IDENTIFYING AND MINIMIZING UNINTENTIONAL ("unconscious") BIAS IN SCIENTIFIC PEER REVIEW

Health Research Alliance

Anna Kaatz, PhD, MPH, MA

Director of Computational Sciences

Tues. March 13, 2018

CENTER FOR WOMEN'S HEALTH RESEARCH
SCHOOL OF MEDICINE AND PUBLIC HEALTH
UNIVERSITY OF WISCONSIN-MADISON

School of Medicine and Public Health
UNIVERSITY OF WISCONSIN-MADISON

Acknowledgements.

National Institutes of Health
Turning Discovery Into Health

Grant #s: R35 GM122557
R01 GM111002
R25 GM083252

National Science Foundation

Grant #s: 1747445
1760092

VCRGE Bridge Award
Hilldale Fellowship
Dept. of Medicine
UW Foundation
ICTR-CTSA
CWHR
WISE

We do Data Science (STEMM) for Social Justice, Health Equity, and Environmental Sustainability

Project Information

DESCRIPTION

Project Number: 1R01GM111002-01 **Former Number:** 1R01OD016764-01

Contact PI / Project Leader: CARNES, MOLLY L. (Contact)
DEVINE, PATRICIA G
FORD, CECILIA E

Title: EXPLORING THE SCIENCE OF SCIENTIFIC REVIEW

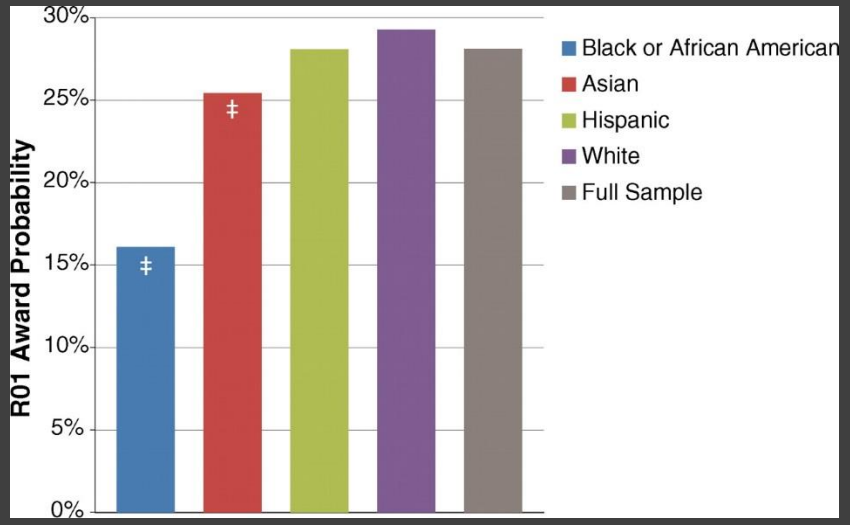
Awardee Organization: UNIVERSITY OF WISCONSIN-MADISON

Abstract Text:

DESCRIPTION (provided by applicant): All humans-no matter how intelligent, egalitarian, or well-intentioned-are susceptible to cognitive biases in the way they make decisions and judge others. Although these biases can operate unintentionally in opposition to one's conscious intentions, personal beliefs, and objective data, they may unwittingly perpetuate social inequalities. Unexplained disparities in R01 funding outcomes by race and gender have raised concern about bias in NIH peer review. This Transformative R01 will examine if and how implicit (i.e., unintentional) bias might occur in R01 peer review through the following three Specific Aims: Specific Aim #1. Identify the extent to which investigator characteristics influence the words and descriptors chosen by R01 peer-reviewers and how text relates to assigned scores. We will validate positive and negative grant evaluation word categories, analyze the text of a national sample of R01 reviews, and compare the grant review text for different investigator characteristics. We hypothesize that categories of words and descriptors will differ in ways that suggest implicitly different evaluation standards by applicant race and gender, even when application scores and funding outcomes are similar. Specific Aim #2. Determine whether investigator race, gender, or institution causally influences the review of identical proposals. We will conduct a randomized, controlled study in which we manipulate characteristics of a grant principal investigator (PI) to assess their influence on grant review outcomes. We will request donations of actual funded R01s and, within each grant, manipulate the PI's gender, race, or home institution. We will then invite reviewers in the appropriate discipline to review the proposals, and we will analyze written reviews and scores. We hypothesize that investigator variables will significantly influence scores and review text such that grants attributed to higher status groups (male, White, prestigious institution) will obtain better scores and text will suppor implicitly different standards of excellence. Specific Aim #3. Examine how interactional patterns among study section members promote receptivity and resistance to discussion topics and associated grant applicants. In audio- and videotapes of constructed study sections, we will investigate the real-time social interactional processes in the discussions of R01 proposals. We will employ conversation analysis to examine the delivery of initial rankings and their rationales, topic development, and the processes through which final rankings are negotiated. This research is innovative because it examines for the first time the complexities of potential bias in NIH peer review. The potential impact is threefold; this research will 1) discover whether certain forms of cognitive bias are or are not consequential in R01 peer review, 2) determine whether quantitative text analysis is a useful measure of implicit bias, and 3) describe and label real-tim grant reviewer interactional patterns. Taken together, the results of our research could set the stage for transformation in peer review throughout NIH.

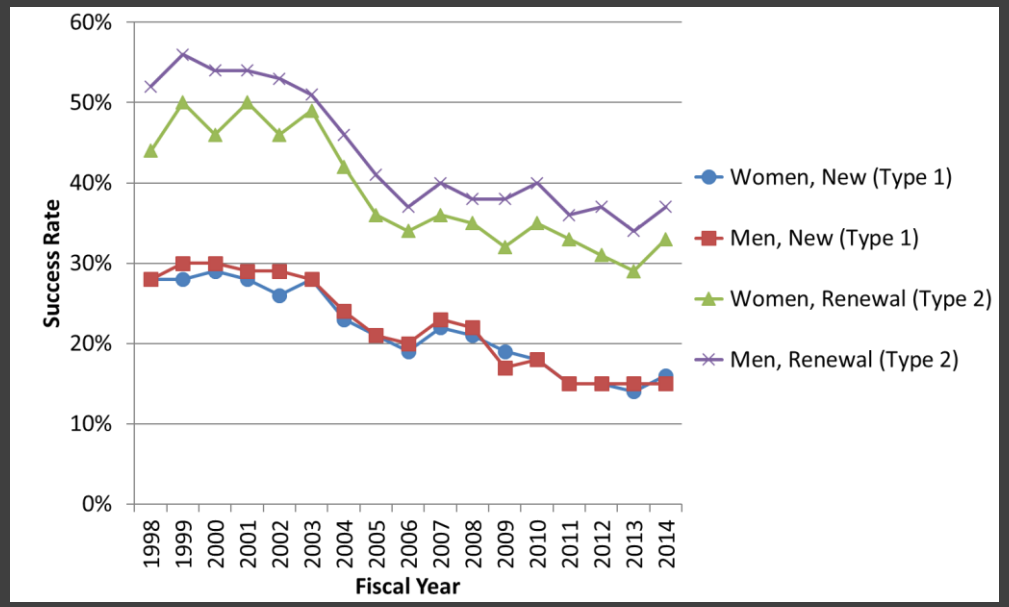
Race and Gender Disparities in R01 Award Rates.

Black PIs Have Lower R01 Award Probabilities Than White PIs



Probability of NIH R01 award by race and ethnicity, FY 2000 to FY 2006 (N = 83,188). Based on data from NIH IMPAC II, DRF, and AAMC Faculty Roster. ‡, P < .001; **, P < .01; *, P < .05; Ginther et al., Science, 2011.

Female PIs Have Lower R01 Renewal (Type 2) Award Rates Than Male PIs



Success Rates for Male and Female Investigators of NIH Type 1 or Type 2 R01 or Equivalent Awards: 1998-2014; NIH, 2015.

Aim 1

Specific Aim #1

- Identify the extent to which investigator characteristics influence the words and descriptors chosen by R01 peer-reviewers and how text relates to assigned scores

- Analysis of R01 Grant critiques and scores from 2010-2014

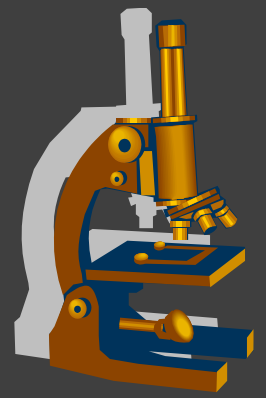
Impact	Score	Descriptor	Strengths/Weaknesses
High Impact	1	Exceptional	Strengths
	2	Outstanding	
	3	Excellent	
Moderate Impact	4	Very Good	
	5	Good	
	6	Satisfactory	
Low Impact	7	Fair	
	8	Marginal	
	9	Poor	Weaknesses

Aim 2

Specific Aim #2

- The first randomized controlled experimental study of the influence of applicant characteristics and application quality on R01 peer review outcomes.

?



Aim 3

Specific Aim #3

- The first study of NIH peer reviewers' interactional patterns during R01 study section meetings.

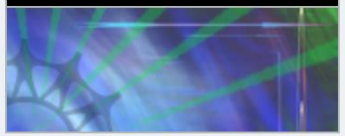
National Science Foundation [US] | https://www.nsf.gov/awardsearch/showAward?AWD_ID=1747445&HistoricalAwards=false

National Science Foundation
WHERE DISCOVERIES BEGIN

SEARCH

HOME | RESEARCH AREAS | FUNDING | AWARDS | DOCUMENT LIBRARY | NEWS | ABOUT NSF

Awards



- Search Awards
- Recent Awards
- Presidential and Honorary Awards
- About Awards

How to Manage Your Award

- Grant Policy Manual
- Grant General Conditions
- Cooperative Agreement Conditions
- Special Conditions
- Federal Demonstration Partnership
- Policy Office Website

Award Abstract #1747445

EAGER: Can Data Mining and Crowd Sourcing Revolutionize the Study of Scientific Peer Review? Generating a National, Open Depository of Grant Review Outcomes from Federal Agencies

NSF Org:	SMA SBE Off Of Multidisciplinary Activities
Initial Amendment Date:	July 28, 2017
Latest Amendment Date:	July 28, 2017
Award Number:	1747445
Award Instrument:	Standard Grant
Program Manager:	Mark K. Fiegener SMA SBE Off Of Multidisciplinary Activities SBE Direct For Social, Behav & Economic Scie
Start Date:	August 1, 2017
End Date:	July 31, 2018 (Estimated)
Awarded Amount to Date:	$149,844.00
Investigator(s):	Anna Kaatz akaatz@wisc.edu (Principal Investigator) You-Geon Lee (Co-Principal Investigator)

Harnessing machine learning algorithms to study scientific grant peer review.

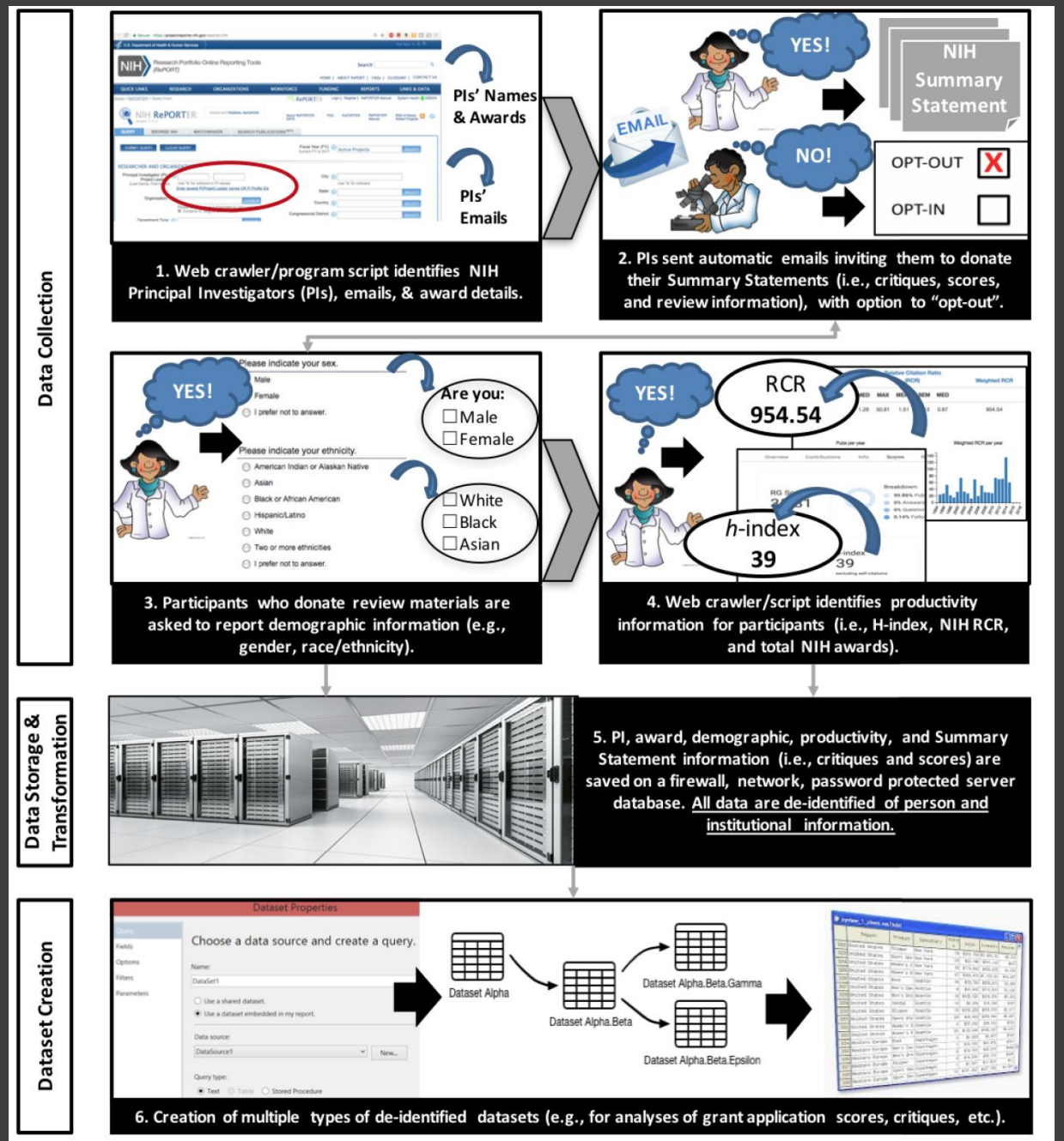


Figure 1. Automated pipeline for collection of review outcomes from applications submitted to the National Institutes of Health (NIH).

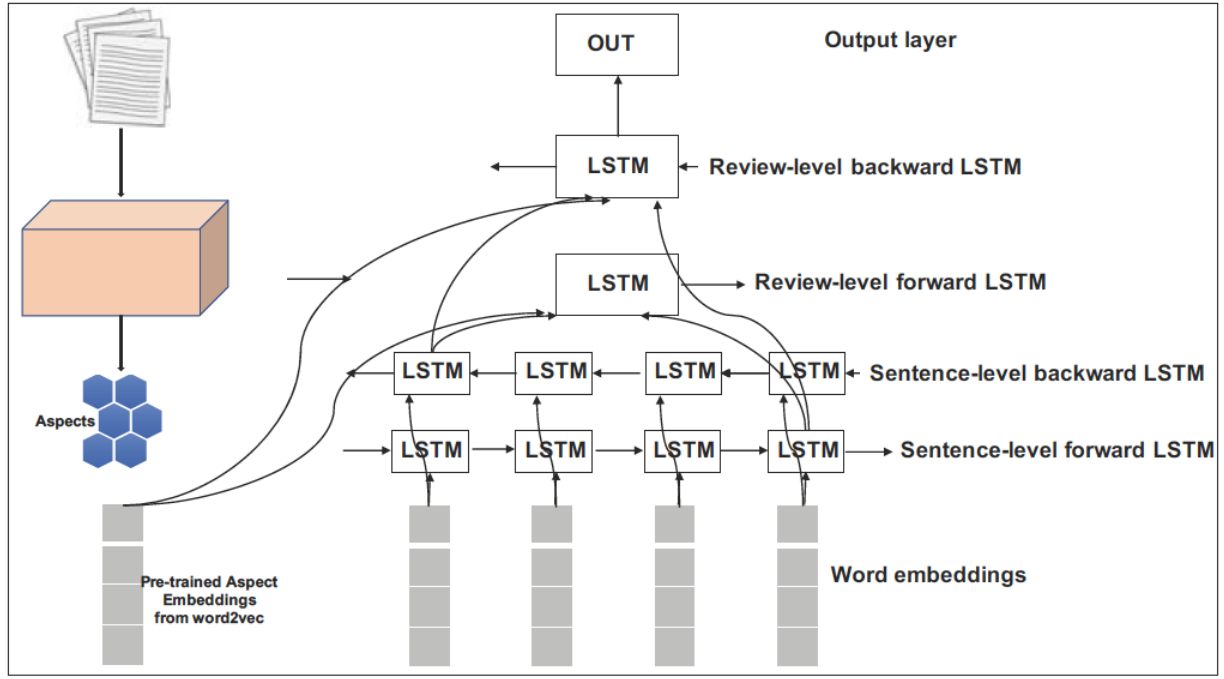


Figure 6. Hierarchical bidirectional LSTM for aspect-based sentiment analysis

Harnessing machine learning algorithms to study scientific grant peer review.

