STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

STScI Dual-Anonymous Peer Reviews

Kenneth Sembach

Director

SoPWG Meeting, 18 February 2021

STScI dual-anonymous peer review

STScI instituted a dual-anonymous review for the Hubble Space Telescope in October 2018 to mitigate unconscious biases present in the review process.

- Proposers do not know identity of reviewers, and reviewers do not know identity of proposers.
- First such peer review in astrophysics
- Reducing gender bias was the impetus for the change, but other biases are mitigated as well.

Given the successful implementation and positive outcome, STScI adopted the dual-anonymous approach for all future HST and JWST peer reviews.

- Others have followed suit (see backup slides)

Gender and HST proposal selection statistics (Cycles 11-21)

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC, **126**:923–934, 2014 October

Gender-Correlated Systematics in HST Proposal Selection

I. NEILL REID

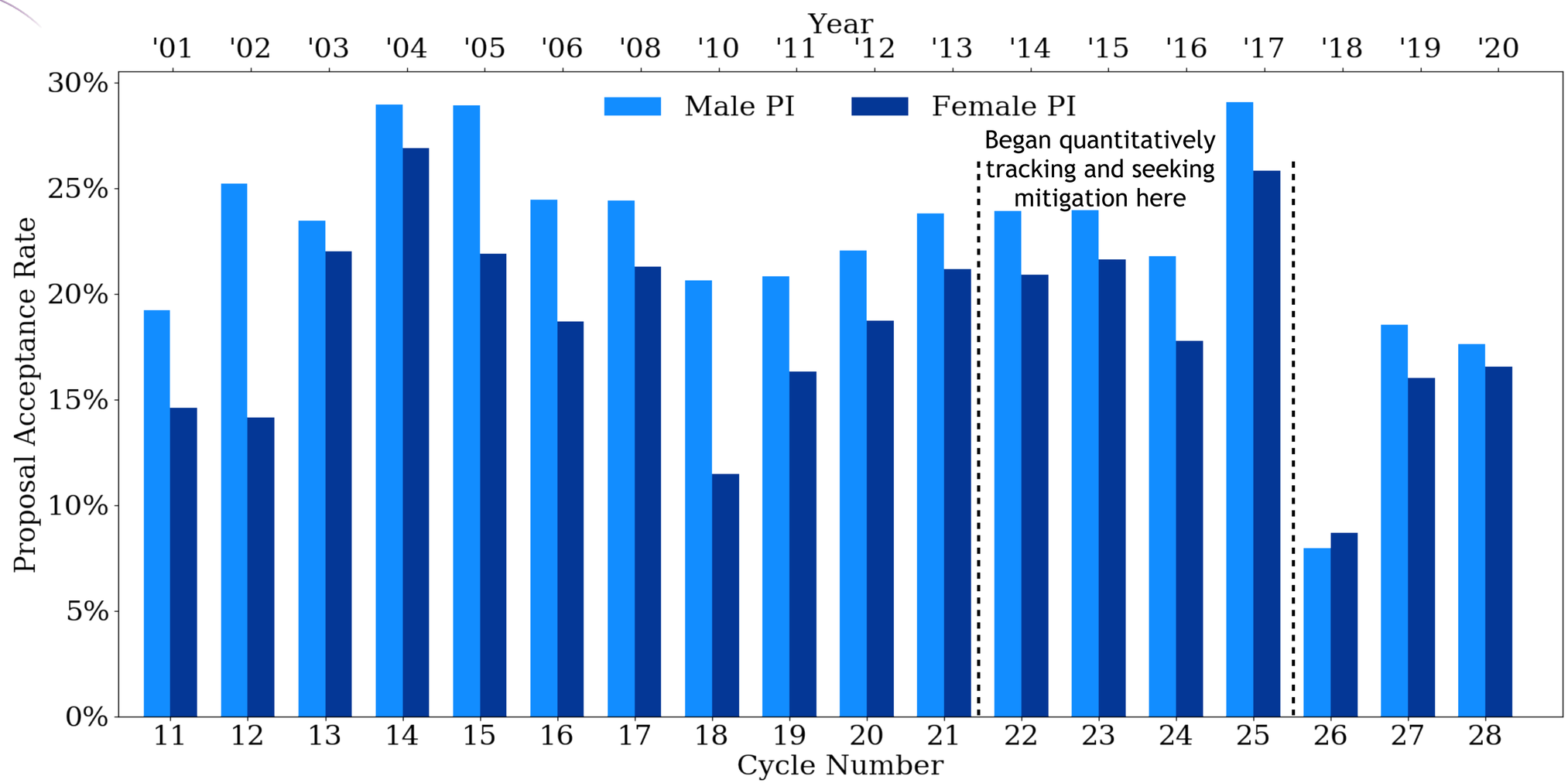
Space Telescope Science Institute, Baltimore, MD 21218

Received 2014 May 12; accepted 2014 September 09; published 2014 October 21

ABSTRACT. Proposal success rates are calculated for HST Cycles 11 through 21 as a function of the gender of the Principal Investigator (PI). In each cycle, proposals with male PIs have a higher success rate, with the disparity greatest for Cycles 12 and 18. The offsets are small enough that they might be ascribed to chance for any single cycle, but the consistent pattern suggests the presence of a systematic effect. Closer inspection of results from Cycles 19, 20, and 21 shows that the systematic difference does not appear to depend on the geographic origin of the proposal nor does it depend on the gender distribution on the review panels. Segregating proposals by the seniority of the PI, the success rates by gender for more recent graduates (Ph.D. since 2000) are more closely comparable. There is also a correlation between success by gender and the average seniority of the review panel for Cycles 19 and 20, but not Cycle 21. We discuss these results and some consequent changes to the proposal format and additions to the HST TAC orientation process.

