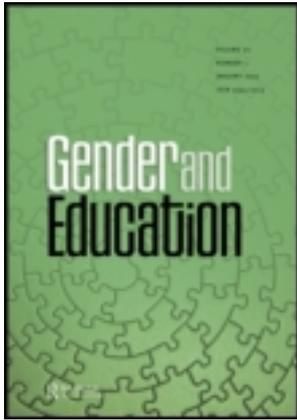
Accepted March 11, 2009 ■

This article was downloaded by: [University of California Davis]

On: 06 March 2014, At: 13:07

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Gender and Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/cgee20>

### Women and science careers: leaky pipeline or gender filter?

Jacob Clark Blickenstaff<sup>a</sup>

<sup>a</sup> Western Washington University, USA

Published online: 15 Aug 2006.

To cite this article: Jacob Clark Blickenstaff (2005) Women and science careers: leaky pipeline or gender filter?, *Gender and Education*, 17:4, 369-386, DOI: [10.1080/09540250500145072](https://doi.org/10.1080/09540250500145072)

To link to this article: <http://dx.doi.org/10.1080/09540250500145072>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Women and science careers: leaky pipeline or gender filter?

Jacob Clark Blickenstaff\*

*Western Washington University, USA*

Women are under-represented in science, technology, engineering and mathematics (STEM) majors and careers in most industrialized countries around the world. This paper explores the broad array of explanations for the absence of women in STEM put forth in the literature of the last 30 years. It is argued that some proposed explanations are without merit and are in fact dangerous, while others do play a part in a complex interaction of factors. It is suggested that the very nature of science may contribute to the removal of women from the 'pipeline'. Recommendations for reform in science education to address this problem are also provided.

## Introduction

A metaphor frequently used to describe the fact that women are under-represented in science, technology, engineering and mathematics (STEM) careers is to propose a 'leaky pipeline' carrying students from secondary school through university and on to a job in STEM. This pipeline leaks students at various stages: students who express interest in science careers sometimes change their minds when applying to colleges and universities and select other areas of study. Others begin their post-secondary education in a STEM program, but change majors before graduation. Finally, some students leave the pipeline after graduating with a STEM degree when they select another field as a career. One interesting feature of these leaks is that women leak out more than men do. The effect of differential leaking is to create a sex-based filter that removes one sex from the stream and leaves the other to arrive at the end of the pipeline. No one in a position of power along the pipeline has consciously decided to filter women out of the STEM stream, but the cumulative effect of many separate but related factors results in the sex imbalance in STEM that is observed today.

There is no dispute that women are under represented in science, technology, engineering and mathematics (STEM) majors and careers<sup>1</sup> (NCES, 2001). Cronin

---

\*Western Washington University, 516 High St., Bellingham, WA, 98225, USA. Email: jacob.blickenstaff@www.edu

and Roger (1999) describe the absence of women in STEM as both progressive (the farther along the pipeline, the fewer women you find) and persistent (the problem has not gone away in spite of treatments). Though the numbers have improved over the last 20 years, there is still a long way to go before women are equally represented.

Some argue that widespread success in technical fields of enquiry is evidence that nothing is wrong with the status quo. There are three compelling reasons to dispute that claim. First, there is the issue of equity. Every person should have an equal opportunity to study and work in the discipline she or he chooses. Second, as long as women are under-represented in STEM, a substantial number of intelligent, talented women are choosing other subject areas in which to study and work. These might be women who could make important contributions to science or engineering if given a chance. Finally, scientific and technical endeavors can only be improved by having a greater diversity of perspectives in the search for knowledge and solutions to human problems. As scientists construct understanding of the world, the ability to see questions and answers from many perspectives will help make scientific explanations more robust and complete.

The challenge of bringing more women into science is not a new problem. Educators have been working for over 20 years to encourage more girls and women to participate in school science through programs like *Girls Into Science and Technology* (GIST) and *Women In Science and Engineering* (WISE) in the UK and *Project Access* in the US (Sonnert, 1995a). Scholars in a wide range of disciplines have paid a great deal of attention to the issue of women and science recently. I have chosen to concentrate on the factors that keep women from studying science or working as scientists. The statistics cited above show that existing efforts to attract women to science have not worked. Why not? What have researchers put forth as explanations for the absence of women in STEM? Which of these explanations are valid and useful in thinking about science teaching? What can and should be done to bring about change?

### Layers in the filter

One of the significant challenges when looking at the literature on girls or women and science is devising a way to organize the very disparate subtopics into meaningful categories. I will be discussing research conducted in many different disciplines and over a span of many decades, so the wide divergence in methodology and focus is perhaps not surprising. I will begin by defining some of the important terms used in this literature, and then explain the taxonomy around which I will organize the remainder of the paper.

When exploring the difference in numerical scores between male and female respondents, the standard statistical measure is the standardized difference in mean scores, or ' $d$ '.  $d$  is calculated by finding the difference between the mean male score and mean female score, and dividing by the average of the male and female standard deviations.<sup>2</sup>  $d$  values of 0 to 0.2 are considered quite small, between 0.2 and 0.5 are small but notable, 0.5 to 0.8 are medium, and  $d$  values greater than 0.8 are considered

to be large (Cole, 1997). Another way of looking at  $d$  is to realize that a  $d$  value of 0.6 means that about 60% of men fall above the combined male–female median, while only about 40% of women fall above that same combined median. Psychometricians often use the  $d$  statistic when designing standardized tests. IQ tests in particular have been carefully manipulated to produce a  $d$  of zero for male and female test takers so that no bias will exist in IQ scores. (This is achieved by using some questions that favor males and others that favor females, which results in no difference in the average scores of men and women.)

In writing this paper I was faced with the challenge of dealing with sex equity in *science* or sex equity in *STEM*. The majority of the research articles I will be discussing deal with science as a particular subject area, rather than addressing all areas of science, technology, engineering and mathematics. A few articles report research in mathematics or engineering courses, and it seems reasonable to include them, since the under-representation of women in those subjects is very similar to the situation in physics and chemistry. ‘Science’ includes both biology and physics, which have very different distributions of men and women as active participants, so in some ways grouping all the sciences together could be a mistake. Because my eventual goal in this paper is to recommend changes in *science* education, my emphasis has been to collect literature focused on science, rather than mathematics, engineering or technology.

I came to this literature review with the idea that women are under-represented in STEM and that the situation ought to be changed. I looked for all the explanations I could find in the literature and attempted to categorize them and evaluate each perspective on its own merits. As a review of every paper written on the topic of gender and STEM would be the work of a lifetime, I have chosen sample papers to describe and analyse in the following pages. Starting from articles in major publications, such as the *Journal of Research in Science Teaching* (Weinburgh, 1995) and the *American Journal of Physics* (Seymour, 1995), I sought out the sources referred to in those initial papers. This broadened my search into the governmental reports, international journals and books treated in this paper.

The subject of women and science has attracted attention from a broad array of researchers, including psychologists, educators, sociologists and scientists. Perhaps because of their widely diverse backgrounds, the many researchers who have tried to understand why women leave STEM careers have proposed widely divergent explanations. I have found the following explanations put forward in the research literature:

1. Biological differences between men and women.
2. Girls’ lack of academic preparation for a science major/career.
3. Girls’ poor attitude toward science and lack of positive experiences with science in childhood.
4. The absence of female scientists/engineers as role models.
5. Science curricula are irrelevant to many girls.
6. The pedagogy of science classes favors male students.

7. A 'chilly climate' exists for girls/women in science classes.
8. Cultural pressure on girls/women to conform to traditional gender roles
9. An inherent masculine worldview in scientific epistemology.

Clearly these explanations lead to very different conclusions about how best to rectify the current situation of women in science, or even if such a remedy is possible. I will argue that some of the above theories hold very little water, while others make some contribution to explaining the under-representation of women in science. As with any complex social situation, in the process of young people choosing a career it is almost impossible to isolate variables completely. Some of these nine explanations interact with each other so intimately that it is impossible to find a single causative factor.

### *Biology*

Psychologists, physiologists and educationalists have explored the biological differences between men and women for over a century. Early attention focused on directly measurable traits like height, arm length, muscle size or the size of the head. Head size was seen as an indicator of brain size, and so indirectly a measure of intelligence. Scientists saw the difference in size of men's and women's brains as the cause of female 'intellectual inferiority' (Sadker & Sadker, 1994; Hyde, 1996). Some researchers even went so far as to compare women's brains to those of gorillas, based on the size and shape of women's skulls (Russet, cited in Birke, 1992). Eventually, the idea of a link between brain size and intelligence was discarded when researchers realized that when corrected for overall body mass, men and women have brains of equal size. Another nail in the coffin was the fact that postmortem examinations showed that some highly respected male geniuses had relatively small brains.

IQ tests have been designed to show no average difference in overall intelligence scores for men and women. But psychologists have broken down intelligence tests into many sub-tests to look at a variety of 'kinds of intelligence' including verbal, mathematical, spatial, and logical abilities. This has enabled researchers to find sex differences in intelligence, resulting in the almost canonical belief in male superiority in mathematical and spatial abilities and female superiority in verbal abilities (Hyde, 1996).

Janet Shibley Hyde conducted two meta-analyses of gender differences in cognitive abilities and summarized her results in Hyde 1996. The only significant differences in ability she found are in mathematical performance ( $d=0.43$ ) and spatial perception ( $d=0.45$ ). Surprisingly, she found the difference in verbal ability was not substantial ( $d=-0.24$ ), in sharp contrast to the conventional wisdom that girls are better with words than boys. The difference she found in spatial perception, while measurable and statistically significant, is not sufficient to explain the degree of under-representation of women in engineering careers. If a spatial ability at or above the ninety-fifth percentile is required to be a successful engineer, the  $d$  values Hyde found would predict that about 7% of men and 3% of women would meet that criterion. So then the degree to which men out perform women in spatial ability could explain a two to one ratio of

males to females in engineering, but not the 20 to one ratio that is observed. Hyde did not find evidence to support any other significant differences in cognitive ability between men and women. Ability may play a small contributing role, but cannot be the only factor causing the under-representation of women in engineering.

Psychologists have explored other intellectual abilities as well. Field dependence research began in the middle of the twentieth century as a way of testing military pilots' ability to maintain an upright position without a visual reference point. Subjects were seated in a dark room in front of a tilted frame and asked to place an illuminated rod vertically within the frame. Those who could accurately place the rod were termed 'field independent' and those who aligned it with the frame were termed 'field dependent'. When sex differences were measured, women were found to be more field dependent than men (Haaken, 1996). Field dependence has been associated with an inability to separate one's self from one's environment, so a person who is field independent has developed a more sophisticated understanding of his/her place in the world. Feminist scholars have criticized the 'field dependence' construct on several points. Both the experimental setting (the male researcher and the female subject alone in a darkened room) and the research task (aligning a rather phallic object within a frame) may have made female subjects uncomfortable and artificially lowered their scores. Early researchers interpretation of scores put women into a no-win situation. If a woman was field dependent, 'Her penis envy seems only to increase her sexual confusion and guilt' (Witkin, cited in Haaken, 1996 p. 297). If instead a woman succeeded at the task and was declared field independent she was described as 'narcissistic with sadistic tendencies' (Haaken, 1996, p. 297). The practical importance of the distinction between field dependence and field independence has also come into question (Haaken, 1996). Haaken also points out that 'field dependence' could be redefined as 'context sensitivity' and that might change how the trait is perceived. While Haaken's work is largely a historical treatment of field dependence research, she chose to criticize the value of the psychological construct, a construct that I believe is open to critique. However, citing research showing that the distinction between field dependent and field independent people is not educationally meaningful would have strengthened Haaken's position.