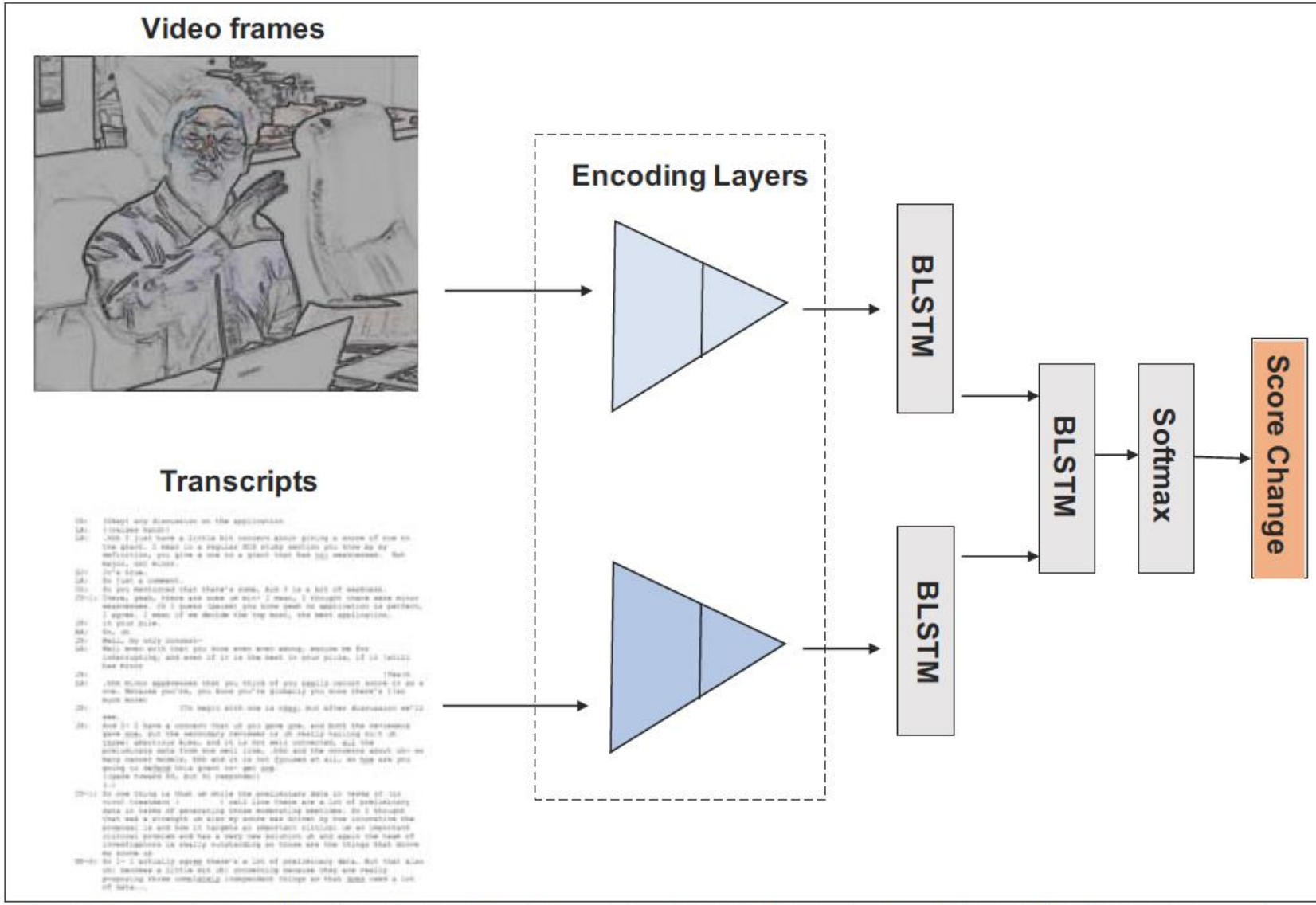
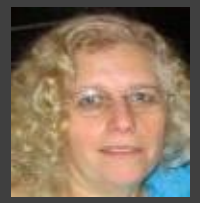


Figure 7. Audiovisual Fusion System to study score change dynamics based on video frame data and transcripts from the study section (Aim 2)

Research Team Members

Molly Carnes, MD, MS
Professor
Director, CWHR
Co-Director, WISELI

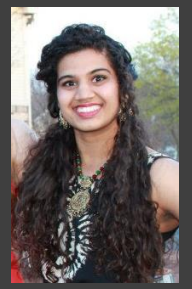


Eve Fine, PhD
Associate Researcher
Women in Science & Engineering Leadership Institute (WISELI)

Jennifer Sheridan, PhD
Associate Scientist
Executive/Research Director, WISELI

Dasta Malikireddy, BS
Graduate Student, Dept. of Computer Science; Research Assistant, Center for Women's Health Research

Skylyn Worzalla
Student Intern, Center for Women's Health Research



Anupama Battcharya,
Student Intern, Center for Women's Health Research

Angela Byars-Winston, PhD
Associate Professor,
Dept. of Medicine
Associate Scientist, CWHR

Wairimu Magua, PhD
Postdoc, Industrial and Systems Engineering

Josh Raclaw, PhD, MA,
Assistant Professor, Westchester University, Honorary Fellow, Center for Women's Health Research

Libby Pier, Phd, MS
Postdoc, Center for Women's Health Research

Cecilia Ford, PhD
Professor
Depts of English and Sociology

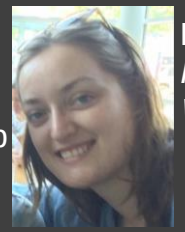
Christine Kolehmainen, MD, MS
Women's Health Physician
Wm. S. Middleton VA Hospital

You-Geon Lee, PhD
Statistican, Center for Women's Health Research, Center for Education Research

Renee Leatherberry,
Financial Officer, Center for Women's Health Research

Jerry Zhu, PhD, MS *Dept of Computer Science.*

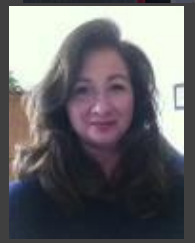
Percey L. Brown, FairPlay workshop facilitator, PhD student Educational Leadership and Policy Analysis



Madeline Jens, BA
Programmer, Center for Women's Health Research

Vicki Leatherberry, BS, *Administrator, Center for Women's Health Research*

Patricia Devine, PhD
Professor
Chair, Dept. of Psychology



Christine Pribbenow, PhD
Associate Scientist
WI Center for Education Research

Sohad Murrar, PhD Candidate, Dept of Psychology, *Center for Women's Health Research*

C. David Paige, PhD,
Professor, Dept. of Computer Science, Dept. of Biostatistics

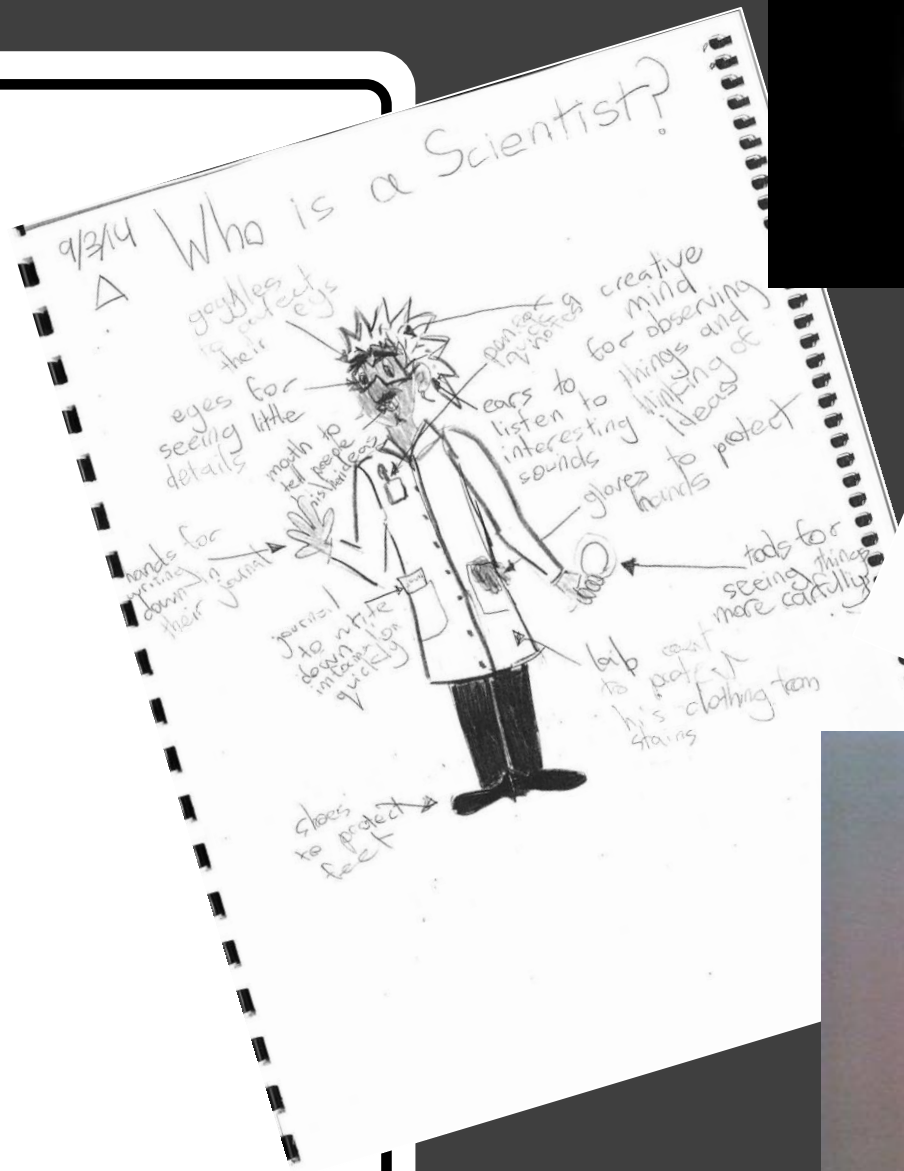
Donald Dantzler, FairPlay workshop facilitator, PhD student Educational Leadership and Policy Analysis

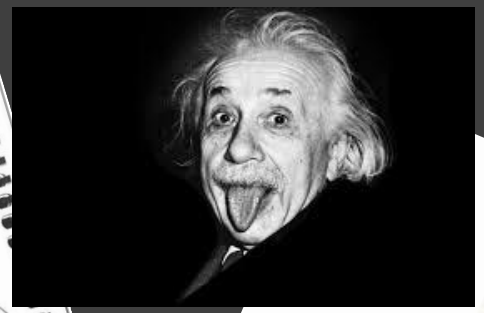
Markus Brauer, PhD Professor, Dept of Psychology

Sharon Topp
Administrator, Center for Women's Health Research

Draw a scientist

Draw a scientist





Why do you think it is important to have a diverse workforce in science, technology, engineering, mathematics, and medicine (STEMM) fields?
to have better representation of women and minorities in STEMM?

Why diversity?

Diverse working groups are more productive, creative, and innovative than homogeneous groups (Herring 2009; Page 2007; van Knippenberg & Michaela 2007; Chang et al., 2003).

Diverse groups engage in a higher level of critical analysis than do homogeneous groups (Sommers 2006; Antonio 2004; Nemeth 1986, 1995).

Diverse scholars and professionals can invigorate and expand disciplines and fields (Schiebinger et al. 2013; Catalyst, 2013).

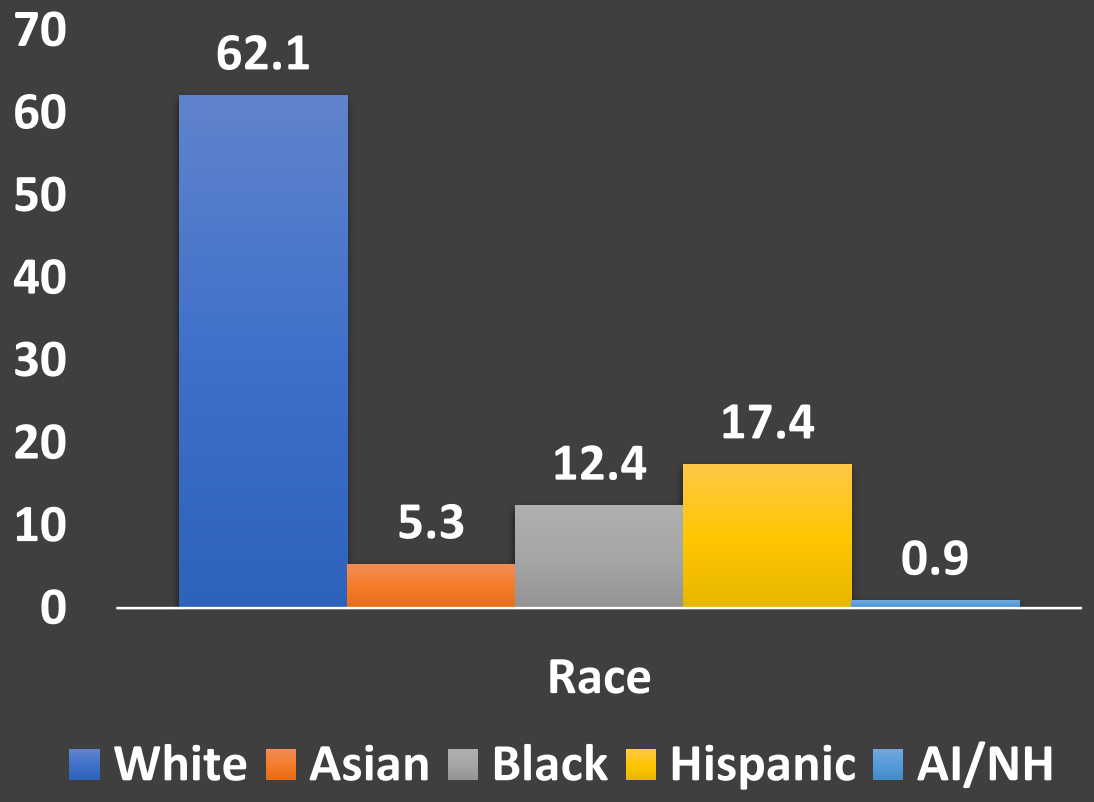
Mentors and role models for all (Nat. Acad. Sci. 2007).

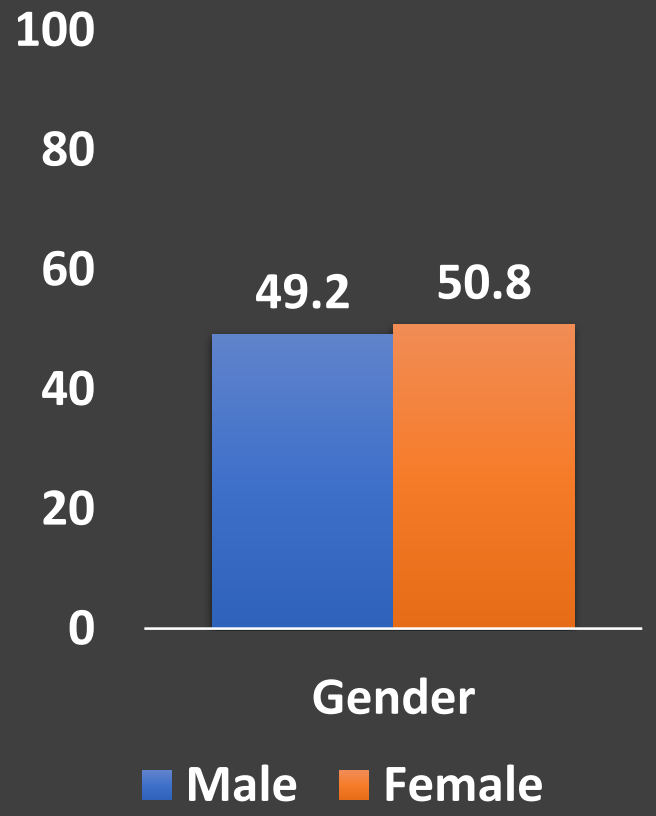
Fairness and equity (Nat. Acad. Sci. 2007).