Online material: color figures

Gender and HST proposal selection statistics



Prior to Cycle 26, proposals led by male principal investigators had a consistently higher success rate than those led by female principal investigators.

Initial attempts to mitigate gender bias

Proposal formats were adjusted to limit principal investigator information

- Cycles 22/23: PI name was removed from front page of proposal and filename, and “Previous Observations” section of proposal was eliminated.
- Cycle 24: Initials replaced forenames in investigator list.
- Cycle 25: Alphabetical listing of investigators, PI not identified

No significant impact on the outcomes

- These incremental changes did not have the desired effect of reducing the acceptance rate offsets.
- The gender bias remained at levels similar to previous cycles.

An expert, external perspective on the HST peer review

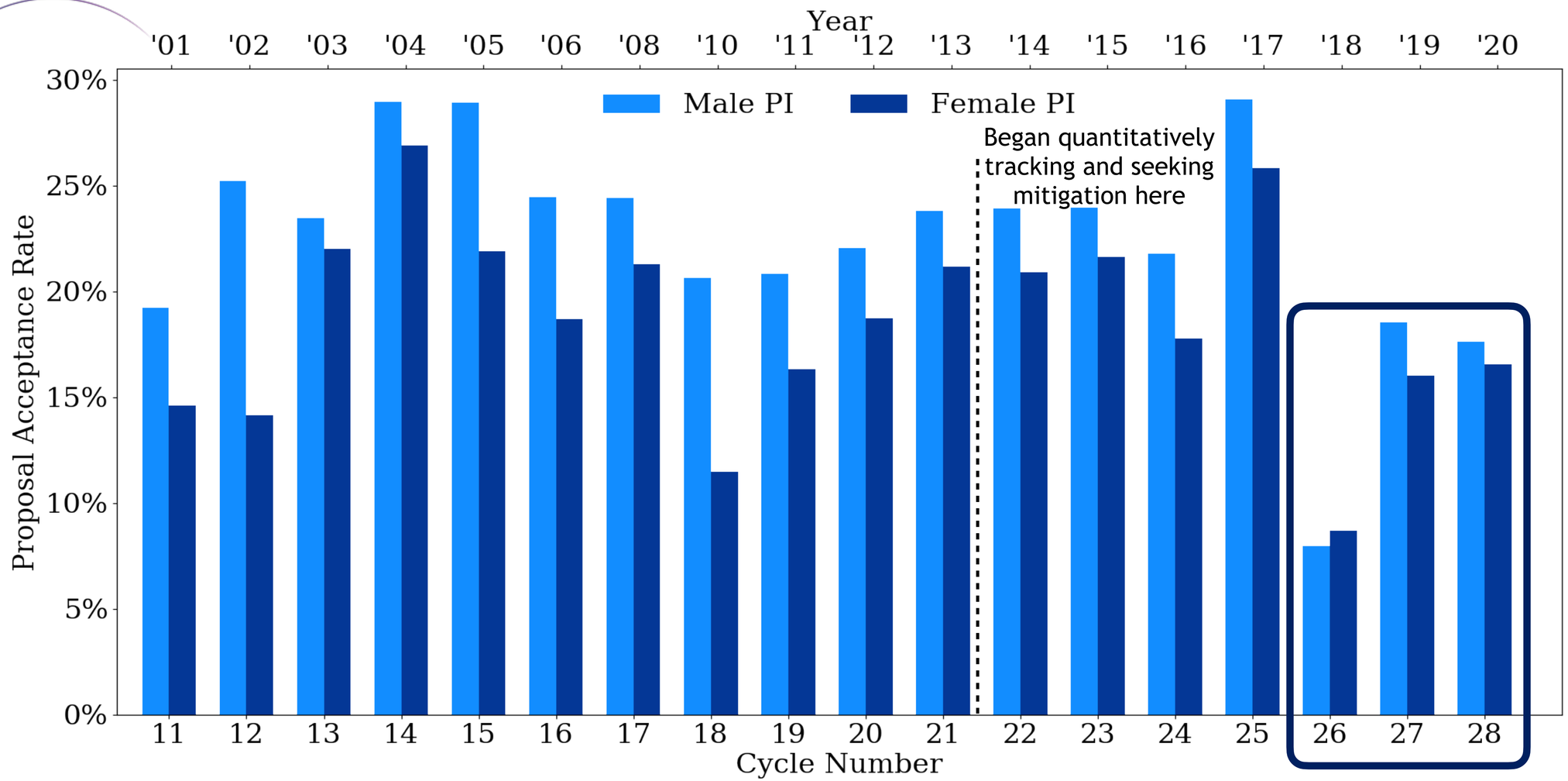
External consultants analyzed grade statistics and observed the time allocation committee process.