It is not unusual to hear 'men and women are different, and that's just a fact of life', to explain the current situation for women in society. It seems clear from the literature that whatever biological differences there are between men and women, there is very little difference in scientific or mathematical ability, and certainly not enough to explain the under-representation of women in STEM careers. There is a danger in continuing to emphasize biological differences between men and women because the tendency is to then argue that if unalterable biological differences exist, then no action need be taken to improve the situation for women.

### *Academic preparation*

Some researchers have argued that women avoid science careers not because of any biological difference between men and women, but rather because girls lack the

academic preparation to be successful science students. The Educational Testing Service (ETS) has conducted a meta-analysis exploring gender differences focused on American school-based tests, like the Scholastic Aptitude Test, National Assessment of Educational Progress, Law Scholl Admissions Test, Medical College Admissions Test, etc (Cole, 1997). The report emphasizes the similarities between boys' and girls' performance and minimizes the differences. Those differences that are reported are that girls tend to be better at writing while boys tend to be better at electricity and mechanical subjects. Cole also found that differences increased with age, that is, tests given in primary school showed very similar results for boys and girls, while secondary school tests showed more substantial differences. This paper is open to criticism, however. When counting topics, verbal ability is separated into several areas, all with some advantage to girls, but 'natural science' is one category, possibly obscuring boys' advantage in physics and chemistry. This way of counting subjects serves to minimize boys' achievements and maximize those of girls. It is also possible that a corporation which exists to administer standardized tests has a vested interest in proving tests are a fair and reliable way of assessing students and predicting their future achievement. A report emphasizing the differences in performance of males and females could have had negative consequences for ETS if it could be seen to support testing bias.

Apart from girls' achievement on standardized tests, it is important to note that many girls do avoid physics and calculus courses in high school and that makes it less likely they will choose a STEM major in college, and that in turn will reduce the number of women earning technical degrees. That, however, is a separate issue from the attrition of women from STEM majors once they have begun. Strenta (1994, cited in Erwin & Marutto, 1998) found that regardless of course grades, women in computer science and math courses were more likely to drop out of college than men. Stewart (1998) found that of those students choosing to study physics at A level (advanced study in the UK that prepares students for university) the females were better prepared than the males, as measured by their course grades and scores on national exams. In spite of their strong preparation, girls still end up leaving science

Brainard and Carlin (1998) conducted a longitudinal study of the persistence of women in undergraduate science and engineering majors at the University of Washington between 1991 and 1996. Using annual surveys of female science and engineering majors, they tracked several cohorts of students through the university to try to understand why some women leave science or engineering and others stay. Students reported a number of barriers blocking their path toward a STEM degree which I will address further in the next section of this paper. Interestingly, there was no difference in performance (as measured by GPA) between those women who remained and those who transferred out of STEM programs.

It seems then, that even when women are equally or better prepared than men for scientific or technical majors, they still drop out of the programs at greater rates. If very well prepared women are still leaving STEM, then there must be other factors causing their departure. Making sure that young women who graduate from high

school are equally well prepared to study science as the young men is a worthy goal, but that alone will not solve the problem of under-representation in higher education and careers.

### *Attitude and early experiences*

If biology and preparation are not substantial factors preventing women from becoming scientists, perhaps the problem is that girls just dislike science or math too much to consider a STEM career. A good deal of research has been done to explore girls' attitudes toward science. Molly Weinburgh (1995) conducted a meta-analysis of attitude research to determine if there is a pattern of difference between male and female feelings about science. Over all, she found that boys do have a more positive attitude toward science than girls, but the effect size is relatively small ( $d=0.2$ ). The effect size was especially small in biology ( $d=0.03$ ), as was expected, since other researchers have noted that girls seem to have more positive attitudes toward biology than other sciences. The largest effect sizes were in general science and earth science ( $d=0.34$ ). The small effect size Weinburgh found in physics ( $d=0.12$ ) was more of a surprise, since physics is perceived as a particularly masculine science. This result may be explained by the fact that the girls who choose to take physics probably have a more positive attitude toward science than girls who do not. Weinburgh also found a moderate correlation between attitude and achievement in science. One of the limitations to meta-analysis papers like this one is that the original studies may not all have been of equal quality, but their results are all given equal weight. No information about the quality of the input research is given in this paper.

In contrast to Weinburgh's quantitative exploration of girls' attitudes, other researchers have come at the question with a qualitative research strategy. In an especially well-designed study, Dale Baker and Rosemary Leary (1995) explored girls' ideas about and attitudes toward science by interviewing 40 girls in primary and secondary grades. In general, they found that girls enjoyed their science experiences in school, but had strong negative feelings about dissection and could not imagine themselves as scientists. The girls also noticed the bias in textbooks and television where very few scientists are depicted as female. Some of the girls made an intriguing distinction between a 'scientist' who studies biology or zoology and a 'scientist scientist' who uses chemicals or works with rockets (p. 18). One of the most common reasons the girls gave to explain their interest in life science as opposed to physical science was their desire to care for people or animals. Some of these issues could be seen as 'public relations' problems for science as a discipline, but I think the 'maleness' of science goes deeper than that, and I will return to this idea in the final section of this paper. Baker and Leary clearly wanted to give female students a chance to verbalize their feelings about school science. Both their method (interviews) and their article, which contains many extended quotes from the students show readers how highly the researchers value children's voices.

Liz Whitelegg (2001) conducted a literature review to explore why females lose interest in school science in the UK. She found that many studies only looked at the

girls in science problem at the secondary level, when science has been separated out as a distinct subject area. Whitelegg was interested in what happens in primary school where science is part of the general curriculum, and is more integrated with language arts and math. Studies of primary school classrooms showed that boys and girls tended to play in gender-segregated groups, and to interact with building toys (like Lego) quite differently. Boys also tended to have greater access to classroom resources than girls, in part because the boys aggressively took control of the materials. In light of these behavior differences, Whitelegg argues that a 'hands off' equal opportunity approach is not enough to generate positive science experiences for girls. At the same time, she was hopeful that the National Curriculum in the UK would help increase the participation of girls in science, as it requires all students study science through age 16. The most recent examination results from summer 2003 show that essentially equal numbers of girls and boys take the science double award exam at GCSE, and girls outperform boys by a small margin<sup>3</sup> (Joint Council for General Qualifications, 2003). As Stewart (1998) found, however, the trend is that as soon as girls have a choice they opt out of science courses.

Clearly there are some important differences between boys' and girls' attitudes toward science. Some of the reasons girls express discomfort with science can be addressed by altering curriculum materials and pedagogy, two areas I will be discussing in later sections. Primary school teachers also have an impact on the opinions children have of science through the activities students are asked to complete. Girls' attitudes toward science seem to be closely tied to other factors that make the subject unattractive, so making some effort to improve girls' views of science could pay off in achievement and retention.

### *Role models*

Men make up the majority of scientists and engineers in most industrialized countries, and even though the percentages vary from one field to another, the overall pattern is clear. This means that there are few role models in science, math, or engineering departments for young female students to follow. A related idea is the concept of a 'critical mass' below which there are not enough people of a particular type to maintain a viable population. A low proportion of women in a discipline probably sends a message to girls that the discipline is unattractive to women, and they should avoid it too. At the same time, simply increasing the number of women working in STEM may not be enough. If the only way that women have been able to 'infiltrate' these professions is through acting as similar to their male colleagues as possible, other women will only gain access through following the typical 'male model'.

The issue of role models seems to interact with others, like the 'chilly climate' since a dearth of female engineers or scientists would likely correspond to a poor environment for those women who are in the department. Eileen Byrne (1993) points out that if a plant doesn't succeed in a garden we ask what it is about the soil, water, sun or fertilizer that is causing the problem, we don't blame the plant first (p. 49).

Looking at issues like the presence of role models in science is a way of looking at the environment that girls and women encounter as they learn science, rather than blaming female students for their situation. At the same time, simply increasing the number of female scientists may not attract more young girls or women to science. Sonnert (1995b) points out that successful female scientists often did not have children, and that's a choice that many young people are unwilling to make, so these practicing scientists cannot provide a model of how to balance a career and a family.

It is very difficult to determine what factors actually contribute to final career choices, even of people who complete science Ph.D.s. Gerhard Sonnert interviewed a large number of science graduates who had earned prestigious postdoctoral positions across the US and found that:

Science careers appeared to be shaped, to a considerable extent, by numerous idiosyncratic events and characteristics that are often insignificant by themselves but become forceful in their accumulation. (Sonnert, 1995a, p. 123)

The events that lead someone to a particular job at a particular institution are too specific to create a map easily that others could follow. It seems that the presence of role models in a science or engineering department would be unlikely to fix the problem of under-representation of women, but could be one part of a solution.

### *Curriculum materials and design*

In the 1970s attention began to be paid to the problem of sex bias in textbooks. Researchers looked at the numbers of males and females depicted in illustrations and photographs in science texts, and found that a majority of the people depicted were male (Walford, 1981). Not only were girls and women largely absent, the pictures they were in frequently showed them as passive observers of science being done by boys. One text used in the UK contained four photographs of women, and three of those four showed women sunbathing or swimming, which also is problematic (Walford, 1981). The negative attention this research brought seems to have had an effect on textbook publishers and authors. The representations of males and females are much more balanced today, though David Sadker has noted that line drawings seem to be lagging behind photographs in the reform (personal communication, May 7, 2001). Other ways that sex bias can show up in texts include the wording of problems and examples, and in the attention paid to the contributions of female scientists. Texts do seem to have improved in these areas as well. It is somewhat difficult to be sure, however, since I was able to find very little recent research in this area. It is interesting to note that a good deal of the current textbook bias work has to do with race and ethnicity, not sex.

In addition to textbooks ignoring women and girls, research has shown that some curricula are more successful with female students than others. Robert Tai and Philip Sadler (2001) studied the performance of 1500 students in 16 different introductory physics courses across the US. They found that women were more successful in introductory physics (as measured by final course grade) at the college or university level

if they had taken a high school physics course that emphasized depth of coverage rather than breadth. Also, when men and women with equivalent backgrounds were compared, the women were more successful than the men in college (algebra based) physics and men were more successful than women in university (calculus based) physics. Both of these results indicate that elements of course design—depth vs. breadth and mathematical level—affect men and women differently. Using the final course grade as a measure of science performance has its drawbacks, as different faculty use different criteria in course grades, and grades may not correlate to long-term learning. However, grades may be the only achievement measure that is widely available. It is worth noting here that these re-designed science classes may be effective for other groups who have historically been unsuccessful in science, including racial and ethnic minorities.

Progress has been made in the effort to eliminate sex bias in school textbooks, but there remains room for improvement, particularly in noting the contributions of little-known women scientists. There is also evidence that science curricula could be made more accessible to girls and women by covering less material in greater depth.