Diversity leads to innovation.

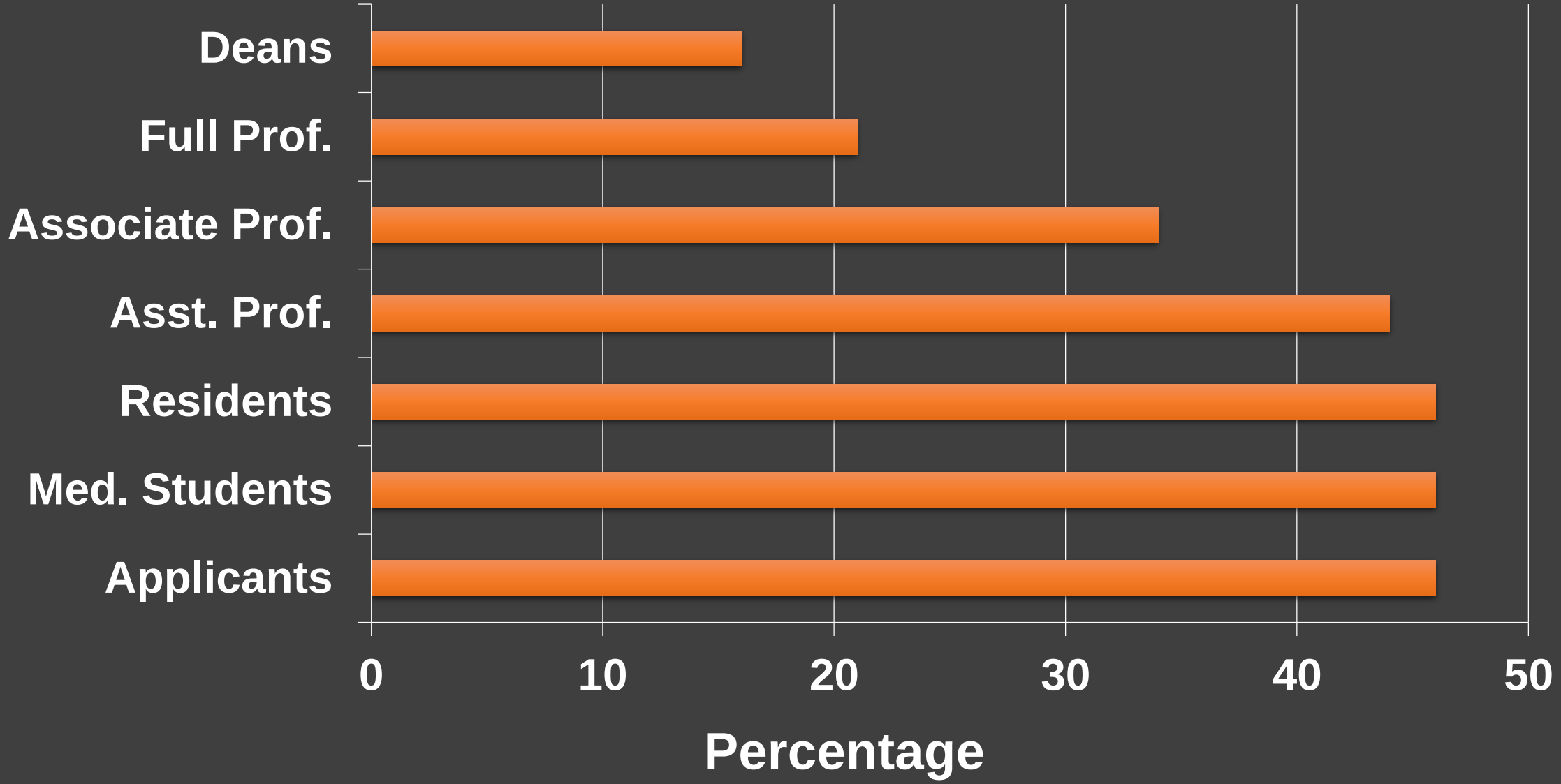
- **Advancing women and racial/ethnic minorities in STEMM is critical for ensuring the future competitiveness of U.S. science and technology.**
    - Research conducted by racial/ethnic minorities and women is linked to technological innovation and is known to address costly education, economic, and health disparities. (National Acad. of Sciences, 2007; 2010; Carnes et al., 2008; Guevara et al., 2013; Roberts et al., 2014)

U.S. Demographics.

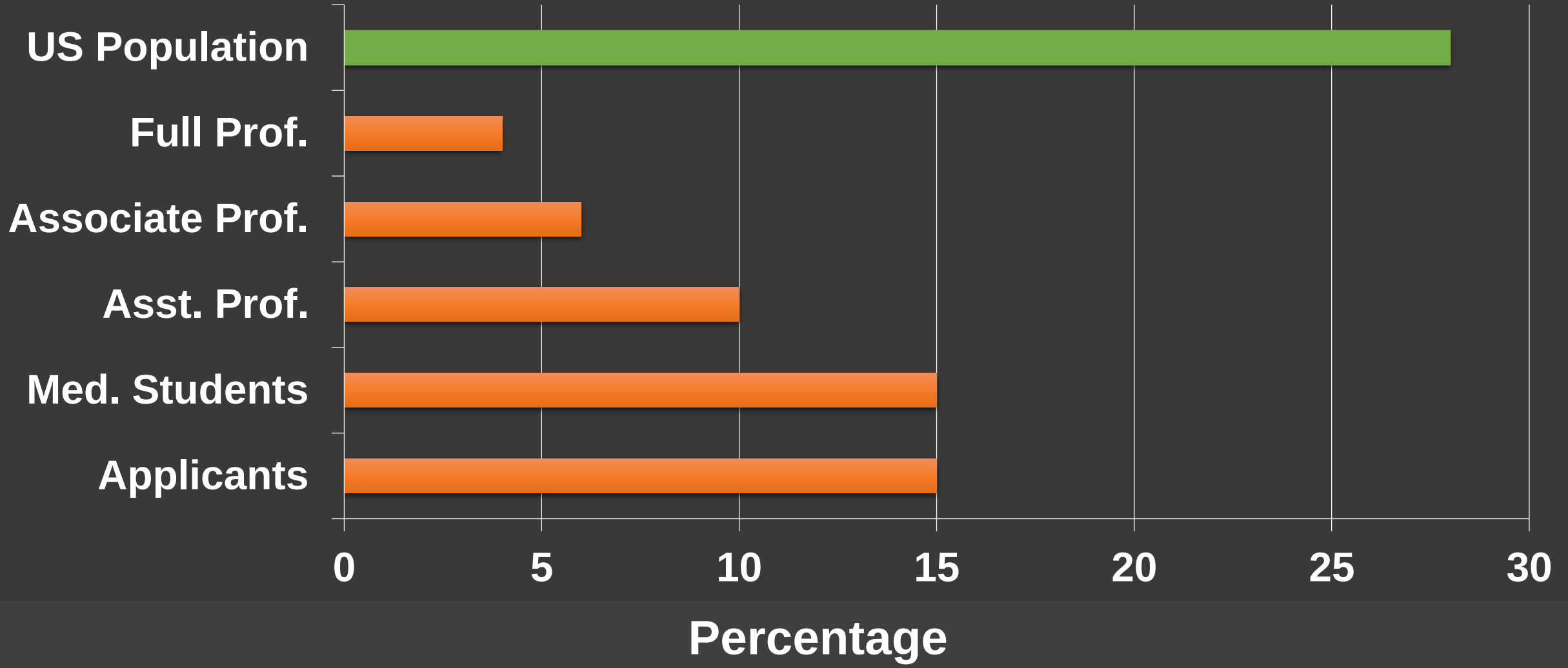




Women's representation in U.S. med. schools by rank (AAMC Benchmarking Report, 2014).

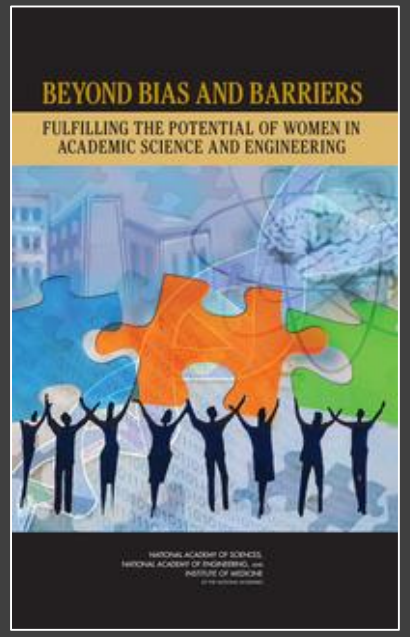


Racial/ethnic minorities (Blk, Hisp, NA) representation in U.S. med. schools by rank (AAMC, 2014, Guevara et al., 2013; Roberts et al., 2014).

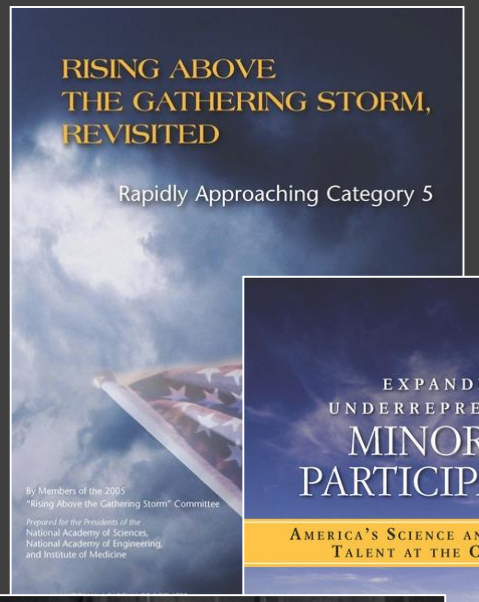


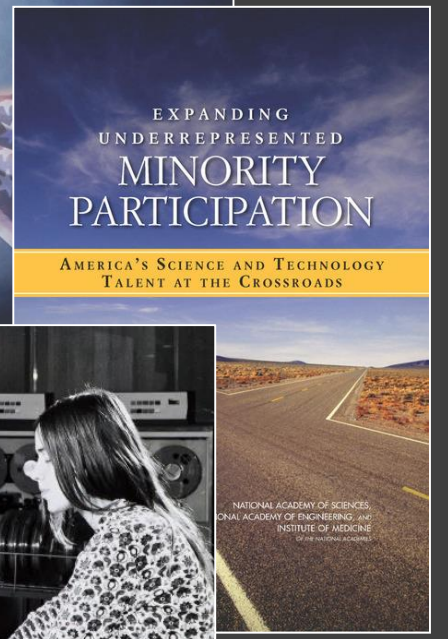
Why haven't we solved this already?

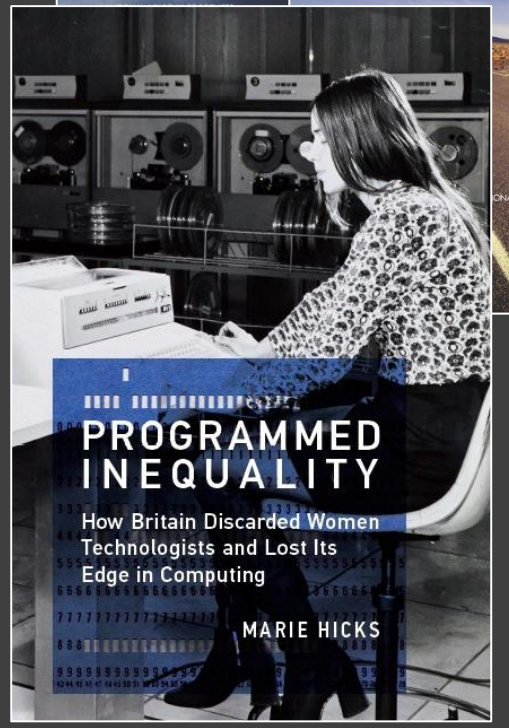
- Title VI (1964) and IX (1972).
- Multiple calls for gender and racial equity for > 40 years.
- National Academy of Sciences concluded that major barriers were:



- ❑ Women and racial/ethnic minorities are equally interested, capable and committed to careers in science and medicine.
- ❑ Stereotypes about gender and race operate to influence self-perception, personal interactions, evaluative processes, and departmental cultures to subtly yet systematically impede women's & racial/ethnic minorities' career advancement.



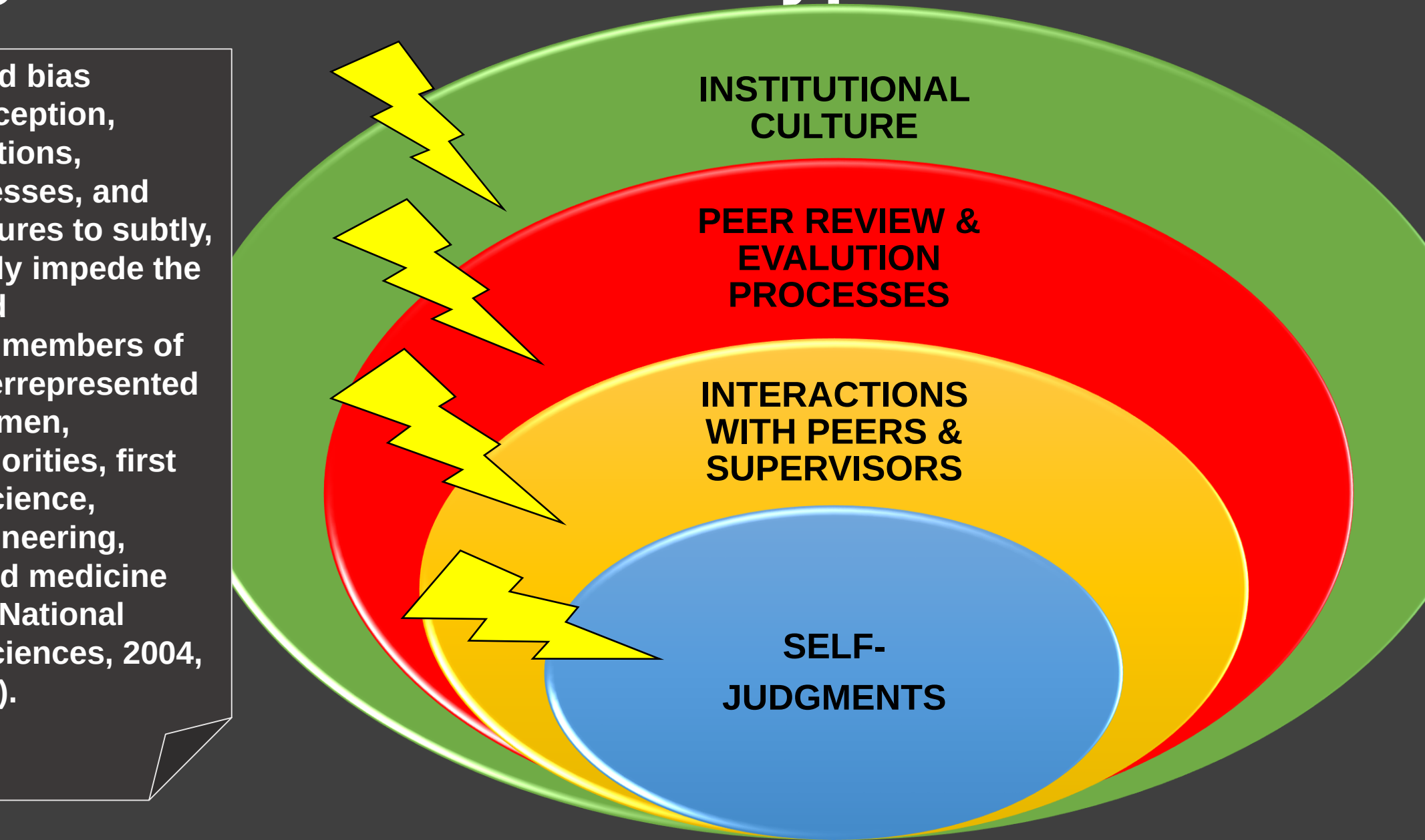




http://mariehicks.net/, 2017

Ecological model of stereotype-based bias.