- Prof. Stefanie Johnson (U. Colorado), Prof. Jessica Kirk (U. Memphis)
- In Cycle 21 (PI plainly identified), male reviewers systematically downgraded female-PI proposals.
- In Cycles 24 and 25 (PI not identified), no systematic difference was found in preliminary grades, but the final results showed the same systematic offsets favoring male-PI proposals.
- Observed Cycle 25 panels, noted that ~60% of proposal conversations included discussion of the PI or team qualifications, previous work, or citation rates

Recommendation (October 2017):

- Fully anonymize the proposal review in Cycle 26.
- STScI established a working group to consider impact on proposers and implementation (see backup slides for website with report/resources).

Gender and HST proposal selection statistics



Cycles 11-25
$\langle\Delta\rangle = 5\%$

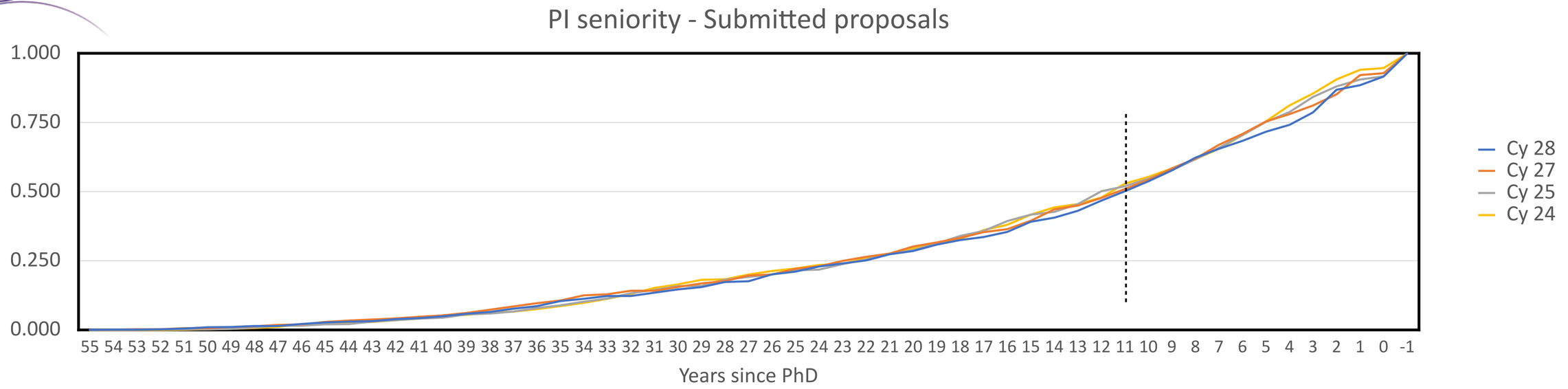
Cycles 26-28
$\langle\Delta\rangle = 1\%$

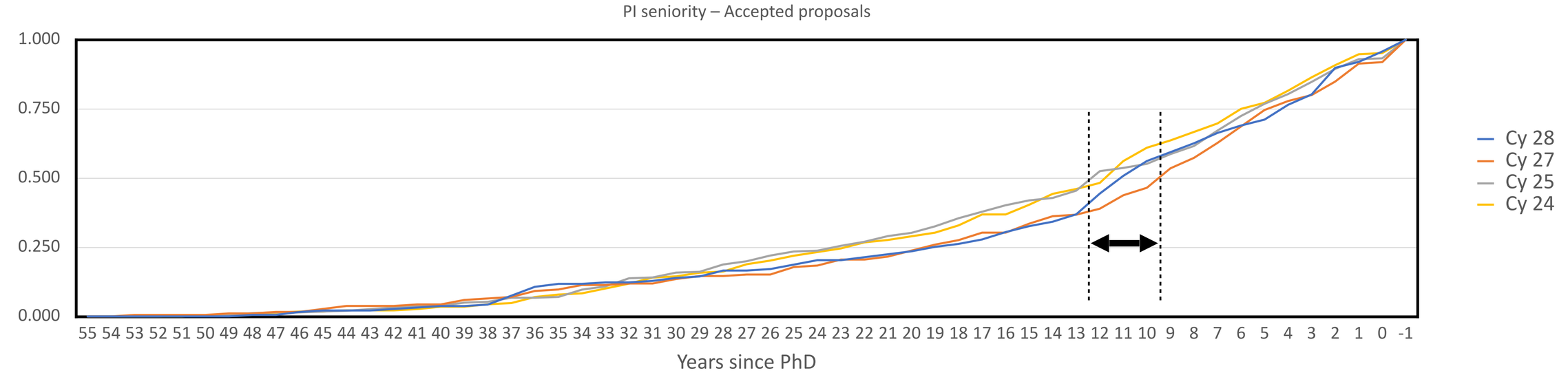
Dual anonymous review

Prior to Cycle 26, proposals led by male principal investigators had a consistently higher success rate than those led by female principal investigators.