### *Pedagogy*

The perception among some teachers and students is that science is simply a boys' subject, there is nothing wrong with that and there is nothing to be done about it (Kelly, 1987). Some might dismiss that attitude as old fashioned, and unimportant since today teachers would claim that they treat boys and girls fairly in their classrooms. This kind of attitude does have an impact on students, however. Margaret Spear found that given identical samples of student work in science, teachers in the UK marked it higher if it appeared to come from a male student than if it came from a female student (Spear, 1987). In this study, Spear gave a set of photocopied papers to teachers who were asked to grade them and provide written comments. Equal numbers of male and female papers were given to each teacher, but the sex-cue labels were switched on half of the pages. Boy's work was consistently rated better in 'scientific accuracy', 'organization of ideas' and 'conciseness' while girls' papers were rated higher than boys' on 'neatness'. In a separate study, Spear found that teachers thought science preparation was more important for boys than for girls (Spear, 1987). At least some teachers still hold old-fashioned ideas about who can be successful in science and these attitudes affect their teaching. It is certainly possible that a similar study conducted today might find different results, since attitudes about boys' and girls' career options have expanded in the last decade. At the same time, many of the teachers who were working in 1987 are still active teachers, and their views are unlikely to have changed substantially.

Elaine Seymour (1995) also found that the quality of teaching in science classes has a substantial impact on the persistence of university science majors. By interviewing a broad cross section of science majors (and students who have left science programs) at over a dozen universities across the US, Seymour was able to determine that pedagogy was a concern of over 90% of students who chose to switch out of science.

Students felt that science instructors were unapproachable and distant, and the over-use of competitive grading systems prevented science students from working collaboratively to improve their understanding. Nearly three-quarters of non-switchers cited concerns about the quality of science teaching as well, indicating that it is a problem for people who remain in science programs as well as those who leave.

Another way that teachers seem to affect student attitudes toward school is their pattern of interactions with students. Researchers have found that girls generally receive less attention from teachers than boys regardless of the subject or age of students (Wilkinson & Marrett, 1985; Sadker & Sadker, 1994). Student–teacher interactions are qualitatively different for boys and girls as well. While boys are asked follow up questions and comments on the ideas represented in their work, girls are more frequently complimented for their looks or for the neatness of their work (Spear, 1987; Eccles & Blumenfeld, 1985; Sadker & Sadker, 1994). Eccles and Blumenfeld also found that though girls asked more questions than boys, teachers gave less feedback (positive or negative) on the girls' answers. Other researchers have noted that boys in science classes call out responses more frequently than girls, and that boys receive more disciplinary attention than girls (Morse & Handley, 1985). Morse and Handley also noted that the differences in participation between boys and girls were more pronounced when children were in Grade 8 than when the same students were in Grade 7. They do not attempt to explain this result, but I suspect that the difference in subject matter may play a role, since the Grade 7 class was on biology, and the Grade 8 class was on earth science. As I will discuss in greater detail in a later section, science disciplines seem to be on a femininity/masculinity continuum which makes biology more open to girls and women.

Though these problems are not limited to science classes, some research has found them to be particularly pronounced in science and math classes, where teachers are more likely to be men, and the majority of students are male (Sadker & Sadker, 1994; Morse & Handley, 1985). The continuing message appears to be that boys' ideas and participation in class are more important than girls'. It seems likely then, that girls would eventually get the message and leave, heading for an environment that recognizes and values their contributions.

Most research into classroom sex bias is open to the criticism that the researcher enters the setting looking for bias and so finds even if it is not present. This result is likely if the only data source is the researcher's observational notes. However, when observations are combined with video recordings and student interviews, multiple data sources should mitigate potential bias and enable other experts to check the researcher's conclusions against the raw data.

R. G. Hacker is particularly sensitive to eliminating expectancy effects from his research. Hacker (1991) looked at teacher–student interactions in several Australian physical science classrooms and found very little difference based on student gender. Those types of interaction that were different for boys and girls, were found to be biased toward girls. Hacker emphasized that the researcher who coded teacher–student interactions did not know the purpose of the observations, so was not biased to find differences that were not there. Hacker does not include the details of the

observation protocol in this article, which necessitated reading an earlier publication with the full description (Hacker, 1984). Hacker's observation instrument leaves out any interactions that are not directly course related, and only addresses the intellectual functions of the interaction, such as 'identifying problems' or 'interpreting data' without considering any possible social functions of interactions. Some important interactions could be missed using this instrument. Hacker also claims that his observer was able to see and interpret the non-verbal interactions of students throughout an entire classroom with a single ceiling-mounted camera. Based on my experience videotaping in classrooms, this does not seem possible, since a lens that would provide a view of the whole classroom would not provide a detailed picture of any part of the room. If classroom bias is really there, anyone should be able to see it, but I am not convinced that Hacker's observational instrument provides a true picture of science classrooms.

The methods teachers use to teach science to young people clearly have an effect on how students perceive the subject. Science pedagogy can re-enforce girls' negative attitudes about science by devaluing the contributions of female students and over-emphasizing rote learning.

### *Chilly climate*

Recent attention to gender differences in the UK has focused on boys' underachievement on exit exams, particularly in English. When the term 'gender gap' is used in the UK, authors are referring to the substantially higher performance of girls, that is in part caused by a 'laddish anti-learning culture'. This gender gap does not translate into high participation in science by girls, however. Warrington and Younger (2000) still found that girls still encountered science teachers with sexist attitudes and low expectations for their achievement. Teachers tended to be overly generous in their predictions of boys' scores on national exams in science, while tending to underestimate girls' scores on the same tests. Girls complained of boys' bad behavior preventing entire classes from conducting lab experiments or other enriching activities. Teachers sometimes expressed the sentiment that though boys behave badly and don't complete their work, 'Boys frequently present more original work, whereas girls copy sentences from textbooks ...' (p. 505). There was also a strange admiration for boys who did not do their work, but still managed to pass exams, while girls who consistently did their assignments lacked 'sparkle'. It is small wonder that girls dislike science and choose other subjects to study at university when their efforts are consistently devalued.

The Brainard and Carlin (1998) longitudinal study described earlier also revealed evidence of a chilly climate in university science. The researchers found that women cited barriers blocking the route to their degrees. (While 20% of freshman reported no barriers, only 3% of seniors could say the same.) These included feelings of isolation and intimidation, as well as a loss in self-confidence as they progressed through their major program. Those women who left the science or engineering track frequently cited poor teaching and their own lack of interest in the subject as their

reasons for leaving. It would be interesting to know more about why some women lost interest in the subject, but this study did not look into that question. Those women who changed majors out of science and engineering had lower self-confidence ratings than their peers who stayed even though both groups had the same GPA.

While the climate for women can be merely cool if they do not feel welcomed by their male peers, at times the atmosphere becomes openly hostile in cases of sexual harassment. Gerhard Sonnert (1995b) reports that of the 191 female fellowship recipients surveyed, 12% had been sexually harassed during their graduate school or early professional experience. One woman reported that on the day of her tenure meeting, a male colleague came to her office and said 'Go to bed with me at noon or I'll vote against you at four'. She received tenure in spite of throwing him out of her office, but there is no indication that she reported his behavior.

The 'chilly climate' for girls and women that seems to exist in many science classrooms I think is largely comprised of the sexist course materials and poor pedagogy that I discussed in prior sections. At the same time, the possibility (some would say probability) of harassing behavior toward women by male colleagues and professors cannot be ignored, and has to be addressed by society at large.

### *Pressure to fill gender roles*

Boys and girls begin learning how to be men and women almost as soon as they are born. Adults teach children how to be a 'grown up' through explicit lessons and through daily interactions with each other. Barrie Thorne deals extensively with the socialization of boys and girls in *Gender play: girls and boys in school* (1993). Thorne describes the process of girls and boys separating into 'opposite sides' both in the classroom and on the playground. Teachers often re-enforce this taking of sides by seating girls on one side of the classroom, and boys on the other, or by setting up class competitions of boys vs. girls. The idea that boys and girls are on opposite sides seems to also infect children's ideas of appropriate career goals and aspirations. Women who have become physicists report that one of the major obstacles in their path on the way was the expectation that they would also be the primary caregiver for their children (Ivie, Czujko, & Stowe, 2002).

As part of their literature review exploring the cause of differences in mathematics performance of boys and girls, Fennema and Peterson (1985) point to the effect of sex-role identity on student achievement.

Mathematics is not seen as an appropriate domain for females. Therefore, achievement by a female in the mathematical domain results in her not fulfilling her sex role identity adequately. She perceives that teachers and peers have lowered expectations of her mathematical success because she is female. ... others see her as somewhat less feminine when she achieves in mathematics ... she thinks others have negative feelings about her success' (Fennema & Peterson, 1985, p. 25)

While it is possible that students' and teachers' perceptions of girls and mathematics have changed since 1985, I think this is still an important issue. Students generally live up to the expectations of their teachers, and want to fill the roles that are expected

of them, so math and science teachers should be taught to have equally high expectations for their male and female students.

*The masculine worldview of science*

There is little argument that at times in the past, scientific theories have been used to further questionable political agendas. Scientific research was done to support White male scientists' ideas that Blacks and women had inferior mental capacity, either because of head size or facial structure. Women were excluded from medical research trials because doctors thought the menstrual cycle would invalidate their results. Men and women with low IQ scores were sterilized to prevent further damage to the gene pool. Virtually all scientists working today would dismiss such work as 'bad science'. Many then argue that scientists have learned how to be more objective today, and therefore science is no longer biased toward the masculine. Future scholars may be just as critical of our science as we are of some work from the late nineteenth century. (Scientific apologists for the tobacco and oil industries come to mind.) Another idea is that science is inherently masculine in its structure, epistemology and methodology. This is the position taken by feminist critics of science like Evelyn Fox-Keller, Jane Gilbert, Sandra Harding, Donna Haraway, Helen Longino and others.

Many science educators seem a bit afraid of feminist science because it challenges the foundations of science, so many literature reviews leave this material out. It is naïve to pretend that science is above criticism, and this material can provide information about how to go about reforming the *science* part of science education. Some argue that women are unsuccessful in science because science itself is a masculine(ist) enterprise. Science is based upon positivist objective rationality, which is seen by some feminist scholars as entirely masculine, and therefore closed (or at least unattractive) to women. There are three dominant perspectives in this category (adapted from Harding, 1991; Kerr, 2001):

- *Feminist standpoint theory*—Sandra Harding argues that science should be done from the perspective of women, because their position outside the dominant social order (as mothers and caregivers) endows women with a more objective view of the world than men have. The vision of those in powerful positions in society is clouded by their power, while the powerless have a less distorted view. Harding argues that knowledge claims should be judged upon the social context of their production.
- *Situated knowledge*—while agreeing with the idea that people have different points of view from which they see the world, Donna Haraway (1991) believes that everyone takes on different standpoints at different times, and people can temporarily take on the standpoint of others. Haraway conceptualizes varied research perspectives as different ways of seeing and argues that every view has both blind spots and focal points.
- *Feminist empiricism*—instead of evaluating knowledge claims on their social context, Helen Longino (1990) argues for 'contextual empiricism', connecting truth claims to the available evidence, but also recommending that scientists use

their political beliefs to guide their theoretical positions. Instead of trying to avoid political bias, scientists should acknowledge the biases they have and work with from their political position.