Stereotype-based bias impacts self perception, personal interactions, evaluation processes, and institutional cultures to subtly, yet systematically impede the participation and advancement of members of historically underrepresented groups (e.g., women, racial/ethnic minorities, first generation) in science, technology, engineering, mathematics, and medicine (STEMM) fields (National Academies of Sciences, 2004, 2007, 2010, 2014).



How to address stereotype-based bias? Multilevel Interventions.

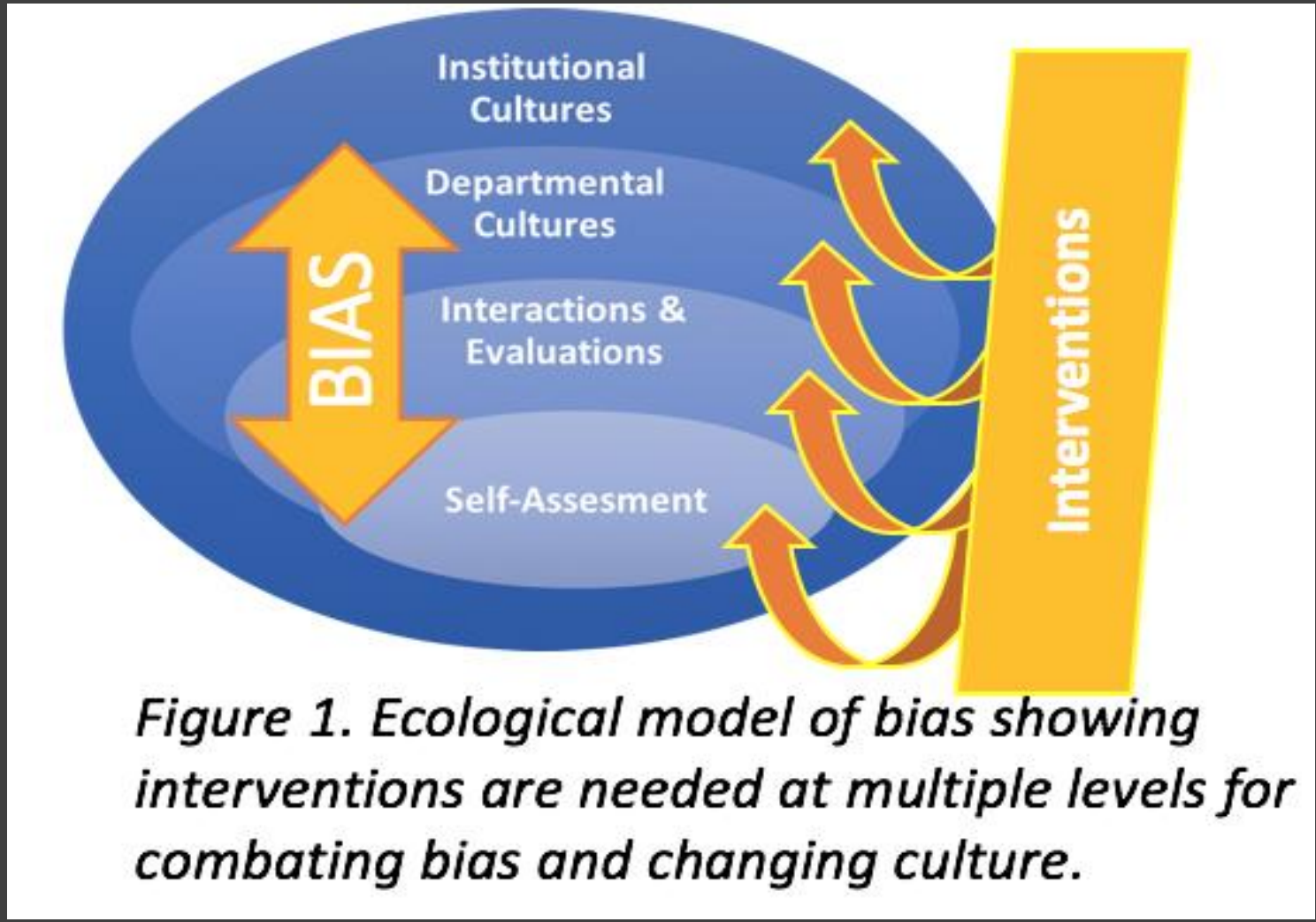


Figure 1. Ecological model of bias showing interventions are needed at multiple levels for combating bias and changing culture.

What is a stereotype?

ster·e·o·type

/ˈsterēəˌtīp/

noun
plural noun: **stereotypes**

1. a widely held but fixed and oversimplified image or idea of a particular type of person or thing.
   "the stereotype of the woman as the carer"
   synonyms: standard/conventional image, received idea, cliché, hackneyed idea, formula
   "the stereotype of the rancher"
   - a person or thing that conforms to a stereotypical image.
     "don't treat anyone as a stereotype"

- Brain observes patterns, over and over…
- Takes mental short-cuts to reserve brainpower for higher order thinking.

Stereotypes can lead to unconscious bias.

- What is unconscious bias?
  - Derives from cultural stereotypes that inform us about what someone is like, what their skills and abilities are, and how they should behave based on their group membership (Nosek et al., 2011).
    - Gender.
    - Race/Ethnicity.
    - Sexual Orientation.
    - Weight.
    - Age.

Group exercise.

What traits or behaviors do you think most people associate with being male or “masculine”?

What traits or behaviors do you think most people associate with being female or “feminine”?

Group exercise.

- What traits do you associate with being a scientist?

Bias arises from stereotypes: Gender.

Women[1]

Aware of the feelings of others
Cheerful
Courteous
Desires to avoid controversy
Desire for friendship
Generous
Grateful
Helpful
Humanitarian
Kind
Need for social acceptance
Need for security
Sentimental
Sociable
Sympathetic
Tactful
Talkative

Men[1]

Able to separate feelings from ideas
Adventurous
Ambitious
Analytical ability
Authoritative
Comfort with aggression
Competitive
Consistent
Curious
Decisive
Desires responsibility
Direct
Dominant
Feelings not easily hurt
Firm
Frank
High self-regard
Independent
Industrious
Intelligent
Knowledgeable about science
Knows the ways of the world
Leadership ability
Logical
Mathematical
Need for achievement
Need for autonomy
Need for recognition
Objective
Persistent
Prompt
Risk-taking
Self-confident
Self-reliant
Steady
Technically skilled
Vigorous
Well-informed

1. Carli et al., 2016.

Role-congruity for White/male & science.

Scientist: Intelligent, Logical, Rationale, Analytical, Independent, Focused, Persistent, Objective, Knowledgeable, Curious, Industrious.

Congruent

Female

Cheerful
Courteous
Generous
Helpful
Kind
Sentimental

Sympathetic
Tactful
Talkative

Male

Ambitious
Intelligent
Logical
Analytical
Decisive

Leadership ability
Objective
Persistent
Technically Skilled

1. Carli et al., 2016.

Role-congruity for White/male & LEADERSHIP.

Leader: Intelligent, Logical, Rationale, Analytical, Independent, Focused, Persistent, Objective, Knowledgeable, Curious, Industrious.

Female

Cheerful
Courteous
Generous
Helpful
Kind
Sentimental
Sympathetic
Tactful
Talkative

Congruent

Male

Ambitious
Intelligent
Logical
Analytical
Decisive
Leadership ability
Objective
Persistent
Technically Skilled

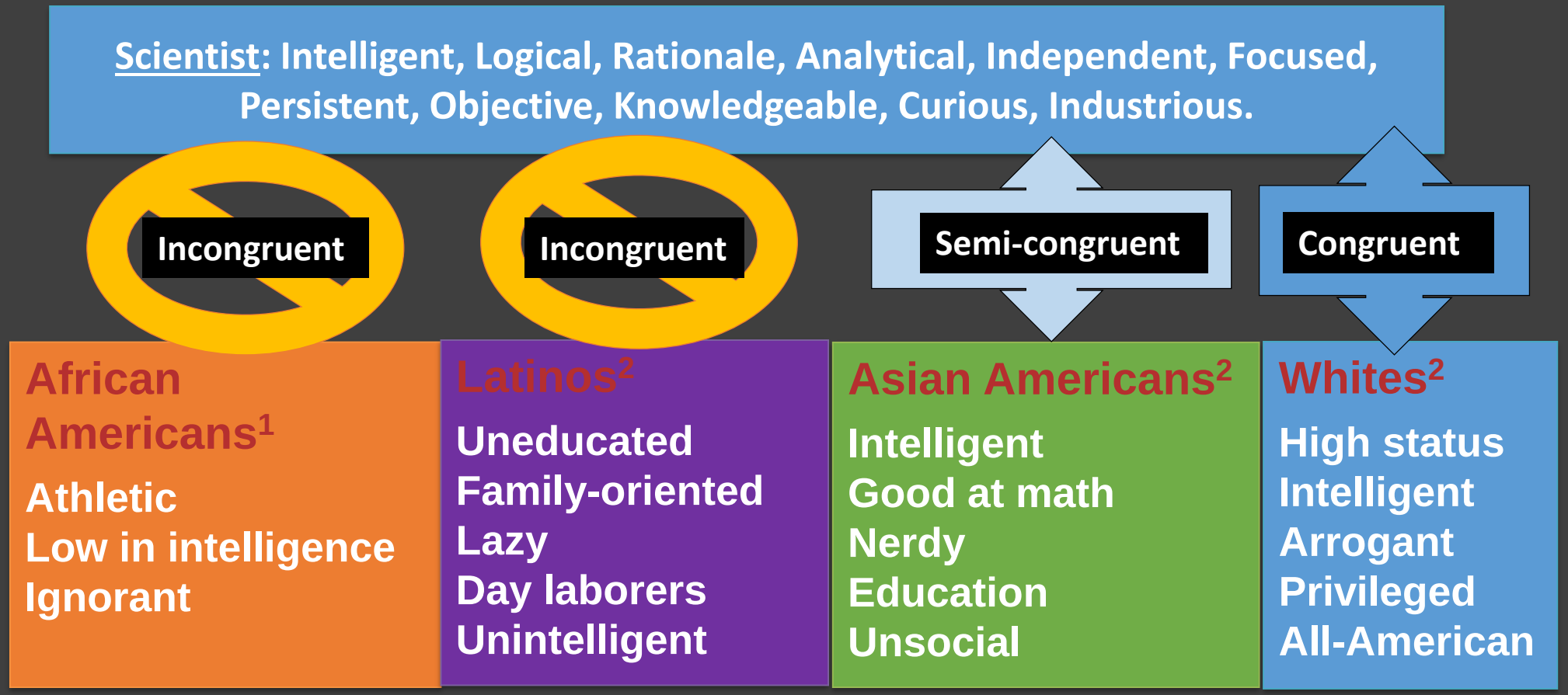
1. Carli et al., 2016.

Bias arises from stereotypes: Race/ethnicity.

African Americans[1]	Asian Americans[2]	Latinos[2]	Whites[2]
Athletic	Intelligent	Poor	High status
Rhythmic	Bad drivers	Have many children	Rich
Low in intelligence	Good at math	Illegal immigrants	Intelligent
Lazy	Nerdy	Dark-skinned	Arrogant
Poor	Shy	Uneducated	Privileged
Loud	Skinny	Family-oriented	Blond
Criminal	Small eyes	Lazy	Racist
Hostile	Education	Day laborers	All-American
Ignorant	Quiet	Unintelligent	Ignorant
		Loud	
		Gangsters	

1. Devine and Elliot, 1995.
2. Ghavami and Peplau., 2015.

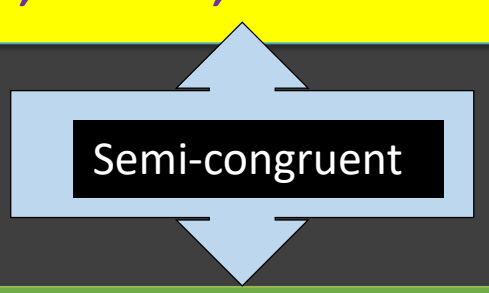
Role-congruity for White/male & science.

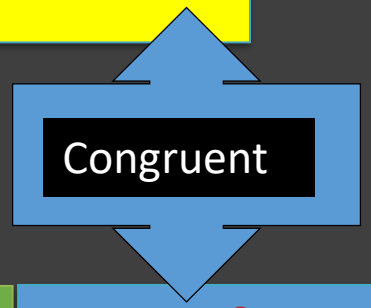


1. Devine and Elliot, 1995.
2. Ghavami and Peplau., 2015.
3. Carli et al., 2016.

Role-congruity for White/male & LEADERSHIP.

LEADER: Intelligent, Logical, Rationale, Analytical, Independent, Focused, Persistent, Objective, Knowledgeable, Curious, Industrious.

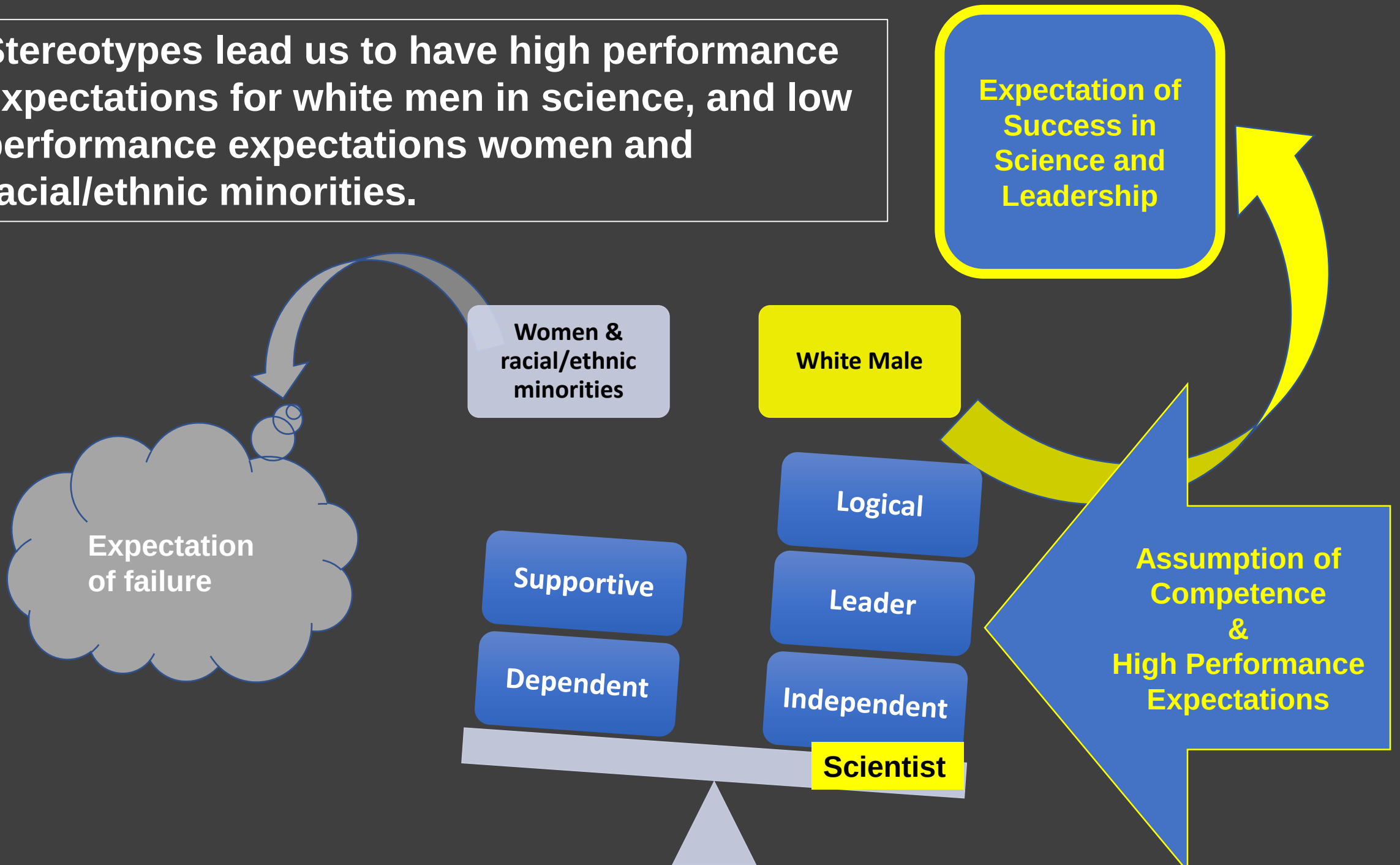




African Americans[1]	Latinos[2]	Asian Americans[2]	Whites[2]
Athletic	Uneducated	Intelligent	High status
Low in intelligence	Family-oriented	Good at math	Intelligent
Ignorant	Lazy	Nerdy	Arrogant
	Day laborers	Education	Privileged
	Unintelligent	Unsocial	All-American

1. Devine and Elliot, 1995.
2. Ghavami and Peplau., 2015.
3. Carli et al., 2016.

- **Stereotypes lead us to have high performance expectations for white men in science, and low performance expectations women and racial/ethnic minorities.**



Stereotype-based bias disadvantages women and racial/ethnic minorities in peer review.