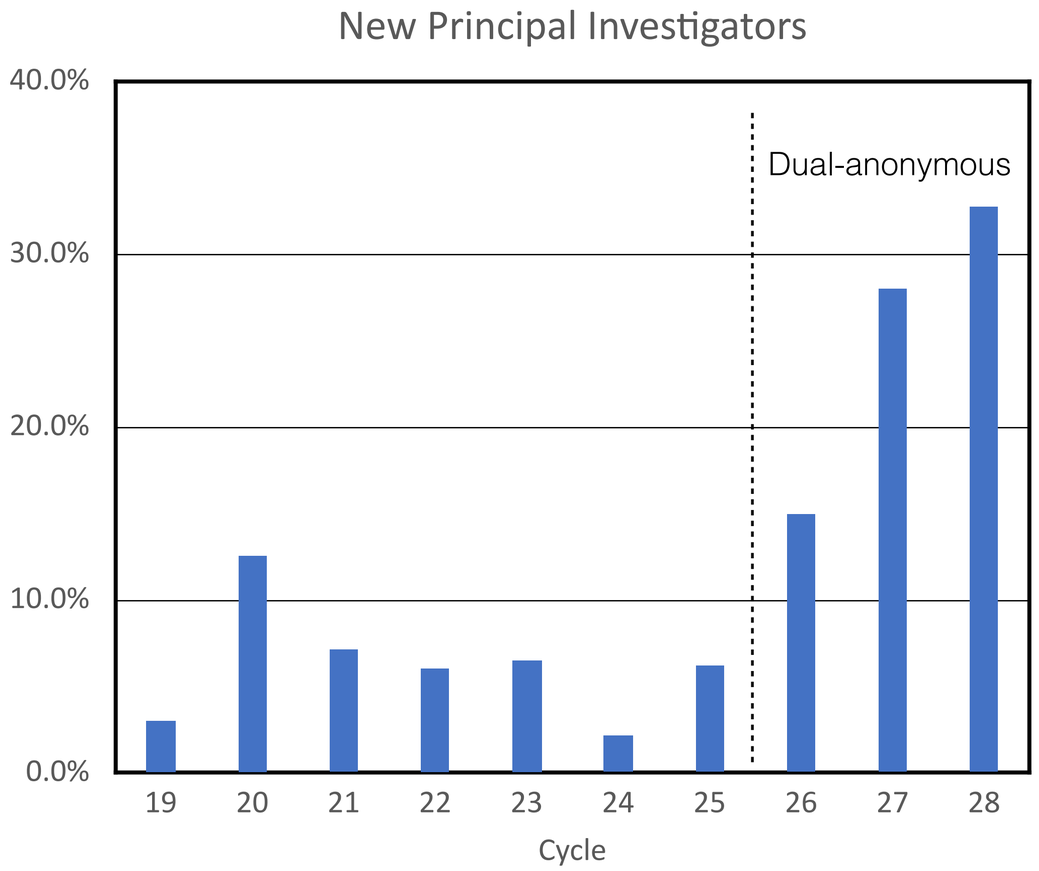
Many biases can be mitigated through dual-anonymous reviews

Principal Investigator seniority





New Principal Investigators – HST Cycles 19-28



Cycle	New PIs	Total accepted proposals	Fraction
28	55	168	33%
27	51	182	28%
26	6	40	15%
25	21	340	6%
24	5	228	2%
23	17	261	7%
22	16	263	6%
21	18	253	7%
20	29	231	13%
19	6	196	3%

HST proposal preparation and review

The HST Cycle 29 Call for Proposals has dual-anonymous guidelines on proposal compliance for proposers and reviewers, examples, and links to relevant content.

- Proposer Guidelines in Anonymous Reviews
- Reviewer Guidelines in Anonymous Reviews
- FAQ on Anonymizing Proposal Reviews
- Example text for anonymous proposing

Intention is NOT to make it impossible to guess the authors of a proposal.

- Aim is to have the reviewers focus on the proposed science, not the scientists.
- Extensive examples are provided on neutral ways of citing previous observations, theoretical results, papers, software tools, etc.

Reviewers are asked to flag potential non-compliant proposals.

- Vast majority of proposals are compliant; only a few (out of thousands) disqualified.
- Feedback is provided to those with minor infractions.

The review process - science evaluation

Proposals have a two-stage science review.

- Preliminary grades submitted prior to the meeting are used to triage the lowest-ranked proposals.
- Final grades are assigned after discussion at the face-to-face panel meeting.

Levelers attend each review panel.

- Levelers are observers tasked with ensuring panel discussions stay focused on science and the science merit of the proposal, not the PI or proposing team.
- Levelers have authority to stop/redirect discussion of a proposal.