Jane Gilbert (2001) argues that since science comes out of enlightenment rationality it is by its very nature masculine. She uses the construct of 'A'/'not A' to set up binary opposites, like rational/emotional, objective/subjective, science/nature, male/female that support her position on the masculinity of science. Gilbert asserts that all of the 'A' qualities are associated with men and science, and all the 'not A' qualities are associated with women. People are taught to value the rational over the emotional, and scientists in particular are trained to divorce their emotions from their work.

Many women respond to these feminist critics of science by pointing out that the above arguments over-emphasize the commonality of women's experience, and the importance of motherhood in the lives of women. Not all women become mothers, and a woman's identity should not be defined first by her participation (or lack of participation) in biological reproduction. Many women who have become successful scientists are quite vehement in their opposition to feminist critiques of science. It is also notable that few feminist critics of science openly discuss the absence of racial or ethnic minorities in their groups. The extent to which sympathetic men can participate in feminist research is another question I find particularly interesting, but that question is beyond the scope of this paper.

Some science disciplines are more masculine<sup>4</sup> than others; biology and other life sciences are at one end of the spectrum, while physics and engineering are at the other end. Esther Saraga and Dorothy Griffiths (1981) point out the sciences most strongly identified as masculine are those that are closely tied to improving economic production and developing weapons, two tasks that male-dominated society has decided are valuable. Biology has historically been less clearly tied to economic development (except in agricultural science, another male province) and weapons than physics or engineering. It will be interesting to see if the perception of biology changes as a result of 'bio-terrorism' becoming a threat to society.

The important message I take from all the critics of masculine science is that political and social power affects the kind of questions that scientists ask and how scientists interpret the answers they obtain. If only one kind of person asks the questions and interprets the results, then the field of scientific inquiry will be narrow and inbred. Science can be improved by broadening the diversity of its practitioners across gender, ethnic and racial lines and science education can be improved by acknowledging the political nature of scientific enquiry.

### **Soluble problems**

Many attempts have been made to explain why women are under-represented in STEM. I believe that I have shown above that some of these explanations are without merit, while others are part of a complex interaction of factors that tend to push girls and women away. Cronin and Roger (1999) point out that initiatives to bring women

and science together focus on one of three areas: attracting women to science, supporting women already in science, or changing science to be more inclusive of women. Some initiatives have emphasized one or two of these possibilities and ignored the other(s). The following are my suggestions to ameliorate the under-representation of women in STEM:

1. Ensure students have equal access to the teacher and classroom resources.
2. Create examples and assignments that emphasize the ways that science can improve the quality of life of living things.
3. Use cooperative groups in class, or at least avoid dividing students by sex for class competitions or in seating arrangements.
4. Eliminate sexist language and imagery in printed materials.
5. Do not tolerate sexist language or behavior in the classroom.
6. Increase depth and reduce breadth in introductory courses.
7. Openly acknowledge the political nature of scientific inquiry.

Clearly my emphasis is on improving science and science education, rather than attempting to 'remediate' girls and women. By implementing the above changes to current classroom practice, schoolteachers would provide an environment where female students will be more likely to succeed in subjects that have traditionally been dominated by males.

## Conclusion

We believe that the relationship of girls to science, and their performance in it, are too complex to be understood in terms of one factor, but that several factors must be integrated in a broader understanding of the social context in which science is carried out, and in which socialization takes place. ... Theories couched in biological terms cannot be sustained ... it is not sufficient just to consider the development of girls in relation to science—the development and practice of science must also be discussed. (Saraga & Griffiths, 1981, p. 85)

The problem of female under-representation in STEM careers and majors is certainly not a simple one. The factors that remove women from the STEM pipeline can be seen as layers in a sex-based filter, though no single issue can be called the primary cause. A complex problem like this requires a multi-faceted solution, and time to allow innovations to take effect. I believe that science educators have a responsibility to change those factors under their control. Over time, individual actions by sympathetic teachers will help women to break down the filter in the STEM pipeline and result in equal participation, which will be good for STEM and good for society in the long term.

## Notes

1. Eighty per cent of engineering and 60% of physical science Bachelor's degrees in the US were awarded to men in 1999–2000. At the Ph.D. level 84% of engineering degrees and 75% of the physical science degrees went to men. Even in biology, where 58% of the Bachelor's degrees were awarded to women, 56% of the Ph.D.s went to men.

2. The equation looks like this:

$$d = \frac{M_m - M_f}{\text{Average}(SD)}$$

3. 55.1% of girls earned a score of C or higher, while 52.5% of boys achieved the same results.
4. I use 'masculine' to refer to both the numerical dominances of men in the practice of the science, and in the perception of the science by the general public.

## References

- Baker, D. & Leary, R. (1995) Letting girls speak out about science, *Journal of Research in Science Teaching*, 32(1), 3–27.
- Birke, L. (1992) In pursuit of difference: scientific studies of women and men, in: G. Kirkup & L. S. Keller (Eds) *Inventing women: science, technology and gender* (Cambridge, MA, Blackwell), 81–102.
- Brainard, S. G. & Carlin, L. (1998) A six-year longitudinal study of undergraduate women in engineering and science, *Journal of Engineering Education*, 87(4), 17–27.
- Byrne, E. M. (1993) *Women and science: the snark syndrome* (London, The Falmer Press).
- Cole, N. S. (1997) *The ETS gender study: how females and males perform in educational settings* (Princeton, NJ, Educational Testing Service).
- Cronin, C. & Roger, A. (1999) Theorizing progress: women in science, engineering, and technology in higher education, *Journal of Research in Science Teaching*, 36(6), 639–661.
- Eccles, J. S. & Blumenfeld, P. (1985) Classroom experiences and student gender: are there differences and do the matter?, in: L. C. Wilkinson & C. B. Marrett (Eds) *Gender influences in classroom interaction* (New York, Academic Press Inc), 79–114.
- Erwin, L. & Maurutto, P. (1998) Beyond access: considering gender deficits in science education, *Gender and Education*, 10(1), 51–69.
- Fennema, E. & Peterson, P. (1985) Autonomous learning behavior: a possible explanation of gender-related differences in mathematics, in: L. C. Wilkinson & C. B. Marrett (Eds) *Gender influences in classroom interaction* (New York, Academic Press Inc), 17–35.
- Gilbert, J. (2001) Science and its 'other': looking underneath 'woman' and 'science for new directions in research on gender and science education, *Gender and Education*, 13(3), 291–305.
- Haaken, J. (1996) Field dependence research: a historical analysis of a psychological construct, in: B. Laslett, S. G. Kohlstedt, H. Longino & E. Hammonds (Eds) *Gender and scientific authority* (Chicago, University of Chicago Press), 282–301.
- Hacker, R. G. (1984) A hierarchy of intellectual development in science, *British Journal of Educational Psychology*, 54, 137–151.
- Hacker, R. G. (1991) Gender differences in science-lesson behaviors, *International Journal of Science Education*, 13(4), 439–445.
- Haraway, D. (1991) *Simians, cyborgs, and women: The reinvention of nature* (New York, Routledge).
- Harding, S. (1991) *Whose science? Whose knowledge? Thinking from women's lives* (Milton Keynes, Open University Press).
- Hyde, J. S. (1996) Meta-analysis and the psychology of gender differences, in: B. Laslett, S. G. Kohlstedt, H. Longino & E. Hammonds (Eds) *Gender and scientific authority* (Chicago, University of Chicago Press).
- Ivie, R. Czujko, R. & Stowe, K. (2002) *Women physicists speak: the 2001 international study of women in physics* (Washington, DC, American Institute of Physics).
- Ivie, R. & Stowe, K. (2000) *Women in physics 2000* (American Institute of Physics Rep. No. R-430) (Washington, DC, American Institute of Physics).

- Joint Council for General Qualifications (2003) *Results 2003*. Available online at: <http://www.jcgq.org.uk/> (accessed 20 October 2003).
- Kelly, A. (1987) The construction of masculine science, in: A. Kelly (Ed.) *Science for girls?* (Milton Keynes, Open University Press), 66–77.
- Keller, E. F. (1985) *Reflections on gender and science* (New Haven, CT, Yale University Press).
- Keller, E. F. (1996) Feminism and science, in: E. F. Keller & H. E. Longino (Eds) *Feminism and science* (New York, Oxford University Press), 1–16.
- Kerr, E. A. (2001) Toward a feminist natural science: linking theory and practice, in: M. Lederman, & I. Bartsch (Eds) *The gender and science reader* (New York, Routledge), 386–406.
- Longino, H. E. (1990) *Science as Social knowledge: values and objectivity in Science Inquiry* (Princeton, NJ, Princeton University Press).
- Morse, L. W. & Handley, H. M. (1985) Listening to adolescents: gender differences in science classroom interaction, in: L. C. Wilkinson & C. B. Marrett (Eds) *Gender influences in classroom interaction* (New York, Academic Press Inc), 37–56.
- National Center for Education Statistics (2001) Postsecondary Institutions in the United States: Fall 2000 and degrees and other awards conferred 1999–2000 (Washington, DC, National Center for Education Statistics).
- Rosser, S. V. (1990) *Female-friendly science: applying women's studies methods and theories to attract students* (New York, Pergamon).
- Sadker, M. & Sadker, D. (1994) *Failing at fairness: how our schools cheat girls* (New York, Simon & Schuster).
- Saraga E. & Griffiths, D. (1981) Biological inevitabilities or political choices? The future for girls in science, in: A. Kelly (Ed.) *The missing half: girls and science education* (Manchester, Manchester University Press), 85–97.
- Seymour, E. (1995) Why undergraduates leave the sciences, *American Journal of Physics*, 63(3), 199–202.
- Sonnert, G. with Holton, G. (1995a) *Gender differences in science careers: the project access study* (New Brunswick, NJ, Rutgers University Press).
- Sonnert, G. with Holton, G. (1995b) *Who succeeds in science?* (New Brunswick, NJ, Rutgers University Press).
- Spear, M. G. (1987) The biasing influence of pupil sex in a science marking exercise, in: A. Kelly (Ed.) *Science for girls?* (Milton Keynes, Open University Press), 46–51.
- Stewart, M. (1998) Gender issues in physics education, *Educational Research*, (40)3, 283–293.
- Tai, R. H. & Sadler, P. M. (2001) Gender differences in introductory undergraduate physics performance: university physics versus college physics in the USA, *International Journal of Science Education*, 23(10), 1017–1037.
- Thorne, B. (1993) *Gender play. Girls and boys in school* (New Brunswick, NJ, Rutgers University Press).
- Walford, G. (1981) Tracking down sexism in physics textbooks, *Physics Education*, 16, 261–265.
- Warrington, M. & Younger, M. (2000) The other side of the gender gap, *Gender and Education*, (12)4, 493–508.
- Weinburgh, M. (1995) Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991, *Journal of Research in Science Teaching*, (32)4, 387–398.
- Whitelegg, L. (2001) Girls in science education: of rice and fruit trees, in: M. Lederman, & I. Bartsch (Eds) *The gender and science reader* (New York, Routledge), 373–382.
- Wilkinson, L. C. & Marrett, C. B. (1985) *Gender influences in classroom interaction* (New York, Academic Press Inc).