Semantic “linguistic” priming.

Competency bias.

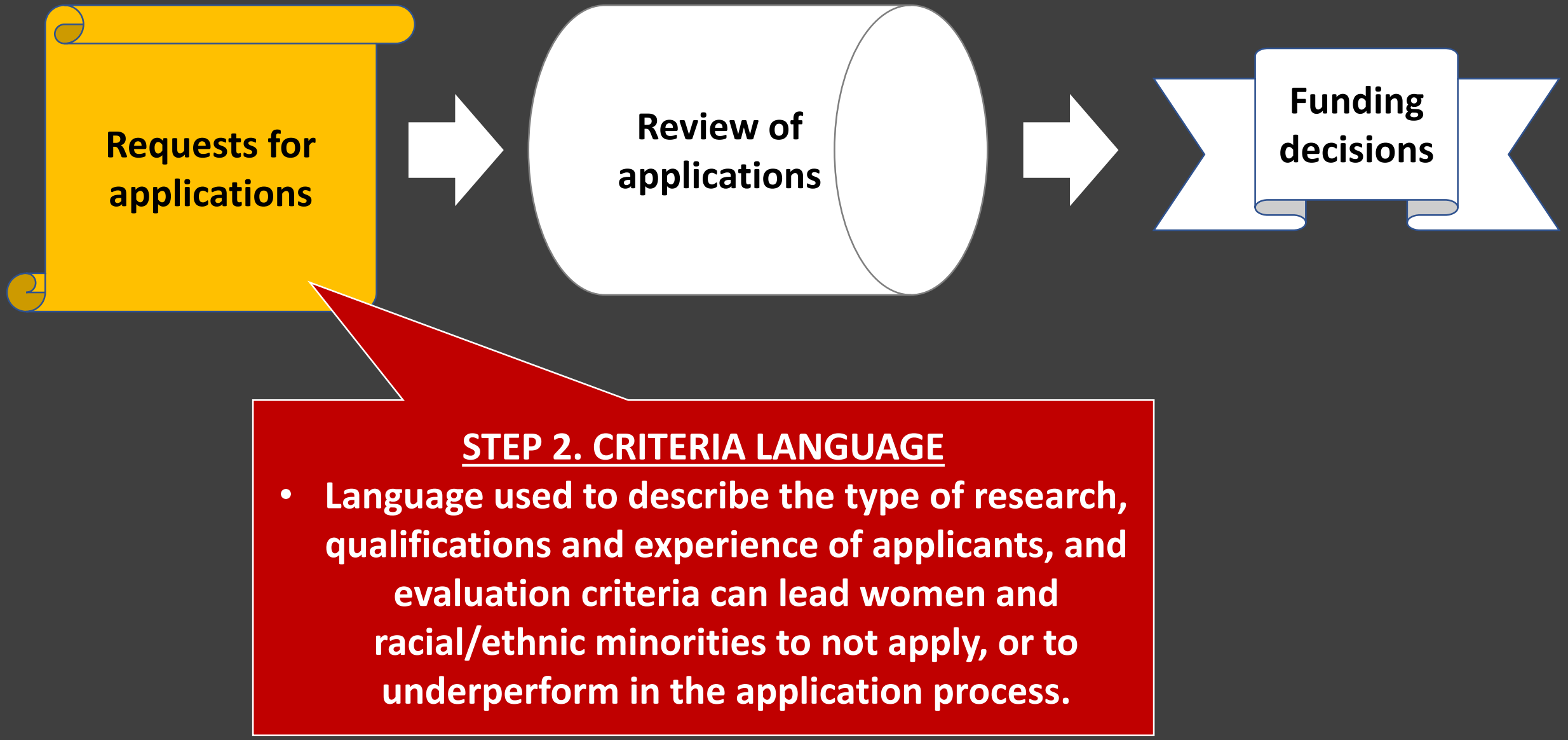
Scientific Peer Review

Scientific Peer Review

Remember to CHECK:

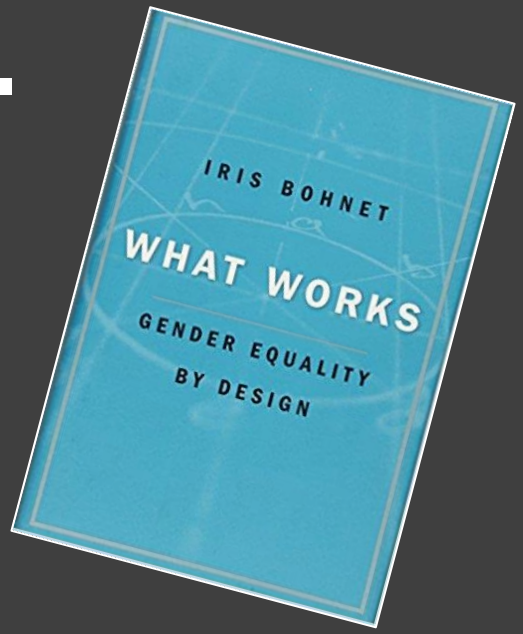
1. Participation
2. Criteria
4. Feedback


Requests for applications
Review of applications
Funding decisions
STEP 1. PARTICIPATION
• Who is getting access to your RFAs?
• Where are you advertising?
• How do you know you are reaching your target applicant pool?
• Do your “ideal” or target applicants have appropriate resources to apply?
• Is it easy to find information about your programs?
• What are the demographics of your applicants? Who are you missing?
What data could you be collecting here to learn more about if you are reaching your target applicant pool?


Requests for applications
Review of applications
Funding decisions
STEP 2. CRITERIA LANGUAGE
• Language used to describe the type of research, qualifications and experience of applicants, and evaluation criteria can lead women and racial/ethnic minorities to not apply, or to underperform in the application process.

Semantic “linguistic” priming.

- **Stereotypic-language causes readers to more easily and effortlessly “call to mind” people who align with those descriptors.**
  - **This is called “*priming*”.**



Women select

Helpful
Supportive
Writing
Mentoring
Clinical
Organizing
Teaching
Women’s health
Sensitive
Mild
Emotional

“Feminine language”

Lower status

Lead us to select women

“Masculine language”

High-risk
Leadership
Strong
Innovative
Experienced
Competent
Ground-breaking
Trailblazing
Independent
Dominant
Tough
Enterprising
Ambitious
Confident
Aggressive

Higher status

Men select

Lead us to select men

NIH Director's Pioneer Awards

Awards to women scientists

- 2004 = 0/9
- 2005 = 6/13 = 43%
- 2006 = 4/13 = 31%
- 2007 = 4/12 = 33%
- 2008 = 4/16 = 25%
- 2009 = 7/18 = 39%

Semantic Priming in NIH Director's Pioneer Award?

2004	≥ 2005
Characteristics of target scientist and research	
Risk-taking emphasized: • "exceptional minds willing and able to explore ideas that were considered risky" • "take…risks" • "aggressive risk-taking" • "high risk/high impact research" • "take intellectual risks" • application URL included "highrisk"	Emphasis on risk removed: • "pioneering approaches" • "potential to produce an unusually high impact" • "ideas that have the potential for high impact" • "highly innovative" • URL no longer includes "risk"

Carnes et al. *J Womens Health*, 2005

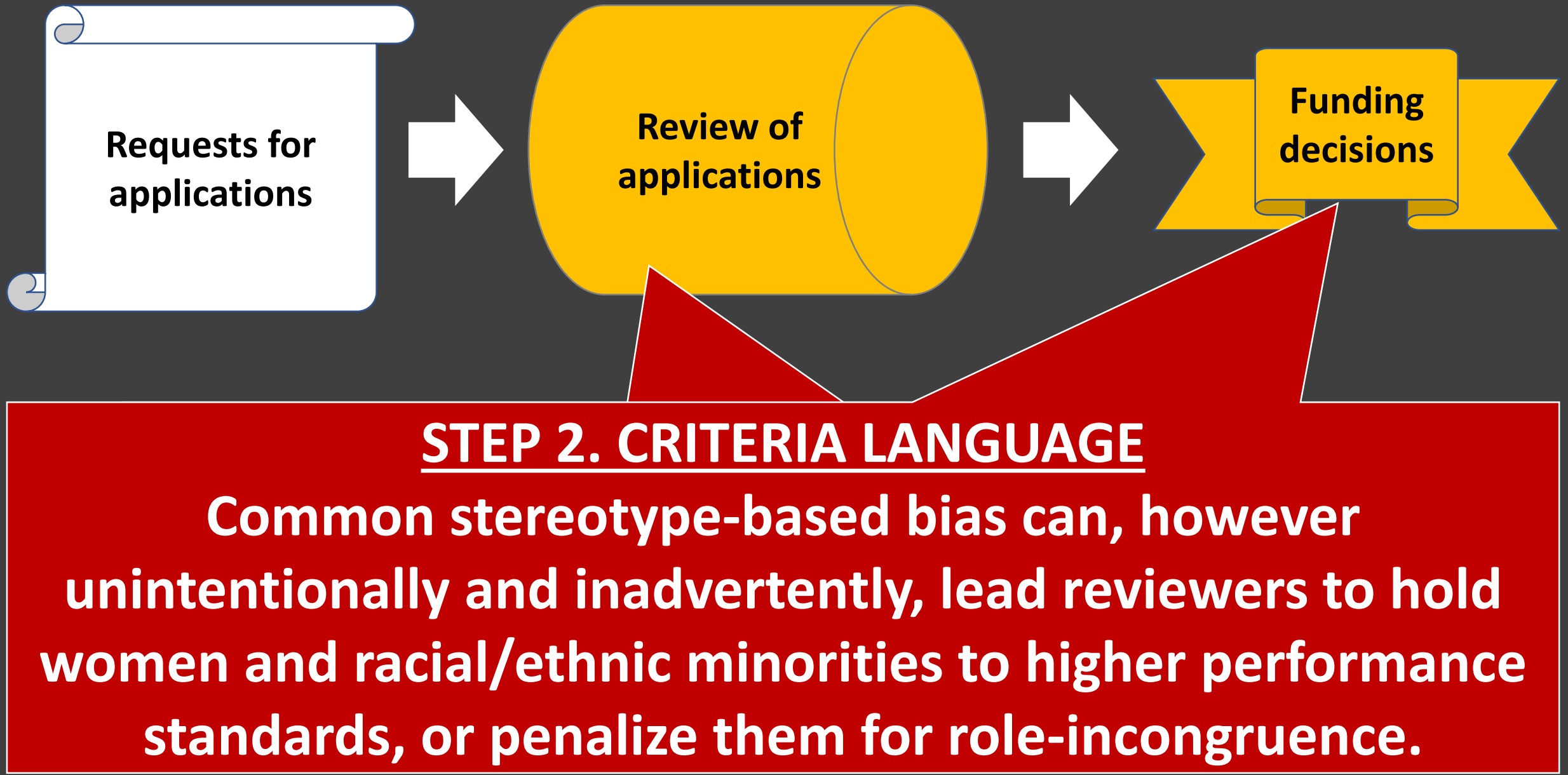
Carnes, M, Geller, S, Fine, E, Sheridan, J and J Handelsman (2005). "NIH Director's Pioneer awards: Could the selection process Be biased against women." J Womens Health 14(8): 684-691

2004 (0/9)	2005 (96/14=43%) 2006 (4/13=31%)
-"High risk", "aggressive" -Potential for scientific leadership -Testimony of intrinsic motivation, enthusiasm, and intellectual energy	-No use of risk language. -Relevance of the research and impact on the scientific field and on the NIH mission. Motivation/enthusiasm/intellectual energy to pursue a challenging problem.

Marchant, A, Bhattacharya, A, and M Carnes (2007). "Can the language of tenure criteria influence women's academic advancement?" J Womens Health 16(7): 998-1003

***Methods*: We used a retrospective, descriptive design to study 24 academic medical centers top-ranked in both NIH funding and Carnegie classification. The main outcome measure was the slope of regression fit to 7-year annual data on percent faculty who are tenured women (1998–2004) relative to the median slope of all 24 institutions.**

***Results*: Medical schools with the word "leader" in tenure criteria were more likely to have slopes below the median slope than schools without the word "leader" (OR = 6.0; CI = 1.02, 35.37; *p* = 0.04).**


Requests for applications
Review of applications
Funding decisions
STEP 2. CRITERIA LANGUAGE
Common stereotype-based bias can, however unintentionally and inadvertently, lead reviewers to hold women and racial/ethnic minorities to higher performance standards, or penalize them for role-incongruence.

Stereotype-based bias disadvantages women and racial/ethnic minorities in peer review.

Semantic “linguistic” priming.

Competency bias.

Consequences of stereotype-based bias.

- Stereotypes lead to assumptions that women and URMs lack competence, so we may:
  - Doubt their ability (Trix and Psenka, 2003; Heilman, 2007).
  - Require more proof (i.e., more accomplishments) to confirm their competence (Biernat, 1997, 2012; Heilman 2004, 2007; Kaatz, 2015).
  - Attribute their accomplishments to others or devalue accomplishments (Heilman, 2007).
  - Require a higher quality of work (Biernat, 1997, 2012; Heilman, 2007; Kaatz, 2015).

Competency bias impacts how we value men and women's ability, work, and accomplishments.