During TAC orientation, reviewers and levelers are briefed on the expectations for discussing the anonymized proposals and are given basic guidance in avoiding/minimizing unconscious (implicit) bias.

The review process - investigator evaluation

Proposers provide a team expertise summary in addition to the standard proposal.

- This summary is not seen by reviewers until after the science ranking.

Reviewers are asked for a binary (qualified/unqualified) assessment on team expertise for those proposals passing the science review.

- Expertise assessment does not change science ranking.
- Assessments are sent to STScI director (selecting official) for adjudication, if needed. No “team unqualified” assessments to date.

Handling conflicts of interest

STScI tracks personal involvement in proposals. Conflicts involving close collaborators and competitors are identified by panelists.

- Identified in advance of the peer review

Conflicts:

- Set on a personal basis: direct participation in a proposal, participation by family members, participation by close collaborators or close competitors (as defined by the reviewer)
- Institutional conflicts no longer flagged.
- Panelist leaves the room during the proposal discussion for any conflict.

Feedback

Outside observers commented that the panel discussions are focused almost exclusively on science rather than anecdotes about teams.

- Johnson and Kirk noted that none of the discussions they witnessed in Cycle 27 discussed the proposers.

Reviewers have noted that the review process is generally less stressful.

- Proposals are in many cases better written.
- More time to review (since less time discussing people)

Two quotes from the HST Cycle 27 TAC Chair (Rachel Sommerville)

"Discussions at both the panel level and TAC level focused predominantly on whether the science was novel, impactful, and feasible with HST, and not on whether the proposers had the expertise to carry out the proposal."

*"Several TAC members noted that they felt that the discussions at both the panel and TAC level seemed **more collegial and less emotionally charged** than previous TACs, perhaps because either positive or negative feelings about the people involved in the proposal were largely removed."*

Why did it take so long for us to get here?

Resistance

- Not in acknowledging that bias existed, but in how (or even if) to mitigate it
- Prominent, and at times coordinated, opposition
- Likely and unlikely opponents

It took time to build a coalition of the willing.

- Incremental steps were a start.
- Working group on anonymous peer review (WGAPR) engaged with the community.
- Space Telescope Institute Council and Space Telescope Users Committee provided helpful feedback and endorsed the dual-anonymous approach.
- NASA has been very supportive (Hertz, Zurbuchen, New, Evans) and shown real leadership in this area as well.

Summary

Proposal reviews are subject to biases.

- HST statistics show a systematic trend with PI gender over many observing cycles.
- To select the best science, focus the review on science not scientists.
- Assess investigator qualifications (if desired) after the science ranking.

We adopted a dual-anonymous approach to mitigate unconscious biases.

- We made a number of proposal format adjustments before moving to the dual anonymous review process.
- We sought expert external advice and community support to support this change.

So far, dual-anonymous proposal reviews seem to be working well.

- We find that gender offset in proposal acceptance rate may be reduced in scale.
- The substantial increase in new (to HST) PIs is interesting/encouraging as well.
- Reviewers and proposers have embraced this change.

Backup / Reference Material

National Aeronautics and
Space Administration
Headquarters
Washington, DC 20546-0001

June 7, 2019

To: Distribution (Astrophysics GO Leads)
From: SMD/Director of Astrophysics
Re: Dual Anonymous Peer Reviews for Astrophysics GO Programs

In June 2018, the Space Telescope Science Institute (STScI) conducted a dual anonymous peer review for Cycle 26 of the Hubble General Observer (GO) program[1]. The dual anonymous peer review addresses many issues of implicit bias. STScI's implementation of dual anonymous peer review was successful in Cycle 26. During June 2019, STScI will be conducting the Hubble Cycle 27 peer review, again using the dual anonymous process. STScI and NASA will review the Cycle 27 experience and outcomes to assess the dual anonymous practice.

In the absence of any contra-indications from the Hubble Cycle 27 peer review, I am directing all NASA Astrophysics GO programs to use dual anonymous peer reviews beginning in CY 2020.

In order to provide all NASA Astrophysics GO program leads with the benefit of STScI's experience, STScI will host a workshop in Fall 2019 to share their practices, lessons learned, and extant documentation with all other missions.

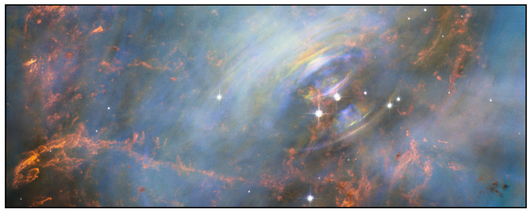
If you have any questions, please address them to your HQ Program Scientist or to me.