# Understanding current causes of women's underrepresentation in science

Stephen J. Ceci and Wendy M. Williams<sup>1</sup>

Department of Human Development, Cornell University, Ithaca, NY 14853

Edited by Richard F. Thompson, University of Southern California, Los Angeles, CA, and approved December 6, 2010 (received for review October 6, 2010)

Explanations for women's underrepresentation in math-intensive fields of science often focus on sex discrimination in grant and manuscript reviewing, interviewing, and hiring. Claims that women scientists suffer discrimination in these arenas rest on a set of studies undergirding policies and programs aimed at remediation. More recent and robust empiricism, however, fails to support assertions of discrimination in these domains. To better understand women's underrepresentation in math-intensive fields and its causes, we reprise claims of discrimination and their evidentiary bases. Based on a review of the past 20 y of data, we suggest that some of these claims are no longer valid and, if uncritically accepted as current causes of women's lack of progress, can delay or prevent understanding of contemporary determinants of women's underrepresentation. We conclude that differential gendered outcomes in the real world result from differences in resources attributable to choices, whether free or constrained, and that such choices could be influenced and better informed through education if resources were so directed. Thus, the ongoing focus on sex discrimination in reviewing, interviewing, and hiring represents costly, misplaced effort: Society is engaged in the present in solving problems of the past, rather than in addressing meaningful limitations deterring women's participation in science, technology, engineering, and mathematics careers today. Addressing today's causes of underrepresentation requires focusing on education and policy changes that will make institutions responsive to differing biological realities of the sexes. Finally, we suggest potential avenues of intervention to increase gender fairness that accord with current, as opposed to historical, findings.

women in science | gender bias | child penalty | peer review

Since 1970, women have made dramatic gains in science. Today, half of all MD degrees and 52% of PhDs in life sciences are awarded to women, as are 57% of PhDs in social sciences, 71% of PhDs to psychologists, and 77% of DVMS to veterinarians.\* Forty years ago, women's presence in most of these fields was several orders of magnitude less; e.g., in 1970 only 13% of PhDs in life sciences went to women (1). In the most math-intensive fields, however, women's growth has been less pronounced (2–4). Among the top 100 US universities, only 8.8–15.8% of tenure-track positions in many math-intensive fields (combined across ranks) are held by women, and female full professors number  $\leq 10\%$ . (*SI Text, S1*)

These figures reveal a problem, but what is its cause? Here, we consider one of the most common alleged causes—discrimination against women in the domains of: (i) manuscript reviewing, (ii) grant funding, and (iii) interviewing/hiring. We reprise the evidence for each and describe counterevidence. We conclude that past initiatives to combat discrimination against women in science appear to have been highly successful. Women's current underrepresentation in math-intensive fields is not caused by discrimination in these domains, but rather to sex differences in resources, abilities, and choices (whether free or constrained). Thus, current initiatives direct energy toward solving past problems rather than current ones. Women's underrepresentation today results from a complex set of interrelated factors, some of which society could meaningfully address if the focus was placed squarely on them. One key to such success is moving beyond historical issues and confronting current ones.

## Claims of Discrimination Against Women Scientists

Recent scientific reports often assert that discrimination against female scientists in hiring, publishing, and funding is a cause of their underrepresentation:

“Substantial research shows that resumes and journal articles were rated lower by male and female reviewers when they were told the author was a woman; similarly, a study of postdoctoral fellowships awarded showed that female awardees needed substantially more publications to achieve the same competency rating as male awardees” (5, p. 1933).

“It is now recognized that biases function at many levels within science including funding allocation, employment, publication, and general research directions” (6, p. 1247).

“Research has pointed to bias in peer review and hiring. For example, Wennerås and Wold found that a female postdoctoral applicant had to . . . publish at least three more papers in a prestigious science journal or an additional 20 papers in lesser-known specialty journals to be judged as productive as a male applicant. . . . The systematic underrating of female applicants could help explain the lower success rate of female scientists in achieving high academic ranks” (7, p. 24).

“An impressive body of controlled experimental [research]. . . shows that, on the average, people are less likely to hire a woman than a man with identical qualifications, are less likely to ascribe credit to a woman than to a man for identical accomplishments. . . .” (8, p. S2).

Such claims of discrimination against women are consistent with claims of glass ceilings, reduction of authorship credit and pay for comparable work, smaller laboratory space, and fewer research resources (9–11). For example, economists analyzing auditions for orchestras found that switching to blind auditions in which juries could not see applicants reduced discrimination against women, explaining one-third of the increase in the proportion of women hired after blind auditions (12). Other examples of discrimination in nonmath fields are similarly striking, e.g., correlations between masculineness of women's first names and likelihood of being awarded judgeships (13), downgrading of psychologists' and sociologists' curriculum vitae when they bear a woman's name (14, 15), or discriminatory pay for female attorneys (16). This evidence from nonmath fields raises the specter that similar biases explain the current dearth of women in math-intensive fields. Below, we describe empirical evidence for claims of discrimination in the domains of publishing, grant reviewing, and hiring. We find the evidence for recent sex discrimination—when it exists—is aberrant, of small magnitude, and is superseded by larger, more sophisticated analyses showing no bias, or occa-

Author contributions: S.J.C. and W.M.W. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence may be addressed. E-mail: [stevececi@cornell.edu](mailto:stevececi@cornell.edu) or [wendywilliams@cornell.edu](mailto:wendywilliams@cornell.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1014871108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1014871108/-DCSupplemental).

\*Feeder pipelines are sometimes even higher than the numbers earning PhDs, indicating some leakage. For example, slightly more than 60% of B.S. degrees in biology are earned by women, so 52% female PhDs represents some pipeline leakage from the 60% of BS degrees.

sionally, bias in favor of women. Although real barriers are still faced by women in science, especially mathematical sciences, our findings suggest that historic forms of discrimination cannot explain current underrepresentation, and that resources should be redirected toward current rather than historical causes of women's underrepresentation in math-based careers.

### Discrimination Against Women in Journal Reviewing

The question of whether sex discrimination exists in getting work published is ideally answerable by examining manuscript acceptance rates of men vs. women, holding constant quality of work. However, quality of work is influenced by available resources. Comparing women and men with comparable resources, we find no sex discrimination in publishing (below). However, a secondary issue is whether resources themselves are, in fact, evenly distributed between the sexes. The answer is that they are not, for a complex constellation of reasons, such as women being more apt to occupy teaching-intensive positions, part-time positions, etc. Thus, the attention devoted to righting perceptions of sex discrimination in reviewing of manuscripts, which as we show, does not in fact exist (*SI Text, S2*), focuses on a spurious issue and detracts from the very real problem that does plague women in publishing—the fact that women more often than men lack resources necessary to produce high-quality work.

Budden and her colleagues published several analyses of gender bias in manuscript reviewing by undergraduates, graduate students, postdocs, and journal reviewers (6, 17–19). In one, they analyzed journal acceptance rates for manuscripts submitted by women to *Behavioral Ecology* after it began blind peer review (18). Acceptance rate for female first-authored manuscripts increased by 7.9% in the 4 y after the onset of blind review, compared with the 4 y prior, suggesting that when reviewers were aware of authors' sex, they were less likely to accept women's manuscripts. Critics argued that the difference between pre- and post-blind-review acceptances cannot be tied to sex discrimination, criticizing this finding on statistical grounds, and noting that increased acceptance rates for women occurred during this period at six other ecology and evolution journals that did not implement blind review and that increases were already apparent in the decade before implementation of blind reviewing at *Behavioral Ecology* (20–22). Whittaker (22) analyzed 1,140 manuscripts submitted to the *Journal of Biogeography* and found no difference in the acceptance/rejection ratio for male and female authors ( $\chi^2 = 1.0637$ ,  $p = 0.3024$ ); he also criticized Budden et al.'s statistical procedures. Others did not replicate Budden et al.'s claim of sex discrimination with other journals (see below). In response, Budden et al. argued such criticisms cannot explain their full findings, although they acknowledged their inability to rule out alternative explanations due to limited data (19).

Less-often cited analyses by Budden et al. did not find evidence of sex discrimination. One (23) compared publication success rates for 2,680 papers with male vs. female first authors at five ecology and evolution journals (with seven editors). There was no sex difference in overall acceptance rates; although one of seven editors' data from one of five journals initially appeared to favor males, this impression was not borne out by more rigorous log-linear analyses, leading to the conclusion that “there is no significant interaction between journal and author gender in their effect on whether manuscripts are accepted for publication” (ref. 23, p. 350). In another study by Budden and her colleagues (17), 989 raters (undergraduates to postdocs) were given an identical manuscript with either no name, author's first initial only, a male name, or a female name. This study also found no sex discrimination. To recap, there is evidence of bias from one of Budden's studies (18), but criticisms have been published on various grounds (20–22); other journals show no such bias (24, 25); and Budden has acknowledged limitations. Further, archival acceptance data from various journals strongly suggest no sex differences (see below).

Another source of evidence is aggregate productivity measures: Are women as likely as men to publish? Many sources of data span numerous fields and cohorts. They show that women are as successful at publishing as men, when comparisons are between men and women with similar resources and characteristics. (Compar-

ing people possessing unequal resources is not a test of discrimination, but a demonstration of limitations imposed by a lack of resources, and women tend to have jobs that provide fewer resources; e.g., teaching- vs. research-intensive positions.) The editors of *Nature Neuroscience* analyzed acceptance rates for 449 authors (24) and found that acceptance rates were statistically indistinguishable between the sexes,  $p = 0.811$ , regardless of how the data were parsed. Other journals (e.g., *Cortex*) have also reported equal acceptance rates (25).

Others have undertaken sophisticated analyses of archival data showing that any sex difference in productivity is due to differences in structural variables that, although correlated with sex, are causally unrelated to it (26–27) because the same fate befalls male authors with similar structural characteristics. Thus, although such variables (e.g., working at teaching-intensive colleges) affect women more often than men, they hinder men's productivity equally. Scientists' publications increase after moving to more prestigious institutions with greater resources (27). In this analysis, males produced 30% more publications than women, but when men tenured at R1 universities were compared with women tenured at R1 universities, the gap fell to 8%, and the difference between men and women full professors at R1s was <5%.

Thus, the critical variable is not sex per se, but rather access to resources, which correlates with sex because women are more likely to work as adjuncts or at teaching-intensive institutions with limited resources (*SI Text, S2*). As evidence, a longitudinal analysis of faculty in 1969, 1973, 1988, and 1993 showed sex differences in productivity steadily declined—from a female:male ratio of 0.580:1 in 1969 to 0.817:1 in 1993 (26). The primary factor affecting women's productivity was structural position. When type of institution, teaching load, funding, and research assistance were factored in, the productivity gap completely disappeared (which is not to say discrimination has not influenced these factors in the real world): “There is very little direct effect of sex on research productivity. . . . men generally have positions superior to those of women, although structural differences by gender have appreciably declined over time. Once sex differences in such positions and resources are taken into account, net differences between men and women in productivity are nil or negligible” (ref. 26, pp. 863–864).