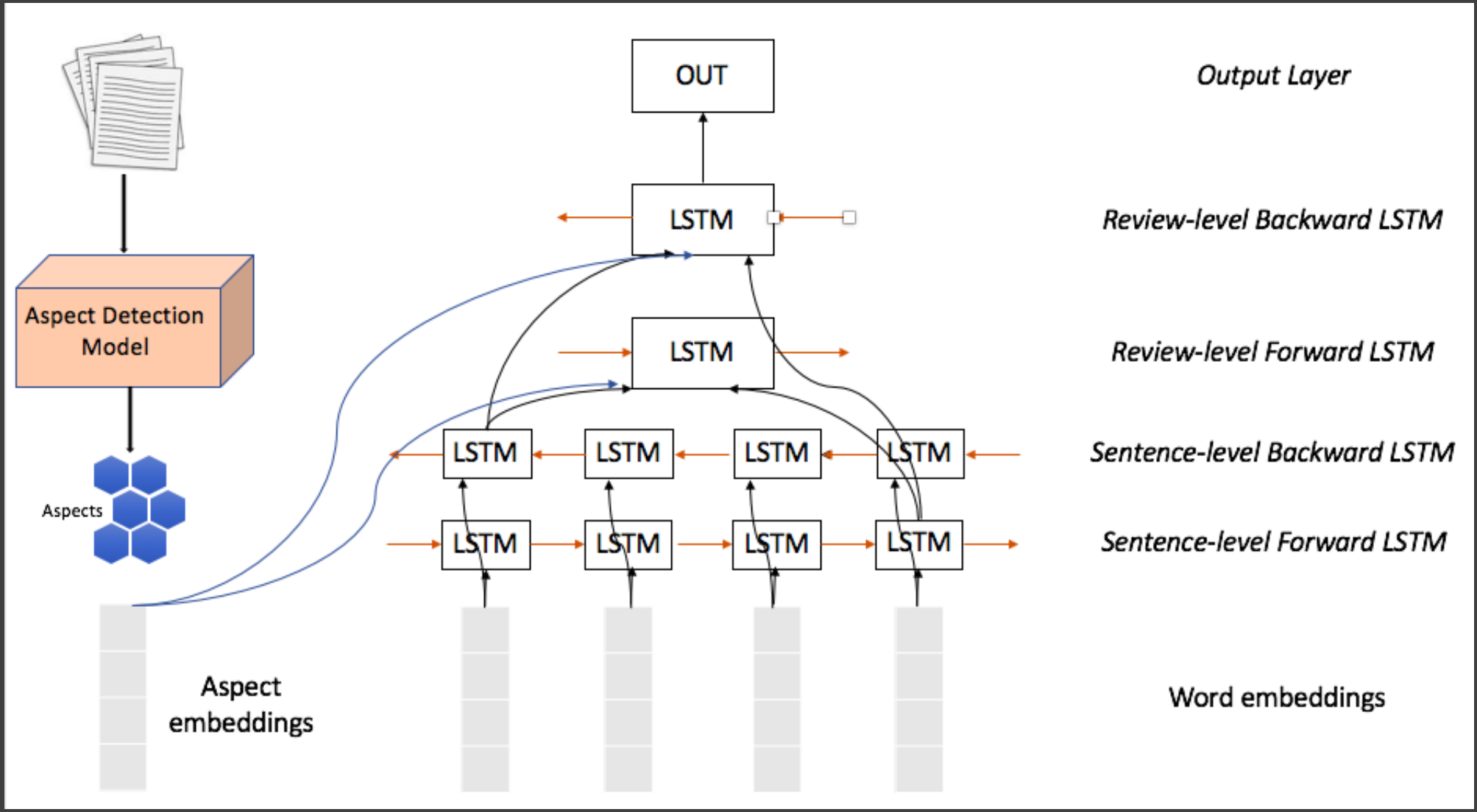
- **Moss-Racusin, C. et al. (2012). "Science faculty's subtle gender biases favor male students." PNAS 109: 16474-16479.**
  - 127 Faculty from Biology, Chemistry and Physics departments participated.
  - Evaluated application randomly assigned male or female name for:
    - Competence, hireability, likeability, starting salary, and willingness to provide mentoring.

- Results:
  Male applicant rated significantly more competent and hirable than female applicant and was granted a higher starting salary and offered more mentoring.

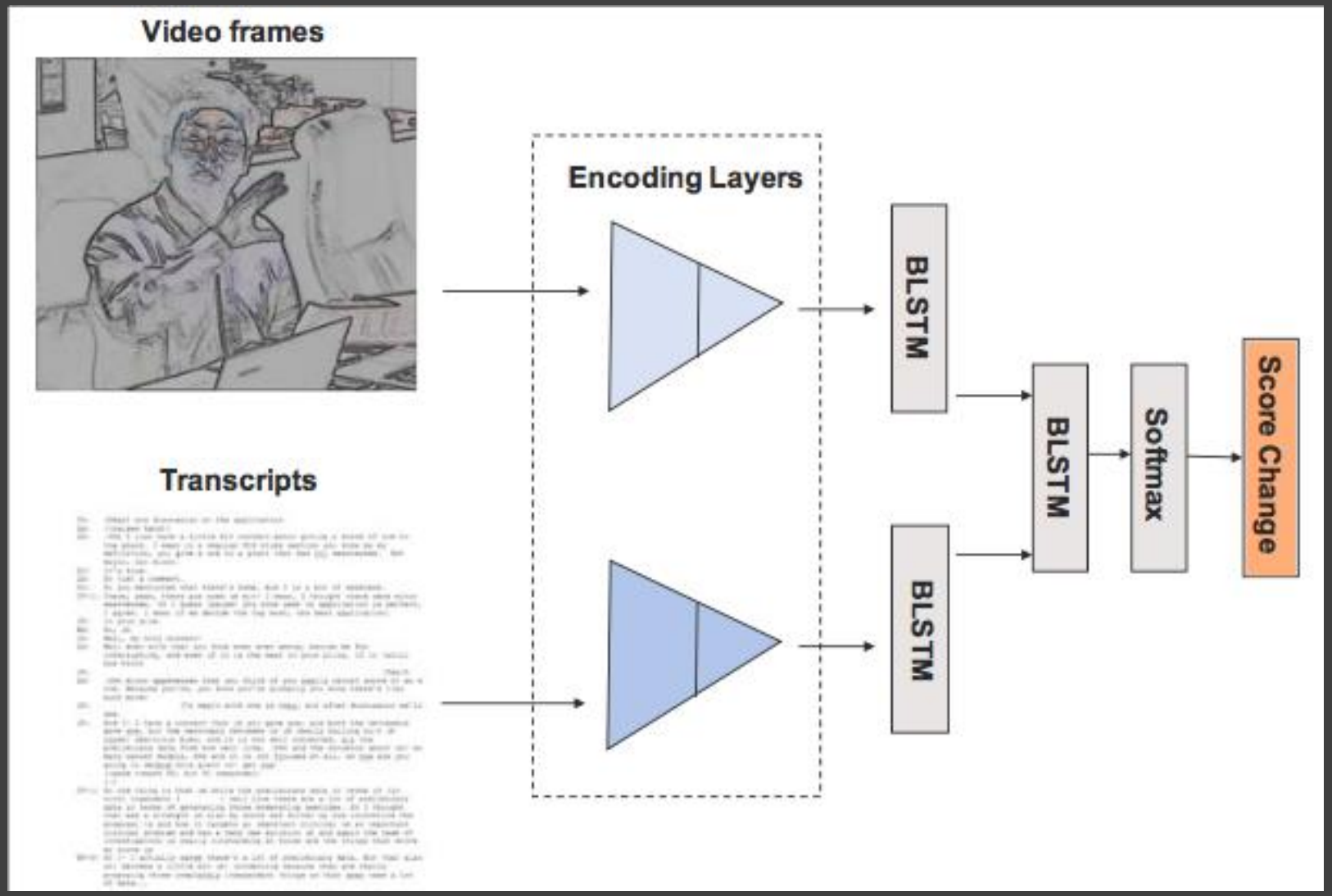
Competency bias impacts evaluation of CVs, content of rec letters, and peer review for research funding.

- Male scientists rated as more hirable, competent, & qualified than female scientists for faculty position despite same CV (Goldberg Design; Steinpreis et al., 1999).
- Letters of recommendation for women in science and med are shorter, have more references to personal life, and contain fewer “outstanding” descriptors (Trix & Psenka, Discourse & Soc, 2003; Schmader et al., 2007).
- Women, and Black physicians and scientists who submit R01 proposals to NIH are significantly less likely than men, and Whites, respectively, to be funded. And text analysis of grant critiques suggests that women may be held to higher standards than men to earn fundable application scores (Ley & Hamilton Science, 2008; Pohlhaus et al., 2011; NIH ,2015; Kaatz et al., 2015; 2016).

Hierarchical LSTMs

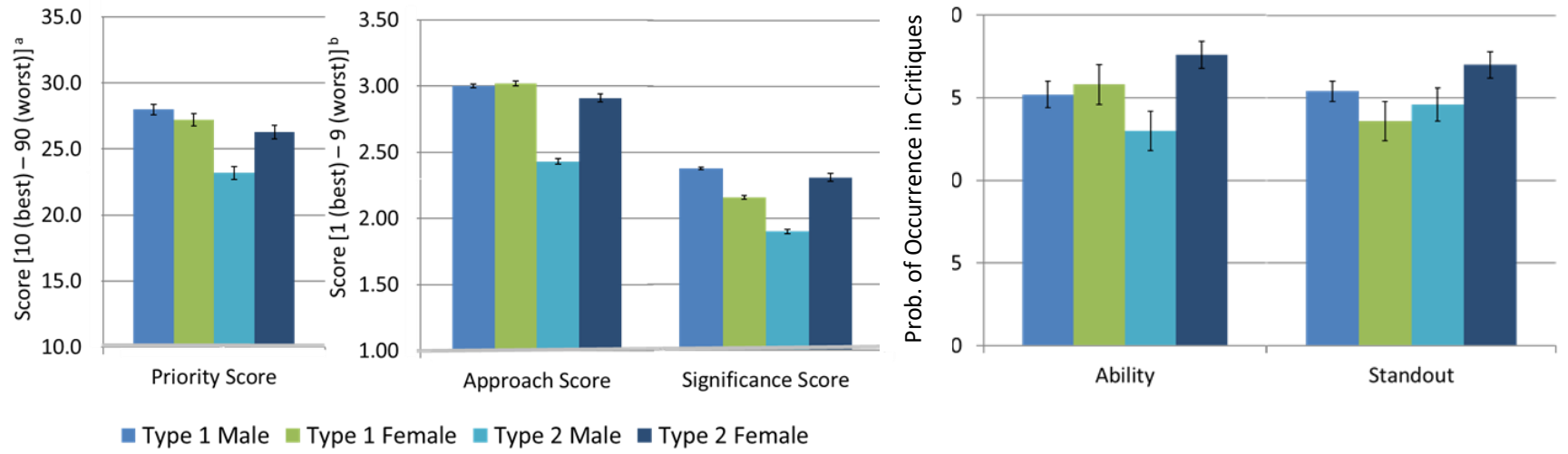


Fusion LSTMs



Kaatz A, Lee Y-G, Potvien A, et al. Analysis of National Institutes of Health R01 Application Critiques, Impact, and Criteria Scores: Does the Sex of the PI Make a Difference? Acad Med. 2016;91(8):1080-1088.

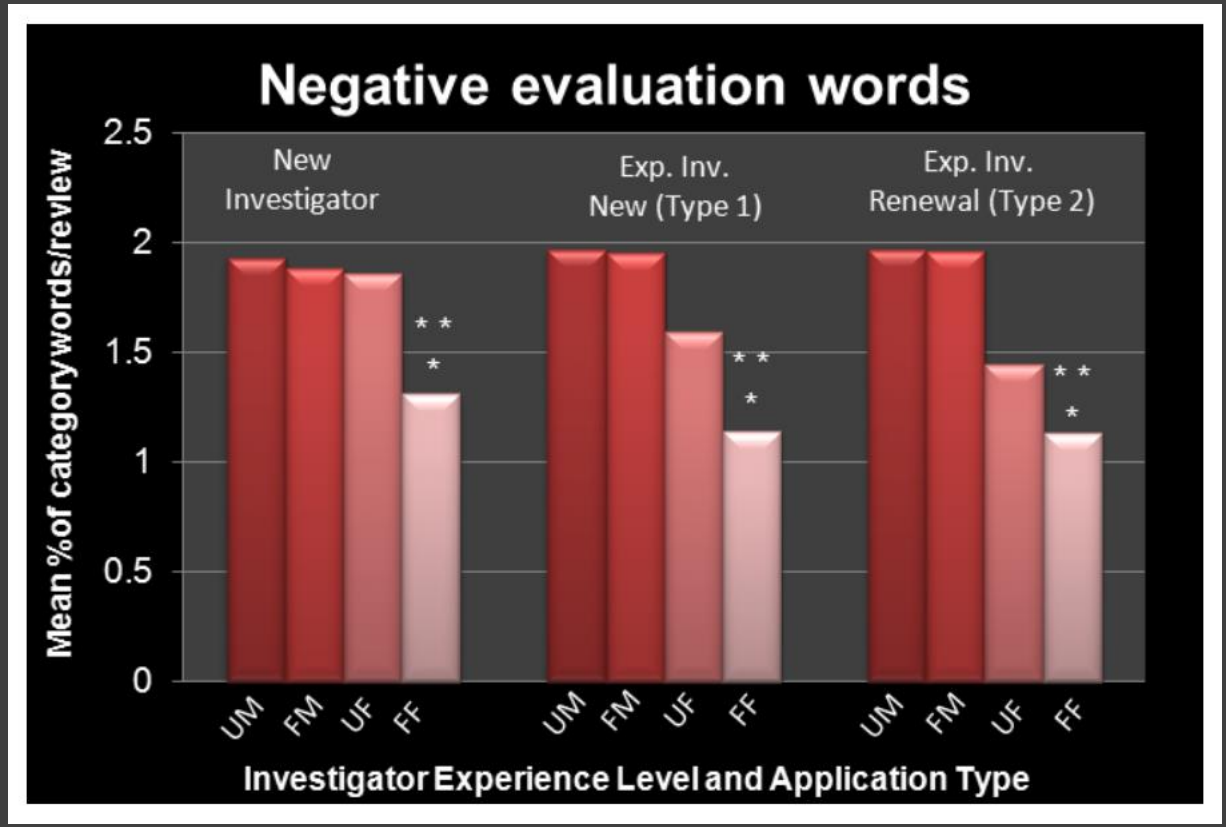
Study of 739 R01 Grant Critiques and Scores, University of Wisconsin-Madison (125 PIs): 2010-2014.

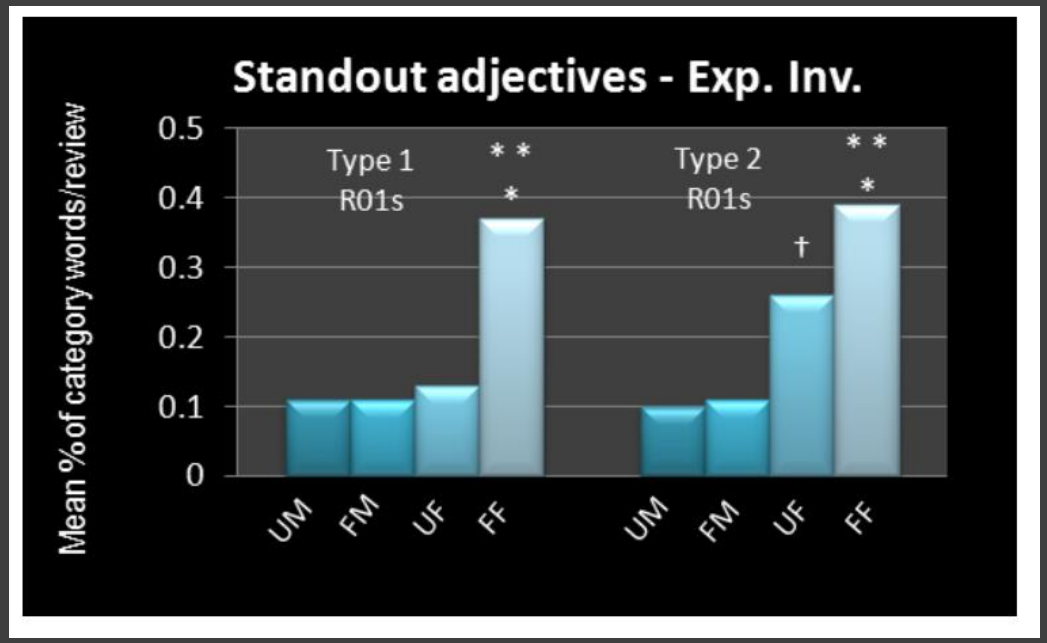


Female PIs' R01 Renewals Assigned Worse Priority, Approach, and Significance Scores, Despite More Praise in the Critiques

NOTE: PIs in sample had similar levels of productivity and background qualifications.

Men's R01 applications funded despite significantly more negative eval words (e.g., illogical, wrong, unclear) in critiques.