Paul Hertz
Director, Astrophysics Missions
Science Mission Directorate

Dual-Anonymous review workshop @ STScI

Dual-Anonymous Review Workshop

Created by Dionne Henricks, last modified by Neill Reid on Feb 03, 2020



Wednesday, September 25, 2019

Location:

Space Telescope Science Institute, Baltimore, Maryland

Description:

STScI has implemented a dual-anonymous proposal system for the HST proposal review process. Under this system, not only are proposers unaware of the identity of the members on the review panel, but the reviewers do not have explicit knowledge of the proposal teams. The goal is to mitigate unconscious bias and allow reviewers to concentrate on the science rather than the scientist. Following the successful implementation with HST, the Director of the NASA Astrophysics Division, Paul Hertz, has directed all Astrophysics GO Programs to implement this process. This workshop will share STScI's experiences in developing and implementing the dual-anonymous review process.

This is a one-day workshop. The morning will be devoted to talks and panel discussions covering the development and implementation of the HST process, including feedback from representatives of the community and the Telescope Allocation Committee panels. The afternoon will provide an opportunity for representatives from other observatories and missions to explore how the process can be adapted for their use.

For further details, please contact Alessandra Aloisi (aloisi@stsci.edu).

Logistics:

Agenda including links to presentations

Participants → NASA (astrophysics GO), AURA (Gemini, NOAO, DKIST, LSST), NRAO, ALMA, ESA, ESO

Registration Information

Workshop webpage:

https://outerspace.stsci.edu/display/DRW/Dual-Anonymous+Review+Workshop

Summary in STScI Newsletter:

https://www.stsci.edu/contents/newsletters/2020-volume-37-issue-01/dual-anonymous-proposal-workshop

PHYSICS TODAY

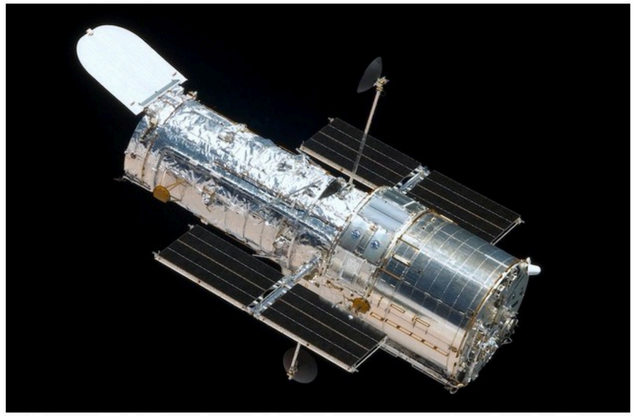
DOI:10.1063/PT.6.3.20190301a

1 Mar 2019 in Commentary & Reviews

Doling out *Hubble* time with dual-anonymous evaluation

A first-of-its-kind peer-review process for allocating time on NASA's workhorse space telescope has the potential to level the playing field for women and other marginalized groups in science.

Lou Strolger
Priyamvada Natarajan



Although it's been nearly 30 years since the launch of the *Hubble Space Telescope*, it's still very difficult for researchers to receive observing time. Credit: NASA

Observing time on the *Hubble Space Telescope* is a scarce resource. Each year only one-fifth of the more than 1000 submitted proposals survive the rigorous system of peer review conducted by the Space Telescope Science Institute (STScI) in Baltimore. With such a competitive process, it's essential to minimize bias based on factors such as gender, race, career stage, institutional size, and geographic origin.

MOST READ

Machine learning meets quantum physics

Commentary: Unity of physics perhaps not as grand as once thought

Ernest Lawrence's brilliant failure

The Planet Nine hypothesis

When condensed--matter physics became king

nature INDEX

Following NASA's lead, researchers are targeting gender bias in instrument time

The switch to double-blind peer reviews could help to ensure that female and early-career researchers get a fair shot at using in-demand equipment.

2 February 2021

Clare Watson

Sven Creutzmann/Mambo Photo/Getty Images

Astronomer Johan Fynbo (L), from Denmark, and Vanessa Doublier, from France, in the control centre of the Very Large Telescope (VLT) in Chile. The European Southern Observatory, which runs operations of the VLT, has switched to double-blind peer reviews for equipment time applications.

In 2019, NASA **switched to double-blind peer reviews** of applications for Hubble Space Telescope time after **research showed** that anonymising proposals practically equalized the chances of success between men and women. It marked a massive shift in review processes after years of **apparent systemic gender bias in allocating scientific resources**.

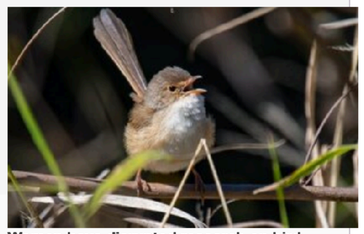
Now the initiative, first introduced by the Space Telescope Science Institute (STScI) in Baltimore, Maryland, which runs operations for the Hubble and James Webb Space Telescope for NASA, is catching on at astronomy observatories, particle physics centres, and computing facilities in Europe, the United States, and Australia.