Similarly, a National Research Council task force concluded that productivity of women science and engineering faculty increased over the last 30 y and is now comparable to men's, the critical factor affecting publications being access to institutional resources (28). Finally, many others also report no sex differences in productivity, controlling for structural variables confounded with sex (e.g., refs. 7 and 8).

In sum, when publication data are controlled for structural position, ensuring that sex differences in manuscript acceptance rates are not conflated with sex differences in resources, there is no difference between the sexes (*SI Text, S3*). Although structural differences present real barriers for many women—and some men—journal reviewers do not reject papers because they are written by women. The preponderance of evidence, including the best and largest studies, indicates no discrimination in reviewing women's manuscripts: Given equivalent resources, men and women do equally well in publishing. A key issue, separable from sex discrimination in manuscript evaluation, is why women occupy positions providing fewer resources and what can be done about this situation. This situation is caused mainly by women's choices, both freely made and constrained by biology and society, such as choices to defer careers to raise children, follow spouses' career moves, care for elderly parents, limit job searches geographically, and enhance work-home balance. Some of these choices are freely made; others are constrained and could be changed (3).

### Discrimination Against Women in Grant Funding

Another domain of alleged sex discrimination is grant and fellowship reviews. In an influential article in *Nature* (cited 212 times),<sup>†</sup> Wennerås and Wold (29) reported that when reviewers

<sup>†</sup>As a more important indicator of its influence, this article is frequently invoked as evidence by proponents of interventions such as NSF's \$130 million ADVANCE initiative.

judged postdoctoral fellowship applications to the Swedish Medical Research Council (MRC) in 1995, the conversion of data into subjective scores was highly prejudiced against women. A *Nature* commentary stated: “The response across the world could be measured on the Richter scale after the revelation that the (MRC) exercised prejudice in its allocation of research fellowships. Six months later, the implications are still being discussed in the newspapers and on radio and TV” (ref. 30, p. 204). The claim of discrimination was based on 62 applications submitted by men and 52 by women: 16 men were funded (25.8%) vs. 4 women (7.7%). Evidence of bias was based on analyses of reviewers’ scores versus objective data (e.g., publications, citations). Reviewers judged each applicant on scientific competence, proposal relevance, and methodology. Women received somewhat lower mean scores than men in all three categories, the largest discrepancy being in scientific competence—2.46 vs. 2.21. The total impact score was most predictive of reviewers’ ratings of scientific competence ( $r^2 = 0.47$ ): A woman needed a 2.6 times higher “total impact measure” score than a man to be judged as competent.

To determine whether women and men were judged equally, Wennerås and Wold (29) assumed an applicant’s scientific competence is linearly correlated with number and quality of published journal articles, leading to an examination of total number of publications, total number of first-author publications, total citations, total impact measure, first-author impact measure, and first-author citations. They reported “a female applicant had to be 2.5 times more productive than the average male applicant to receive the same competence score” (ref. 30, p. 342). The most productive female applicants—those with 100 or more impact points (based on number of publications and how frequently the journals are cited)—was the only group of women judged as competent as men, but only the least productive men who had fewer than 20 impact points. However, the authors’ conclusion of bias against women was challenged on statistical, methodological, and conceptual grounds (*SI Text, S4*). Analyses of funding societies in Europe and North America failed to find bias against female applicants during this same period (1994–1995). One analysis of nearly 3,000 grant applications to the British Wellcome Trust and MRC for the period 1993–1996 revealed no evidence of sex bias in approval rates (31). Another reported no sex bias in grant approvals to the UK Biological Sciences Research Council (32). A third (33) reported results of  $\approx 8,000$  grant applications to the MRC of Canada, also failing to find sex differences. Although there are occasional instances of sex effects in these reports, they are rare, of small magnitude, and are as often in favor of women as against them; the largest aberrations were not close to Wennerås and Wold’s finding (29) that women had to be 2.5 times more productive than men to obtain similar scores.

Consider that Dickson (32) reported that in 1996, the success rates for male and female grant applicants to the UK’s Biotechnology and Biological Sciences Research Council were 24% and 19%, and 26% and 29% to the MRC, respectively. Grant et al. (31) reported that “the award rates for both sexes are approximately the same. . . . Neither is there any evidence that women need a more impressive publication record than men to be successful in either organization’s competitions” (p. 438). In fact, successful female candidates for project grants published on average 11.2 papers vs. 13.8 papers by males; successful female candidates for senior research fellowships published 11.8 papers vs. 14.3 for men.

In Canada, Friesen (33) found that, with one exception, grants were also gender-neutral: For the largest program, the success rates for men and women were 26.6% and 25.4% (n.s.). For the prestigious MRC-Canada Scholarship awards, which provided five years of salary for new PhDs, there was likewise no significant sex difference (14% for men vs. 16.6% for women). The sole sex difference favoring men was for the category of PhD students doing postdoc training and health professionals undertaking research training: 16.3% for men vs. 12.9% for women ( $p < 0.05$ ). This difference is small, and of five competitions, two had virtually identical approval rates. Finally, Sandstrom and Hallsten (34) analyzed more recent data from the Swedish MRC and found that the gender bias reported by Wennerås and Wold (29) had re-

versed itself, so that there was a small but significant effect in favor of funding women’s grants compared to men’s with the same score. They analyzed 280 grant applications in 2004, 118 from female principal investigators, and found no evidence of gender bias: “Surprisingly, none of the productivity measures interact with gender: Male and female PIs are judged similarly with reference to productivity. When we control for all productivity measures and interactions. . . . It appears that female PIs receive a bonus compared to male PIs” (34, p. 186). Perhaps this lack of sex difference is due to Wennerås and Wold’s 1997 paper’s impact, but this possibility does not explain why even larger, more encompassing studies preceding theirs found no sex differences.

Thus, a decade after Wennerås and Wold’s report (29), the Cochrane Methodology Review Group concluded that other than Wennerås and Wold’s study, “a number of other studies carried out in similar contexts found no evidence of (sex discrimination)” (35, p. 2). However, perhaps studies with samples big enough to use the most sensitive measurement framework might unearth sex differences. It is especially useful, therefore, to examine studies meeting this standard. Six large-scale analyses have been published, on net providing compelling counterevidence to sex discrimination claims.

The first large-scale analysis (36) assessed gender bias at the National Science Foundation (NSF), the National Institutes of Health (NIH), and the US Department of Agriculture. This study concluded there was no gender bias in awarding of grants at these three federal agencies (*SI Text, S5*). At all three agencies, which fund tens of thousands of US scientists, men’s and women’s grants were approved at the same rate. [Relatedly, Leboy’s (37) much smaller report of success rates for first-time RO1 grants at the National Institutes of Health for men and women revealed identical success rates for new submissions between 1998–2004 and very similar continuation grants by 2004.] Clearly, female PIs did not have to be 2.5 times better than males to be approved.

The second large-scale analysis is based on the Australian Research Council, which processes >3,000 applications annually in all areas of science. Marsh and his colleagues published several analyses of these data, using sophisticated measurement frameworks (38–40). In a multilevel analysis of 10,023 reviews by 6,233 reviewers of 2,331 proposals (38, 39), they found that although only 15.3% of applicants for grants were female PIs, their success was almost perfectly proportional (15.2%). When sex of only first-named investigator was considered, the success rate was identical: 21% both for men and women. Detailed analyses of second- and third-named researchers also indicated an absence of sex differences in success as did analyses based on mean external ratings and final panel committee ratings (38). No gender effect was found in any of the nine disciplines, and there was no gender bias as a function of reviewer or applicant sex or their interaction. [Subsequently, Marsh et al. (39) extended these results, again reporting no sex differences in approval rates that again generalized across disciplines, as well as to reviewers nominated by PIs vs. chosen by the agency, and to the country of the reviewer.]

The third large-scale analysis provides an exception to the consistent failure to find sex differences. Bornmann and Daniel (41) examined 1,022 applications for predoctoral fellowships and 134 for postdoctoral fellowships to a German Foundation between 1985 and 1990. They found no evidence for gender bias in approving postdoc fellowships, but evidence for bias in approving pre-PhD fellowships, with males more likely to be approved (*SI Text, S6*). A follow-up study by Bornmann and his colleagues (42, 43) used a sophisticated multilevel approach to a metaanalysis, based on their comprehensive collection of studies of peer reviews for grants. This time Bornmann found a difference in favor of men, albeit extremely small—an effect size of <0.04 (an odds ratio of only 1.07:1.00). Only 1 of 66 sex-difference effect sizes was significant. Once again, no bias was found for postdocs, the group studied by Wennerås and Wold (29).

The fourth large-scale study was based on all grant proposals submitted to the Economics Program at NSF during the years 1987–1990. Broder (44) analyzed 6,764 reviews. Consistent with the dearth of women economists in the late 1980s, only 9.3% of the PIs and 7.6% of the reviewers were women. However, Broder found

female PIs fared well when rated by male reviewers at NSF, but less well when rated by female reviewers, a finding she suggested may have worked against increased representation of women.

The fifth large-scale analysis attempted to reconcile contradictory findings by Marsh et al. vs. Bornmann et al. These teams joined forces and conducted a reanalysis of Bornmann and Daniel's (41) data to resolve their differences. Marsh, Bornmann et al. (45) applied the most powerful analytic approach to date, leading both camps to agree that there was no evidence of sex differences favoring men in any category. In fact, they found evidence favoring women, after controlling for discipline and country. These results were robust, with little study-to-study variation and a lack of interactions: "This noneffect of gender generalized across discipline, the different countries (and funding agencies) considered here, and the publication year" (45, p. 1311). Regarding sex differences in approval of fellowship applications, there was a small but statistically significant difference in favor of men. This finding is the closest any of the analyses have come to replicating Wennerås and Wold (29) (*SI Text, S7*), although the magnitude of this finding was not nearly as pronounced. The joint team interpreted this single finding in favor of men as an aberration from an otherwise unambiguous pattern of no sex advantages or even slight female advantage, and the lack of sex differences generalized over country and discipline (45).

Finally, the sixth large-scale analysis of funding was conducted on >100,000 NIH submissions in six biomedical categories between 1996 and 2007 (46). The percentages of submissions funded were largely equivalent, with men favored slightly in some categories and women favored in others, leading the authors to conclude: "Men and women have near-equal NIH funding success at all stages of their careers, which makes it very unlikely that female attrition is due to negative selection from NIH grant-funding decisions" (p. 1473).

To recap, the weight of evidence overwhelmingly points to a gender-fair grant review process. There are occasional small aberrations, sometimes favoring men and sometimes favoring women; all of the smaller-scale studies failed to replicate Wennerås and Wold's provocative findings, and all but one of the large-scale studies did as well—however, this one study was reversed after a more ambitious joint reanalysis (45). Despite this overwhelming counterevidence, numerous organizations continue to suggest grant review is discriminatory (47), thus diverting attention from legitimate factors limiting women's participation in math-based careers.