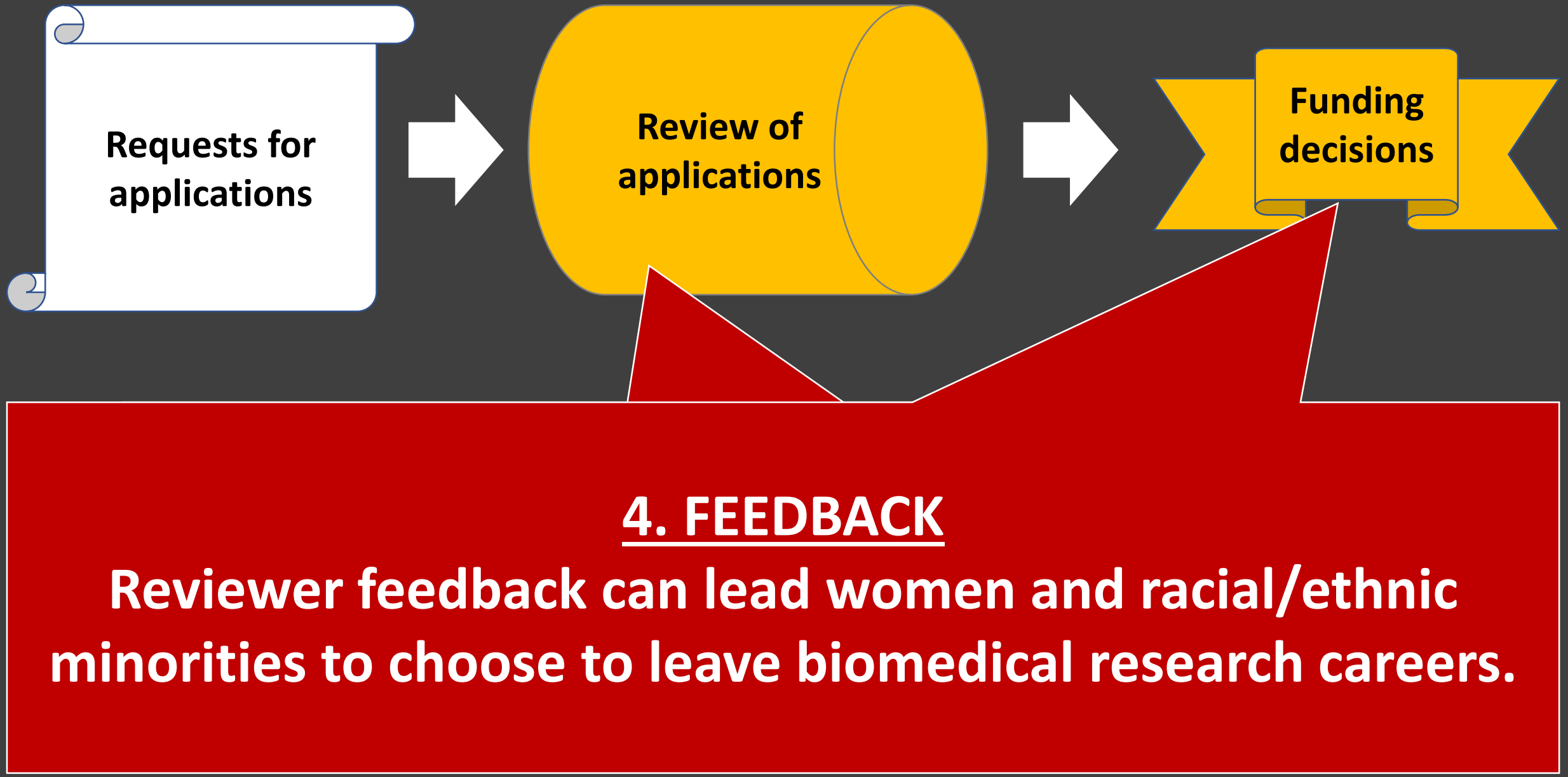




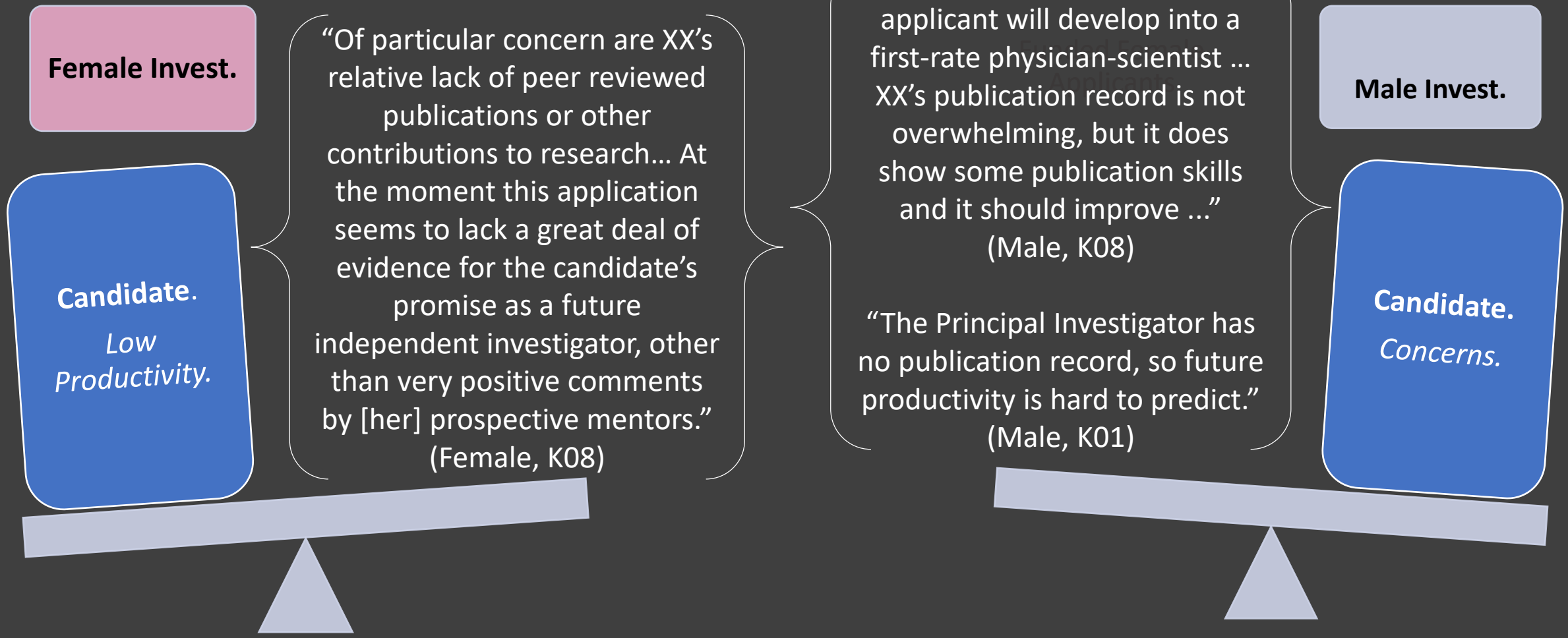
To be funded, women's critiques had to contain significantly more standout adjectives (e.g., outstanding, exceptional, excellent).

Text analysis of NIH grant critiques provides a window in to reviewers' cognitive processing.

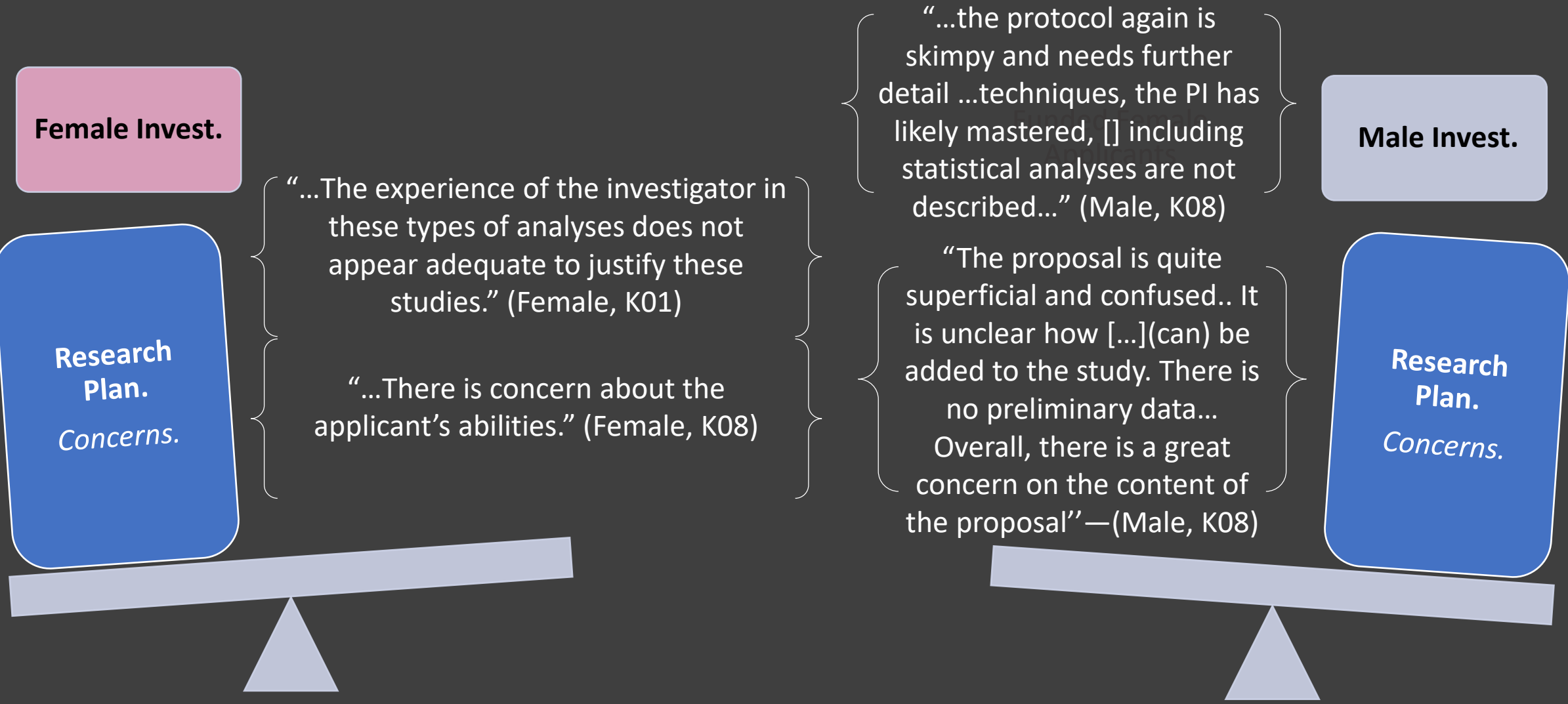
STUDIES OF NIH REVIEWERS' CRITIQUES
Kaatz, A., W. Magua, D. R. Zimmerman and M. Carnes (2015). "A quantitative linguistic analysis of National Institutes of Health R01 application critiques from investigators at one institution."Acad Med 90(1): 69-75.
Kaatz, A., Y. G. Lee, a. Potvien, W. Magua, A. Filut, A. Bhattacharya, R. Leatherberry, X. Zhu and M. Carnes (2016). "Analysis of National Institutes of Health R01 application critiques, impact, and criteria scores: Does sex of the principal investigator make a difference?"Acad Med 91(8): 1080-1088.
Kaatz, A., M. Dattalo, C. Regner, A. Filut and M. Carnes (2016). "Patterns of feedback on the bridge to independence: A qualitative thematic analysis of NIH mentored career development award application critiques."J Womens Health25(1): 78-90.
Magua, W., X. Zhu, A. Bhattacharya, A. Filut, A. Potvien, R. Leatherberry, Y. G. Lee, M. Jens, D. Malikireddy, M. Carnes and A. Kaatz (2017). "Are female applicants disadvantaged in National Institutes of Health peer review? Combining Algorithmic Text mining and qualitative methods to detect evaluative differences in R01 reviewers' critiques."J Womens Health (Larchmt)**26**(5): 560-570.


Requests for applications
Review of applications
Funding decisions
4. FEEDBACK
Reviewer feedback can lead women and racial/ethnic minorities to choose to leave biomedical research careers.

Candidate.


Female Invest.
Candidate.
Low Productivity.
"Of particular concern are XX's relative lack of peer reviewed publications or other contributions to research… At the moment this application seems to lack a great deal of evidence for the candidate's promise as a future independent investigator, other than very positive comments by [her] prospective mentors." (Female, K08)
"… there is confidence that the applicant will develop into a first-rate physician-scientist … XX's publication record is not overwhelming, but it does show some publication skills and it should improve ..." (Male, K08)
"The Principal Investigator has no publication record, so future productivity is hard to predict." (Male, K01)
Male Invest.
Candidate.
Concerns.

Research Plan. Assumption of competence for men



Research Plan: Assumption of competence for men.

Funded Female Applicants

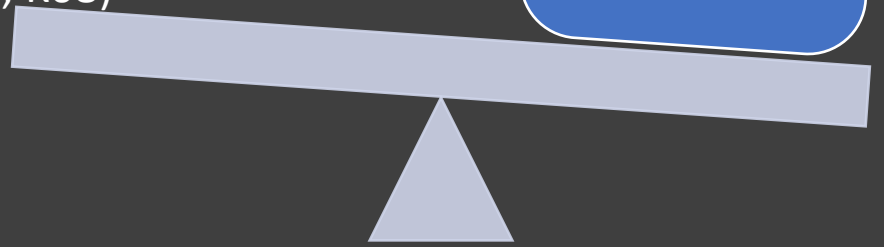
Male Invest.

"…the protocol again is skimpy and needs further detail …techniques, the PI has likely mastered, [] including statistical analyses are not described…" (Male, K08)

"The proposal is quite superficial and confused.. It is unclear how […](can) be added to the study. There is no preliminary data… Overall, there is a great concern on the content of the proposal''— (Male, K08)

Research Plan.

Concerns.

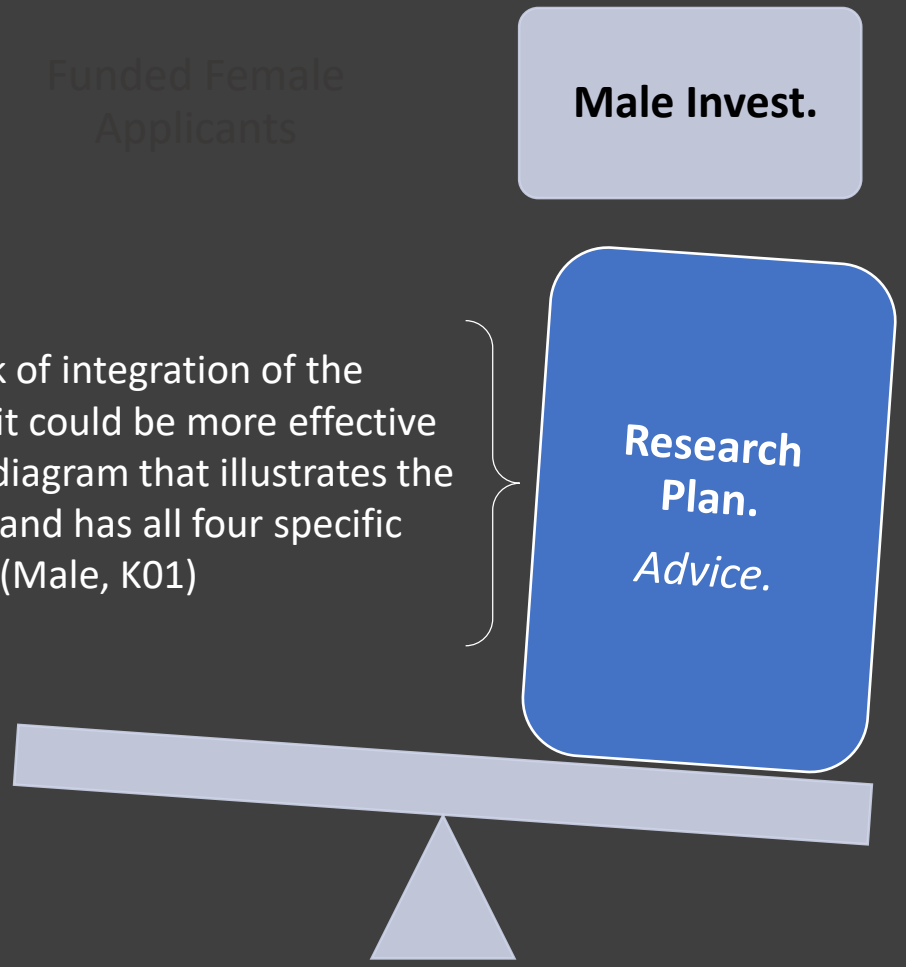


Research Plan.