Researchers applying to use certain telescopes, synchrotrons, and supercomputers are now

Related articles

Female researchers in Australia less likely to win major medical grants than males
30 October 2019
Bianca Nogrady

Women rival men in scientific research publications and citations
17 March 2020
Jon Brock



Women have disrupted research on bird song, and show how diversity can improve all fields of science
16 September 2020
Kevin Omland, Evangeline Rose & Karan Odom

Astro2020 Activities, Projects, or State of the Profession Consideration

White Paper

Adopting Dual-Anonymous Practices in the Reviews for Resource Allocation in Astronomy

Type of Activity: ☐ Ground Based Project ☐ Space Based Project
☐ Infrastructure Activity ☐ Technological Development Activity
☑ State of the Profession Consideration ☑ Other: *Recommendations to agencies/organizations*

Principal Author:
Name: Louis-Gregory Strolger
Institution: Space Telescope Science Institute
Email: strolger@stsci.edu
Phone: (410) 388-1348
Co-authors: I. Neill Reid (STScI)

https://ui.adsabs.harvard.edu/abs/2019BAAS...51g.272S/abstract

Abstract: In an effort to reduce biases, Space Telescope Science Institute (STScI) adopted a dual-anonymous review in its time allocation process in which the identities of both the proposers and the reviewers are withheld until after the science program is selected. The results of the change (after two cycles of implementation) have so far been encouraging, with the success rates of female PIs no longer consistently lagging below the success rates of their male counterparts. Coupled with other examples like the large increase in the fraction awards to new investigators in Cycle 27, and the increasing trend in the fraction of submissions led by women, the adoption of the dual-anonymous review is poised to increase the participation of women and other underrepresented groups in the use of the *Hubble Space Telescope*, improving access to this highly impactful research resource.

Gender-related systematics in reviews at other observatories

Gender-Related Systematics in the NRAO and ALMA Proposal Review Processes

Carol J. Lonsdale, Frederic R. Schwab and Gareth Hunt

National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903

Abstract

A study has been made of the evidence for gender-related systematics in the proposal review processes for the four facilities operated by NRAO: the Jansky Very Large Array (JVLA; hereafter VLA), the Very Long Baseline Array (VLBA), the Green Bank Telescope (GBT) and the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile which is operated by NRAO/AUI in partnership with the European Southern Observatory (ESO) and the National Astronomical Observatories of Japan (NAOJ), in cooperation with the Republic of Chile. A significant gender-related effect is found in the proposal rankings in favor of men over women in the ALMA Proposal Review Processes (PRP) for ALMA Cycles 2-4, with reliability of 99.998% that the underlying rank distributions for male and female PIs are not the same. The effect is largest and most significant for ALMA Cycle 3. A similar overall result is found for the other three NRAO telescopes over proposal Semesters 2012A-2017A, but with lower reliability level overall (98.3%), and with some reversals across semesters in the trend for better performance in the rankings for male PIs. The results align with similar studies recently completed for the HST (Reid 2014) and the ESO proposal review processes (Patat 2016). No correlations are found between the gender-related proposal ranking trends and the gender fractions on review panels. The HST and ESO proposal reviews have come to different conclusions from each other on the role of seniority on the gender-related proposal outcomes at those observatories. The currently available data for the ALMA and NRAO user base do not allow us to investigate the important question of the dependence on the gender-related trends of the seniority of the principal investigators.

Lonsdale et al. 2016, arXiv:161104795

Draft version October 5, 2016
Preprint typeset using LaTeX style AASTeX6 v. 1.0

GENDER SYSTEMATICS IN TELESCOPE TIME ALLOCATION AT ESO

Ferdinando Patat[1]
European Southern Observatory
K. Schwarzschildstr. 2
D-85748 Garching b. München, Germany

[1]fpatat@eso.org

ABSTRACT

The results of a comprehensive statistical analysis of gender systematics in the time allocation process at European Southern Observatory (ESO) are presented. The sample on which the study is based includes more than 13000 Normal and Short proposals, submitted by about 3000 principal investigators (PI) over eight years. The genders of PIs, and of the panel members of the Observing Programmes Committee (OPC), were used, together with their career level, to analyse the grade distributions and the proposal success rates. Proposals submitted by female PIs show a significantly lower probability of being allocated time. The proposal success rates (defined as number of top ranked runs over requested runs) are 16.0±0.6% and 22.0±0.4% for females and males, respectively. To a significant extent the disparity is related to different input distributions in terms of career level. The seniority of male PIs is significantly higher than that of female PIs, with only 34% of the female PIs being professionally employed astronomers (compared to 53% for male PIs). A small, but statistically significant, gender-dependent behaviour is measured for the OPC referees: both genders show the same systematics, but they are larger for males than females. The PI female/male fraction is very close to 30/70; although far from parity, the fraction is higher than that observed, for instance, among IAU membership.