The pattern of null sex effects reviewed above is based on funding decisions since the mid-1980s. This period may differ from that >25 y ago. Perhaps sexism was more common at agencies then and women's grants had to be superior to men's to be funded. Such sexism would be unsurprising given other evidence of sexism from this earlier era. Still, sexist reviews cannot be the source of today's dearth of women entering assistant professorships in math-intensive fields. In contrast to claims of anti-female bias among funding agencies quoted in the Introduction, this review indicates a level playing field over the last two decades.

### Discrimination Against Women in Hiring

If the underrepresentation of women in math-intensive fields is not due to biased journal or grant reviews, perhaps it results from biased interviewing and hiring decisions? A study of mock-search-committee recommendations for hiring of psychology professors (15) is often invoked for suspecting it does. In this study, 238 psychologists reviewed fictitious assistant professor candidates and more advanced job seekers. The authors used the same CV, varying applicant sex, and found that both female and male reviewers favored CVs with male assistant professor names, although they did not favor men for the more advanced post. Similarly, women on business teams receive less credit than men for identical work. For stereotypically male tasks, if there is ambiguity about the quality of a woman's contribution to a joint task, it is downplayed (48). Both male and female judges rated a hypothetical worker's performance worse when they thought the worker was female. These results, coupled with findings that nonblind auditions for positions in orchestras discriminated against women (12) and that 18-y-old college males favored resumes with male names

for summer jobs when they were similar, although not identical to, those with female names (14), suggest comparable discrimination may be responsible for the dearth of women entering math-intensive fields.

Although none of this evidence involved discrimination in math-intensive fields, it would be unlikely for sex discrimination to occur in all fields except math-intensive ones. A Government Accounting Office (GAO) report notes that women in math-intensive fields express feelings of isolation, dissatisfaction, and discrimination, "assertions that we also heard during many of our site visits to selected campuses" (ref. 49, p. 4).<sup>‡</sup> This report touches on several factors supported by various analyses as being relevant to women's underrepresentation (50). These factors include women being more likely to prefer working fewer hours and at part-time positions to achieve work-family balance. Although 77% of female and 81% of male graduate students believed a full-time career is "important" or "extremely important" (51), sex differences emerged after additional questioning, with 31% of women (vs. only 9% of men) feeling that working part time for a period is "important" or "extremely important". For having a permanent part-time career, the respective proportions were 19% for women and 9% for men (51). Similarly, in the United Kingdom for 2006–2007, female academics were significantly more likely than males to work part-time, 41.8% vs. 26.8% (25).

Such sex differences reflect preferences and choices, whether freely made or constrained by gendered expectations, and result in more women in teaching-intensive, part-time posts where research resources are scarce. Relatedly, the GAO report mentions studies of pay differentials, demonstrating that nearly all current salary differences can be accounted for by factors other than discrimination, such as women being disproportionately employed at teaching-intensive institutions paying less and providing less time for research. Historically, however, this was not true; women, particularly senior women, lagged behind men in pay and promotion (52, 53) (*SI Text, S8*). Ginther and Kahn (54) analyzed promotion and pay data, noting that historic asymmetries favoring males largely disappeared by the early 2000s, with current asymmetries due to nongender factors. Others have also found that after controlling for structural variables such as status of university, discipline, and presence of young children (which affects women disproportionately), there is no evidence of discriminatory treatment, because women and men in the same circumstances (e.g., same type of institution, discipline, and amount of experience) fare equivalently. Again, although these variables affect men and women similarly, they disadvantage women more in practice, because more women work at teaching-intensive jobs. A National Center for Education Statistics study found that among full-time faculty, women were more likely to work in 2-y institutions (33% vs. 23%), and men in research universities (20% vs. 14%). Whether this is a consequence of choices freely made, or constrained by gendered expectations related to work-family balance coupled with inflexibility in tenure-track timetables and employment options, is worthy of study.

Finally, an in-depth analysis of academic interviewing, hiring, institutional resources, and climate at R1 universities in six areas of natural science by an NRC task force (55) found that, among PhDs applying for tenure-track jobs, women were slightly more likely than men to be invited to interview and offered jobs: "If women applied for positions at R1 institutions, they had a better chance of being interviewed and receiving offers than male job candidates" (ref. 55, p. 5). These results are inconsistent with initiatives promoting gender sensitivity training for search committees and grant panels, which assume bias in funding and hiring of women (ref. 47, also see refs. 11, 56, and 57). Such initiatives target historical rather than current problems facing women scientists.

<sup>‡</sup>Analyses of >8,000 University of California graduate students' responses by Mason and Goulden (50) document the important role played by family formation in female graduate students' decisions to switch out of tenure-track careers in science. For example, married female doctoral students with children are 35% less likely to enter a tenure-track position after receiving a Ph.D. than married men with children and they are 27% less likely than men to achieve tenure. See <http://ucfamilyedge.berkeley.edu/grad%20life%20survey.html>.

## Conclusion: Redirecting Energies Toward Today's Causes of Underrepresentation

Despite frequent assertions that women's current underrepresentation in math-intensive fields is caused by sex discrimination by grant agencies, journal reviewers, and search committees, the evidence shows women fare as well as men in hiring, funding, and publishing (given comparable resources). That women tend to occupy positions offering fewer resources is not due to women being bypassed in interviewing and hiring or being denied grants and journal publications because of their sex. It is due primarily to factors surrounding family formation and childrearing, gendered expectations, lifestyle choices, and career preferences—some originating before or during adolescence (3, 50, 54, 58) (*SI Text, S9*)—and secondarily to sex differences at the extreme right tail of mathematics performance on tests used as gateways to graduate school admission (*SI Text, S10*). As noted, women in math-intensive fields are interviewed and hired slightly in excess of their representation among PhDs applying for tenure-track positions. The primary factors in women's underrepresentation are preferences and choices—both freely made and constrained: “Women choose at a young age not to pursue math-intensive careers, with few adolescent girls expressing desires to be engineers or physicists, preferring instead to be medical doctors, veterinarians, biologists, psychologists, and lawyers. Females make this choice despite earning higher math and science grades than males throughout schooling” (3).

Although women earn a large portion of undergraduate degrees in all science and math fields, disproportionately fewer matriculate in math-intensive graduate fields, preferring biology, medicine, and nonscience fields (law, humanities)—even when math ability is held constant. Of women who matriculate in math-intensive graduate fields, more drop out or change majors. Even among those who complete doctorates in math fields, fewer apply for tenure-track posts than do male counterparts. And the leakage of women continues even after starting careers as assistant professors—especially in math and physical sciences, and this trend continues as women advance through the ranks:

“Although the reasons for this attrition are not well understood, it appears to have less to do with discrimination or ability than with fertility decisions and lifestyle choices, both freely made and constrained. The tenure structure in academe demands that women having children make their greatest intellectual contributions contemporaneously with their greatest physical and emotional achievements, a feat not expected of men. When women opt out of full-time careers to have and rear children, this is a choice—constrained by biology—that men are not required to make” (3, p. 4).

To the extent that women's choices are freely made and women are satisfied with the outcomes, then we have no problem. However, to the extent that these choices are constrained by biology and/or society, and women are dissatisfied with the outcomes, or women's talent is not actualized, then we most emphatically have a problem. With a redirection of resources, this problem might be addressed by education and outreach to young women and girls and to academic administrators. Past strategies to remediate women's underrepresentation can be viewed as a success story; however, continuing to advocate strategies successful in the past to combat shortages of women in math-based fields today mistakes the current causes of women's underrepresentation.

If not discrimination, what is the cause of women's underrepresentation? Today, the dearth of women in math-based fields is related to three factors, one of which (fertility/lifestyle choices) hinders women in all fields, not just mathematical ones, whereas the others (career preferences and ability differences) impact women in math-based fields. Regarding the role of math-related career preferences, adolescent girls often prefer careers focusing on people as opposed to things, and this preference accounts for their burgeoning numbers in such fields as medicine and biology, and their smaller presence in math-intensive fields such as computer science, physics, engineering, chemistry, and mathematics, even when math ability is equated. In a recent metaanalysis of >500,000 participants, the male-female effect size for preferring people vs. things overall was  $d > 0.90$ , and for engineering, 1.1,

both substantial differences (59). One strategy to broaden girls' interests and aspirations involves providing them with realistic information about career opportunities and exposing them to role models in math-based fields. This intervention is not meant to dissuade girls from aspiring to be physicians, veterinarians, and biologists, fields in which women are becoming a majority, but rather to ensure they do not opt out of inorganic fields because of misinformation or stereotypes.

Regarding the role of math-ability differences, potentially influenced by both socialization and biology, twice as many men as women are found in the top 1% of the math score distribution (e.g., SAT-M, GRE-Q). A 30-y study of 1.6 million talent search participants revealed the male-female ratio of SAT-M scores in the top 0.01% has remained relatively stable since the mid-1990s at roughly 4:1 (60). This upper-tail difference is more pronounced for spatial ability (61) due partly to sex differences in variances in cognitive abilities (4). However, ability differences are a secondary explanation for the dearth of women in math-intensive fields because, even given these differences, we would still expect more women in these fields (e.g., a 4:1 ratio would engender 20% female professors in, say, engineering, and a 2:1 ratio would lead to 33%, whereas actual percentages of women are lower (62; *SI Text, S10*).

The third factor influencing underrepresentation affects women in all fields: fertility choices and work-home balance issues. However, this challenge is exacerbated in math-intensive fields because the number of women is smaller to begin with. Attrition at each stage (from undergraduate to graduate school to tenure track) further reduces an already small number. There are significant sex differences in hours worked and lifestyle preferences (58), and having children early in one's career exerts more downward pressure on pretenure women than men (4, 52, 53). The tenure system has strong disincentives for women to have children; these disincentives are why more women in the academy are childless than men, and even women on tenure track with children are twice as likely as men to say they had fewer children than desired (50). Not only is it more common for male academic scientists to have children than for female scientists, but males with children are more likely to be tenured than females with children. Compared with males, new female PhDs are less likely to apply for tenure-track posts; and among those who do apply, females are more likely to terminate for family reasons (55). The GAO report (49) noted that many women PhD students stated during compliance visits that they would not seek tenure-track positions (*SI Text, S11*). In sum, the most salient reasons for women's underrepresentation today are career preferences and fertility/lifestyle choices, both free and constrained.

The GAO report lists strategies, such as stopping tenure clocks for family formation and tenure-track positions seguing from part-time to full-time. Gender Equity Committees have suggested adjusting the length of time to work on grants to accommodate child-rearing, no-cost grant extensions, supplements to hire postdocs to maintain momentum during family leave, reduction in teaching responsibilities for women with newborns, grants for retooling after leaves of absence, couples-hiring, and childcare to attend professional meetings (47, 50, 63). The UC-Berkeley's “Family Edge” provides high-quality childcare and emergency backup care, summer camps and school break care, and reentry postdocs and instructs committees to ignore family-related gaps in CVs. Research into these strategies is needed to identify which are promising.