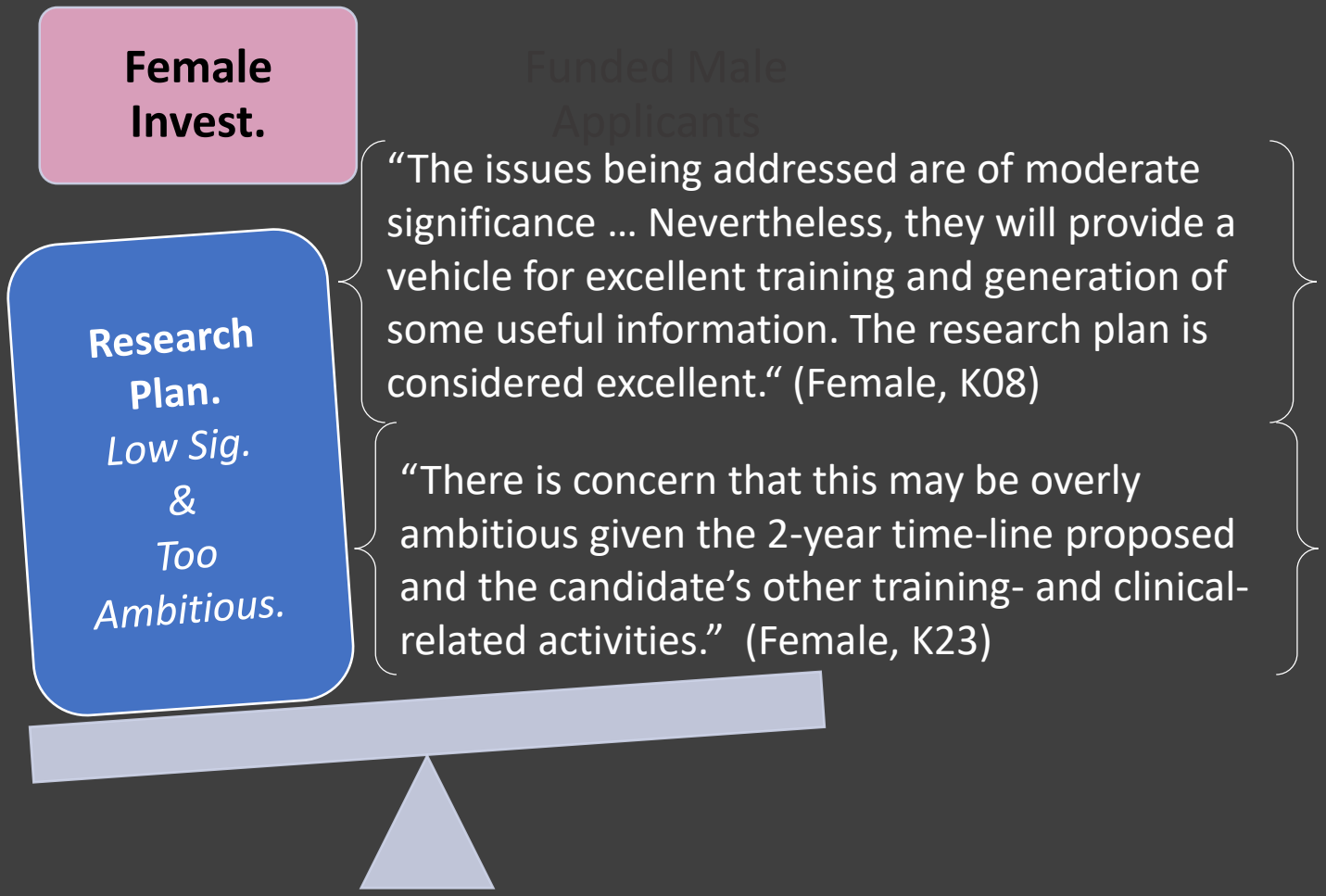
Funded Female Applicants

Male Invest.

"…A general concern is lack of integration of the research plan. Specifically, it could be more effective to present, for example, a diagram that illustrates the proposed causal pathways and has all four specific aims embedded within it." (Male, K01)

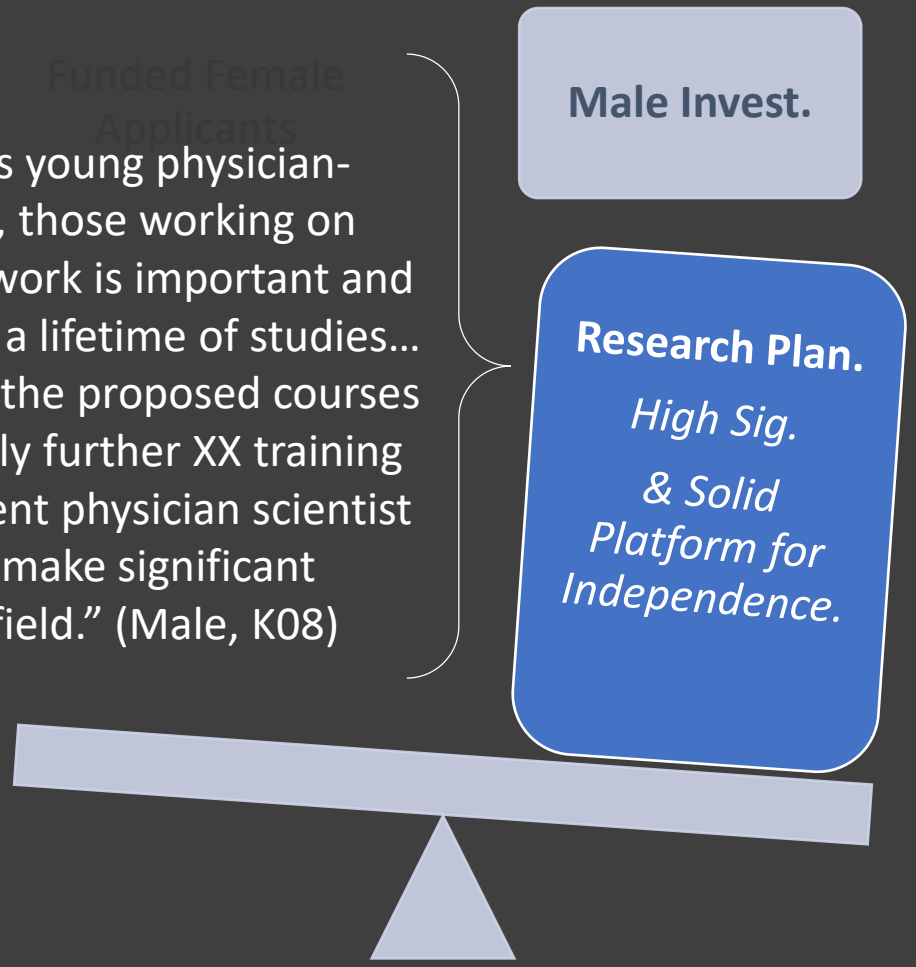


Research Plan.



Research Plan.

"The field of XX needs young physician-scientists, particularly, those working on [disease]… The focus of work is important and doable and could lead to a lifetime of studies… It is highly expected that the proposed courses and research will not only further XX training to become an independent physician scientist but will allow XX to make significant contributions in the field." (Male, K08)



Discussion.

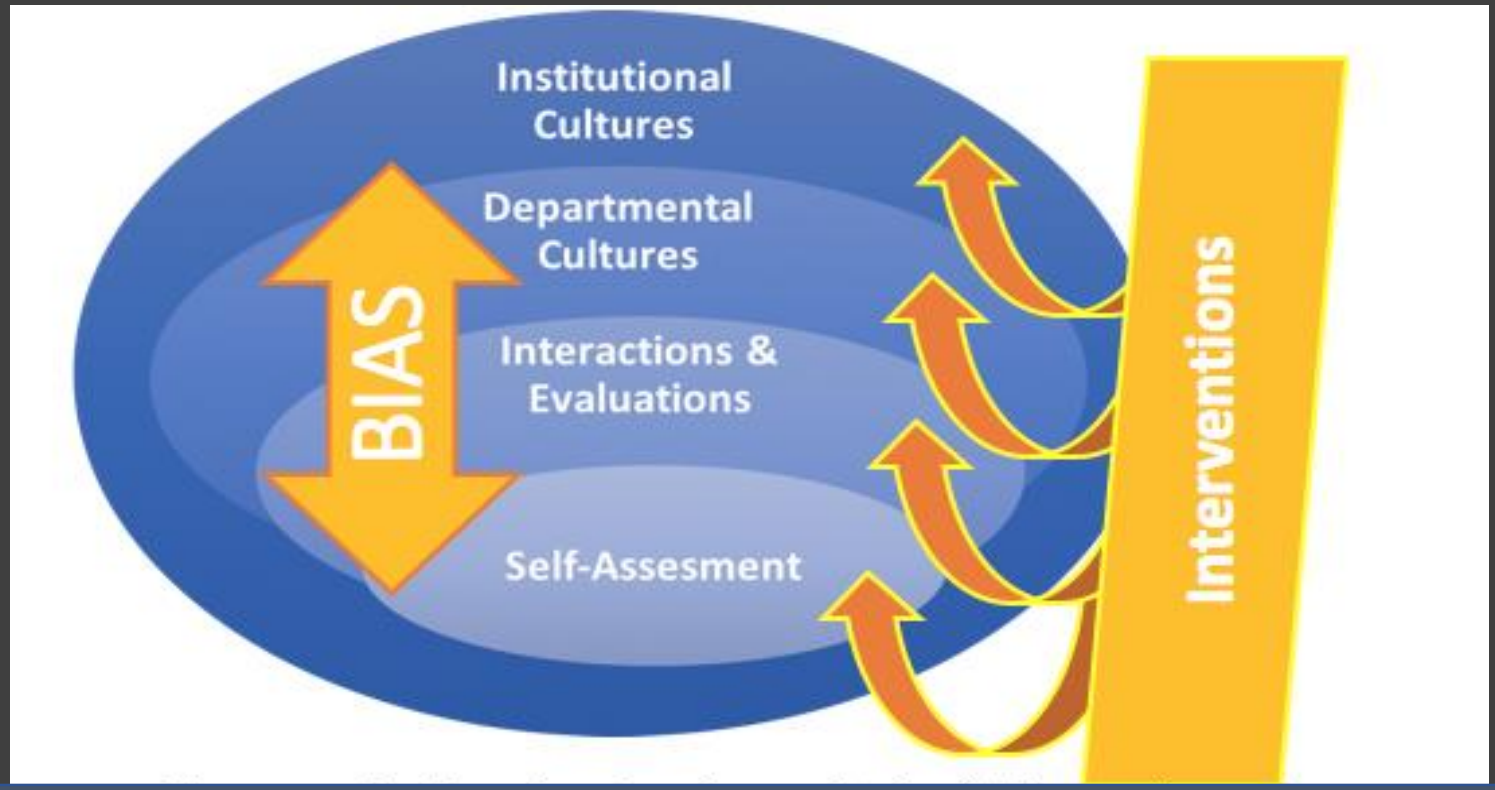
Three major “take-aways:”
1. Male and female K awardees may receive different feedback from NIH peer reviewers.
2. Only the female K awardees who were not exposed to critical feedback (i.e., who were “outstanding”) went on to obtain a subsequent R-level award.
3. Peer reviewers appear to assign different value to the research proposed by male and female K award applicants.

Follow-up studies.

Testing generalizability of study results and the impact of K award reviewer feedback on research career interest and persistence.
• **<u>Study 1:</u> Quantitative text analysis of a large national sample of K award application critiques to examine the extent to which results from this study generalize.**
• **<u>Study 2:</u> Experiment testing the impact of NIH peer reviewers' feedback (different types of critical remarks) on NIH K awardees' interest in research and decisions to apply for subsequent research awards (Biernat & Danaher, 2012).**

Think of a time when someone has made an incorrect assumption about you based on your membership in a group (e.g., gender, race, age, etc.) OR when you have made an incorrect assumption about someone (e.g., patient, colleague, etc.) based on their membership in a group.

How to address stereotype-based bias? Multilevel Interventions.



What can we do? Acknowledge bias.

Shift in Conceptual Framework

Old Framework:

Bias/Prejudice is bad so if I think or act with bias, I am a bad person.

New Framework:

Prejudiced thoughts and actions are habits that we all have and breaking these habits requires more than good intentions.

Implement evaluation practices that minimize bias.

Evidence-based practices.
• Establish and prioritize clear and specific evaluation criteria prior to evaluation. (Uhlman & Cohen, 2005,2007)
• Avoid abstract descriptors that promote stereotype assumptions. (Rubini & Menegatti, 2008; Wigboldus, Semin, & Spears, 2000)
• Where possible, consider removing identifiers from application materials, e.g., name, address, institution. (Goldin & Rouse, 2000)
• Spend sufficient time and attention on assessing each individual on the established criteria. (Martell, 1991)
• Justify decisions on the basis of established criteria and evidence from the evaluation materials.

What can institutions do to mitigate bias against women in hiring settings?

At least 1 RCT = level 1 evidence

- Infuse environment with statements that research evidence shows equivalent gender competence in relevant roles
- Encourage raters to take adequate time
- Allow applicants to provide individuating evidence of job-relevant competency
- Work for applicant pool to have at least 25% women
- Do not ask about parenthood status
- Use structured vs. unstructured interview questions
- Avoid man-suffix job titles (e.g. use chair rather than chairman)
- Use inclusion vs. exclusion strategy for selection from final list
- Implement training workshops for personnel decision-makers

Isaac, Lee, & Carnes. Acad Med, 84:1440-46, 2009

What can we do?

- Become “bias literate.”
  - **Defined by Sevo and Chubin as learning a vocabulary about bias and a skill set to recognize, understand, converse about, and intentionally practice cognitive and behavioral strategies to mitigate the impact of group stereotypes on judgment and decision making (Sevo and Chubin, 2008).**
  - **Used by Devine *et al.* to successfully reduce implicit race bias in students (Devine, Forscher, Austin and Cox, 2012) and by Carnes *et al*. to reduce biased behavior and improve department climate for faculty (Carnes *et al.*, 2015) in academic science fields.**

What can we do? Two strategies that don't work.

Stereotype suppression (i.e., attempting to be "gender blind").

- **Monteith MJ, Sherman JW, Devine PG. Suppression as a stereotype control strategy. Pers Soc Psychol Rev. 1998;2:63–82.**

Strong belief in one's ability to make objective judgments.

- **Uhlmann EL, Cohen GL. "I think it, therefore it's true": Effects of self perceived objectivity on hiring discrimination. Organ Behav Hum Decision Proc. 2007;104:207–223.**

Both of these have been shown to enhance the infuence of stereotype-based bias on judgment.

Breaking the "bias habit."

- **Strategies that work:**
  - **question your own objectivity.**

    Uhlmann and Cohen, "I think it, therefore it is true: Effects of self-perceived objectivity on hiring discrimination," *Organizational Behavior and Human Decision Processes, 104;2 (*2007): 207-223

  - **stereotype replacement** (e.g., if girls are being portrayed as bad at math, identify this as a gender stereotype and consciously replace it with accurate information).

    Devine PG, Forscher PS, Austin AJ, Cox **32** WTL. Long-term reduction in implicit race prejudice: A prejudice habit-breaking intervention. J Exp Soc Psychol. **33** 2012;48:1267–1278.

  - **positive counterstereotype imaging** (e.g., before evaluating job applicants for a position traditionally held by men, imagine in detail an effective woman leader or scientist).

    Blair IV, Ma JE, Lenton AP. Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. J Pers Soc Psychol. 2001;81:828–841.

  - **perspective taking** (e.g., imagine in detail what it is like to be a person in a stereotyped group).

    Galinsky AD, Moskowitz GB. Perspective- taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. J Pers Soc Psychol. 2000;78:708–724.

  - **individuation** (e.g., gather specific information about a student or applicant to prevent group stereotypes from leading to potentially inaccurate assumptions).

    Heilman ME. Information as a deterrent against sex discrimination: The effects of applicant sex and information type on preliminary employment decisions. Organ Behav Hum Perform. 1984;33:174–186.

  - increasing opportunities for contact with **counterstereotypic exemplars** (e.g., meet with senior women faculty to discuss their ideas and vision).

    Allport GW. The Nature of Prejudice. Reading, Mass: Addison-Wesley Publishing Company; 1979.

Practice the right message.

Recite this mantra:

"The vast majority of people try to overcome their stereotypic preconceptions."

This message reduced weight, age, and gender bias vs. a message that
we all have bias.

Duguid & Thomas-Hunt, 2015

UW-Madison fosters growth mindsets

"Everyone has bias"

Vs.

"We are all working to reduce bias"

What makes us the same...is that we are all different

www.lacse.fr

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
l'acsé
l'agence nationale pour la cohésion sociale et l'égalité des chances