Keywords: sociology of astronomy – history and philosophy of astronomy

Patat 2016, Messenger, 165, 2

Working Group on Anonymizing Proposal Reviews

https://outerspace.stsci.edu/display/APRWG

Anonymous-Double Blind Review Process

Annotated Bibliography/Excerpts

1. **Budden, A. E. *et al.* Double-blind review favours increased representation of female authors. *Trends Ecol. Evol. (Amst.)* 23, 4–6 (2008). doi:10.1016/j.tree.2007.07.008**

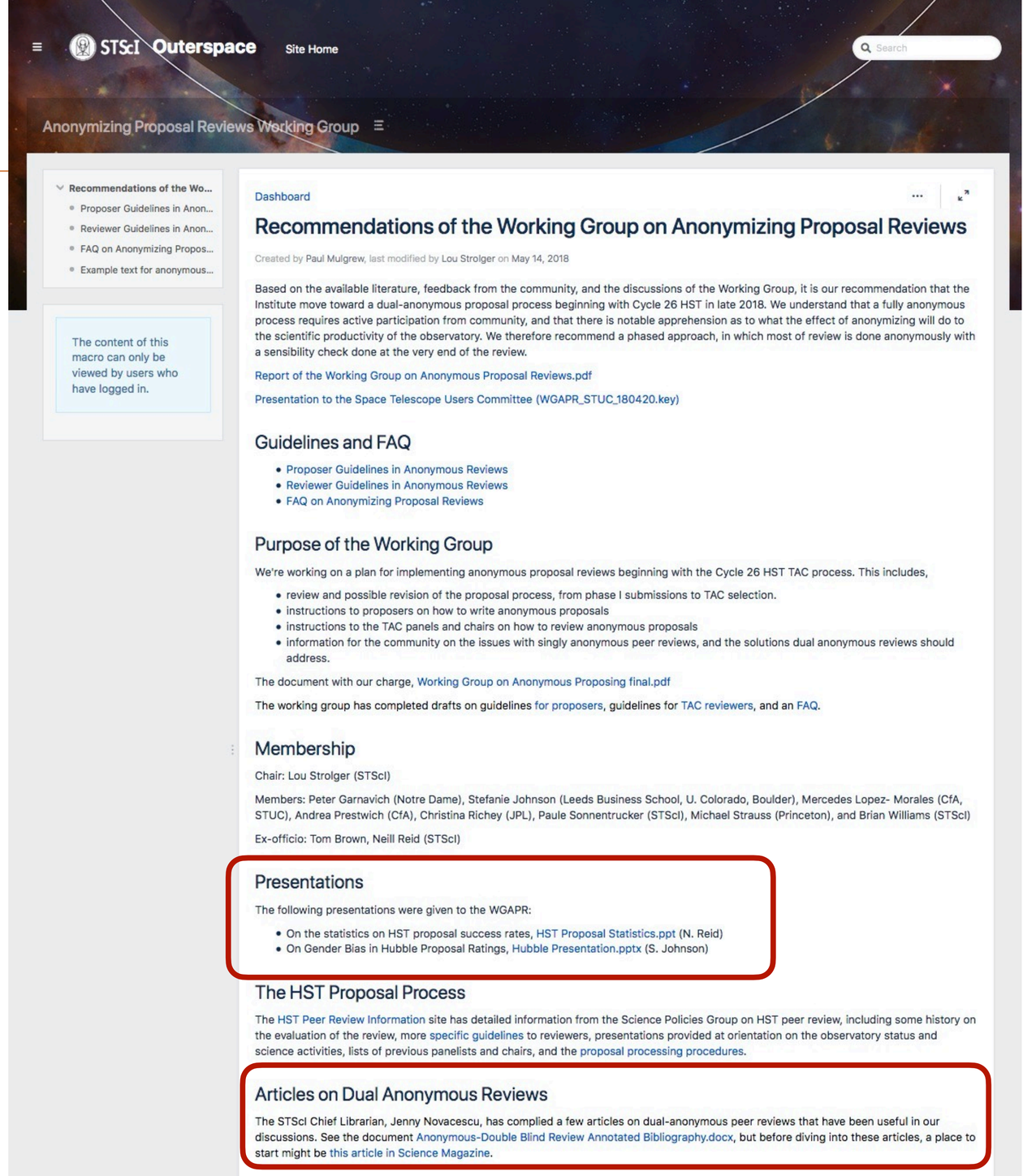
Abstract: Double-blind peer review, in which neither author nor reviewer identity are revealed, is rarely practised in ecology or evolution journals. However, in 2001, double-blind review was introduced by the journal Behavioral Ecology. Following this policy change, there was a significant increase in female first-authored papers, a pattern not observed in a very similar journal that provides reviewers with author information. No negative effects could be identified, suggesting that double-blind review should be considered by other journals.

Double-blind review is frequently criticized on the grounds that it involves an increased administrative load and that authors can be readily identified. However, the more compelling issue is whether double-blind review makes a difference. In light of our study, and evidence that the ecology and evolutionary biology community support double-blind review [12], now might be the time for journals to revisit this issue.

2. **Darling, E. S. Use of double-blind peer review to increase author diversity. *Conserv. Biol.* 29, 297–299 (2015). doi:10.1111/cobi.12333**

Recommended: A brief research letter that discusses merits and limitations of double-blind review in increasing author diversity.

3. **Guglielmi, Giorgia. Gender bias tilts success of grant applications: But it goes away when reviewers focus on the science. Sci. 554, 14-15 (2018). doi: 10.1038/d41586-018-01212-0**



HST oversubscription by cycle