Federal agencies and universities could play an important role by funding studies on the differing lifecourses of women's and men's careers to determine whether the traditional timing of hiring, tenure, and promotion may deny society and science the contributions of talented women. Perhaps women in scientific fields generally have greater impact later in their careers when family needs are less intense, even if they were less productive earlier because of family-balancing conflicts, as research has shown in biology (64). If this finding can be generalized to today's cohort of women in math-intensive fields, universities might explore options for offering women part-time tenure-track jobs (with concomitantly longer periods of time in which to amass a tenure portfolio), posts that could segue to full-time

once women were ready. However, implementing such flexible options will require motivation and commitment of resources, and raises important questions that research will need to resolve (e.g., the impact on graduate students and postdocs working with part-time faculty; ways to “game” the part-time option for tenure). The linear career path of the modal male scientist of the past may not be the only route to success, and departments and universities should be encouraged and funded to experiment

with alternate lifecourse options. A partnership between the academy and federal funding agencies could be instrumental in researching such alternatives.

**ACKNOWLEDGMENTS.** We acknowledge critical feedback from David Dunning, Kevin Clermont, Emily Sherwin, Ronald Ehrenberg, Valerie Reyna, Michael Macy, and anonymous reviewers. This work was funded by National Institutes of Health Grant R01 NS069792-01.

- Burrelli J (2008) *Thirty-three years of women in S&E faculty positions*. InfoBrief NSF 08-308. (National Science Foundation, Division of Science Resources Statistics, Arlington, VA).
- Ceci SJ, Williams WM (2010) *The Mathematics of Sex: How Biology and Society Conspire to Limit Talented Women and Girls* (Oxford Univ Press, New York).
- Ceci SJ, Williams WM (2010) Sex differences in math-intensive fields. *Curr Dir Psychol Sci* 19:275–279.
- Ceci SJ, Williams WM, Barnett SM (2009) Women's underrepresentation in science: Sociocultural and biological considerations. *Psychol Bull* 135:218–261.
- Chesler NC, Barabino G, Bhatia SN, Richards-Kortum R (2010) The pipeline still leaks and more than you think: a status report on gender diversity in biomedical engineering. *Ann Biomed Eng* 38:1928–1935.
- Lortie CJ, et al. (2007) Publication bias and merit in ecology. *Oikos* 116:1247–1253.
- Hill C, Corbett C, St. Rose A (2010) *Why So Few? Women in Science, Technology, Engineering, and Mathematics* (Amer Assoc Univ Women, Washington, DC).
- NAS Committee on Science, Engineering, and Public Policy (2006) *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering* (Natl Acad Press, Washington, DC).
- Faculty Newsletter MIT (1999) A Study on the Status of Women Faculty in Science at MIT. <http://web.mit.edu/fnl/women/women.html>. Accessed June 10, 2008.
- McLaughlin T (December 27, 2006) Top Wall Street jobs still elude women, minorities. *Boston Globe*. Available at [http://www.boston.com/news/education/higher/articles/2006/12/27/top\\_wall\\_street\\_jobs\\_still\\_elude\\_women\\_minorities](http://www.boston.com/news/education/higher/articles/2006/12/27/top_wall_street_jobs_still_elude_women_minorities). Retrieved September 19, 2008.
- Rhode D (1997) *Speaking of Sex: The Denial of Gender Inequality* (Harvard Univ Press, Cambridge, MA).
- Goldin C, Rouse C (2000) Orchestrating impartiality: The impact of “blind” auditions on female musicians. *Am Econ Rev* 90:715–741.
- Coffey B, McLaughlin P (2009) Do masculine names help female lawyers become judges? *Am Law Econ Rev* 11:112–133.
- Foschi M, Lai L, Sigerson K (1994) Gender and double standards in the assessment of job applicant. *Soc Psychol Q* 57:326–339.
- Steinpreis R, Anders K, Ritzke D (1999) The impact of gender on the review of the CVs of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41: 509–528.
- Angel M, Whang E, Banker R, Lopez J (2010) Statistical Evidence on the Gender Gap in Law Firm Partner Compensation. Temple University Legal Studies Research Paper No. 2010-24, <http://ssrn.com/abstract=1674630>.
- Borsuk R, et al. (2009) To name or not to name: The effect of changing author gender on peer review. *BioSci* 59:985–989.
- Budden AE, et al. (2008) Double-blind review favours increased representation of female authors. *Trends Ecol Evol* 23:4–6.
- Budden A, et al. (2008) Response to Webb et al. *Trends Ecol Evol* 23:353–355.
- Webb TJ, O'Hara B, Freckleton RP (2008) Does double-blind review benefit female authors? *Trends Ecol Evol* 23:351–353, author reply 353–354.
- Hammerschmidt K, Reinhardt K, Rolff J (2008) Does double-blind review favor female authors? *Frontiers in Ecol and the Environ* 6:354.
- Whittaker RJ (2008) Journal review and gender equality: A critical comment on Budden et al. *Trends Ecol Evol* 23:478–479, author reply 480.
- Tregenza T (2002) Gender bias in the refereeing process? *Trends Ecol Evol* 17: 349–350.
- Nature Neuroscience (2006) Women in neuroscience. *Nature Neurosci* 9:853.
- Brooks J, Della Sala S (2009) Re-addressing gender bias in Cortex publications. *Cortex* 45:1126–1137.
- Xie Y, Shauman K (1998) Sex differences in research productivity: New evidence about an old puzzle. *Amer Soc Rev* 63:847–870.
- Allison P, Long S (1990) Departmental effects on scientific productivity. *Am Sociol Rev* 55:119–125.
- Committee on Gender Differences in the Careers of Science, Engineering, and Mathematics Faculty (2009) *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. (Natl Acad Press, Washington, DC).
- Wenneras C, Wold A (1997) Nepotism and sexism in peer-review. *Nature* 387:341–343.
- Abbott A (1997) Equality not taken for granted. *Nature* 390:204.
- Grant J, Burden S, Breen G (1997) No evidence of sexism in peer review. *Nature* 390: 438.
- Dickson D (1997) Female scientists wanted — apply to UK research councils. *Nature* 390:431.
- Friesen HG (1998) Equal opportunities in Canada. *Nature* 391:326.
- Sandstrom U, Hallsten M (2008) Persistent nepotism in peer-review. *Scientometr* 74: 175–189.
- Demicheli V, Di Pietrantonj C (2007) Peer review for improving the quality of grant applications. *Cochrane Database Syst Rev* 2:MR000003.
- RAND (2005) Is there gender bias in federal grant programs? (RAND Infrastructure, Safety, and Environment Brief RB-9147-NSF). Available at [http://rand.org/pubs/research\\_briefs/RB9147/RAND\\_RB9147.pdf](http://rand.org/pubs/research_briefs/RB9147/RAND_RB9147.pdf). Accessed March 1, 2008.
- Leboy P (2008) Fixing the leaky pipeline. *The Scientist* 22:67–70.
- Jayasinghe U, Marsh H, Bond N (2003) Multilevel cross-classified modeling approach to peer-review of grant proposals. *JR Stat Soc* 166A:279–300.
- Marsh HW, Jayasinghe UW, Bond NW (2008a) Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *Am Psychol* 63: 160–168.
- Marsh HW, Jayasinghe U, Bond N (2008b) *Peer Reviews of Grant Applications* (SELF Res Centre, Univ Oxford, Oxford).
- Bornmann L, Daniel H-D (2005) Selection of research fellowship recipients by committee peer review. *Scientometr* 63:297–320.
- Bornmann L (2007) Bias cut: Women, it seems, often get a raw deal in science—So how can discrimination be tackled? *Nature* 445:566.
- Bornmann L, Mutz R, Daniel HD (2007) Gender differences in grant peer review: A meta-analysis. *J of Informetr* 1:226–238.
- Broder I (1993) Review of NSF economics proposals: Gender and institutional patterns. *Am Econ Rev* 83:964–970.
- Marsh HW, Bornmann L, Mutz R, Daniel H-D, O'Mara A (2009) Gender effects in the peer reviews of grant proposals. *Rev Educ Res* 79:1290–1326.
- Ley TJ, Hamilton BH (2008) Sociology. The gender gap in NIH grant applications. *Science* 322:1472–1474.
- Comm on Status of Women in Physics of the American Physics Society (2007) Strengthening the Physics Enterprise in Universities and National Laboratories. *Gazette*, 26:2. Available at <http://www.aps.org/programs/women/reports/gazette/loader.cfm?csModule=security/getfile&PageID=102599>. Accessed January 11, 2011.
- Heilman ME, Haynes MC (2005) No credit where credit is due: Attributional rationalization of women's success in male-female teams. *J Appl Psychol* 90:905–916.
- General Accounting Office (July 2004) Gender issues: Women's participation in the sciences has increased, but agencies need to do more to ensure compliance with title IX. GAO 04-639. Available at [www.gao.gov/cgi-bin/getprpt?GAO-04-639](http://www.gao.gov/cgi-bin/getprpt?GAO-04-639).
- Mason MA, Goulden M (2009) UC Doctoral Student Career and Life Survey (Univ of Calif, Berkeley). Available at <http://ucfamilyedge.berkeley.edu/grad%20life%20survey.html>. Accessed January 3, 2011.
- Lubinski D, Benbow CP, Shea DL, Eftekhari-Sanjani H, Halvorson MB (2001) Men and women at promise for scientific excellence: Similarity not dissimilarity. *Psychol Sci* 12: 309–317.
- Ginther D (2004) Why women earn less than men. *AWIS Mag* 33:1–5.
- Ginther DK (February 2001). Does Science Discriminate against Women? (Fed Res Bank of Atlanta, Atlanta), Working Paper 2001-02.
- Ginther DK, Kahn S (2006) Does Science Promote Women?. *Science and Engineering Careers in the United States: An Analysis of Markets and Employment*, eds Freeman RB, Goroff DL (Natl Bur Econ Res, Chicago), pp 163–194.
- Faculty Committee on Women in Science, Engineering, and Medicine (2010). *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. (Natl Acad Press, Washington DC).
- Sommers CH (2008) Why can't a woman be more like a man? *The American*, Available at [www.american.com/archive/2008/march-april-magazine-contents/why-can2019t-a-woman-be-more-like-a-man](http://www.american.com/archive/2008/march-april-magazine-contents/why-can2019t-a-woman-be-more-like-a-man). Accessed September 19, 2008.
- Tierney J (June 14, 2010) Legislation won't close gender gap in science. *New York Times*, D4. Available at: [http://www.nytimes.com/2010/06/15/science/15tier.html?\\_r=1&scq=stephen%20ceci&st=cse](http://www.nytimes.com/2010/06/15/science/15tier.html?_r=1&scq=stephen%20ceci&st=cse). Accessed 1/13/2011.
- Ferriman K, Lubinski D, Benbow CP (2009) Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: Developmental changes and gender differences during emerging adulthood and parenthood. *J Pers Soc Psychol* 97:517–532.
- Su R, Rounds J, Armstrong PI (2009) Men and things, women and people: A meta-analysis of sex differences in interests. *Psychol Bull* 135:859–884.
- Wai J, Cacchio M, Putallaz M, Makel M (2010) Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence* 38:412–423.
- Wai J, Lubinski D, Benbow CP (2009) Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *J Ed Psych* 101:817–835.
- Valla J, Ceci S Can sex differences in science be tied to the long reach of prenatal hormones? *Perspect Psychol Sci*, in press.
- Sher M (2006) Dual career couples: Problem of opportunity? *CSWP Gazette* 25:1–2, 9.
- Long J (1992) Measures of sex differences in scientific productivity. *Soc Forces* 71: 159–178.