**STScI | SPACE TELESCOPE SCIENCE INSTITUTE**

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

# HST dual anonymous proposal review Implementation & recent statistics

Neill Reid & Louis Strolger, STScI

October 29 2019

# Topics

- Background
  - HST proposal process
  - Statistical results & actions through Cycle 25

- Developing the dual anonymous process

- Implementation & Results from Cycles 26 & 27

  [Based extensively on presentations made at the Dual  Anonymous Workshop at STScI on September 25th 2019]

# Background

# HST Proposal Review Process: overview

- Annual proposal review (most cycles)
  - Smaller proposals are distributed to topical panels
    - Solar System. Exoplanets & disks, stellar physics, stellar populations, galaxies & IGM, black holes & their hosts, cosmology
    - Typically 8 panel members + chair
    - STScI staff provide panel support
    - Larger proposals are reviewed by super-TAC comprised of TAC chair, panels chairs & at-large
  - Two-stage review process
    - Preliminary reviews prior to the meeting
      - 5-6 reviews per proposal ➔ individual grades combined ➔ ranked list
      - Proposals in lower 40% ruled out from discussion (but can be revived)
    - Remaining proposals are discussed and re-graded at face-to-face meeting
      - All un-conflicted panelists grade proposals ➔ ranked list
      - Panels can adjust ranked list to allow for science balance
      - Final ranked list presented as a recommendation to the Director
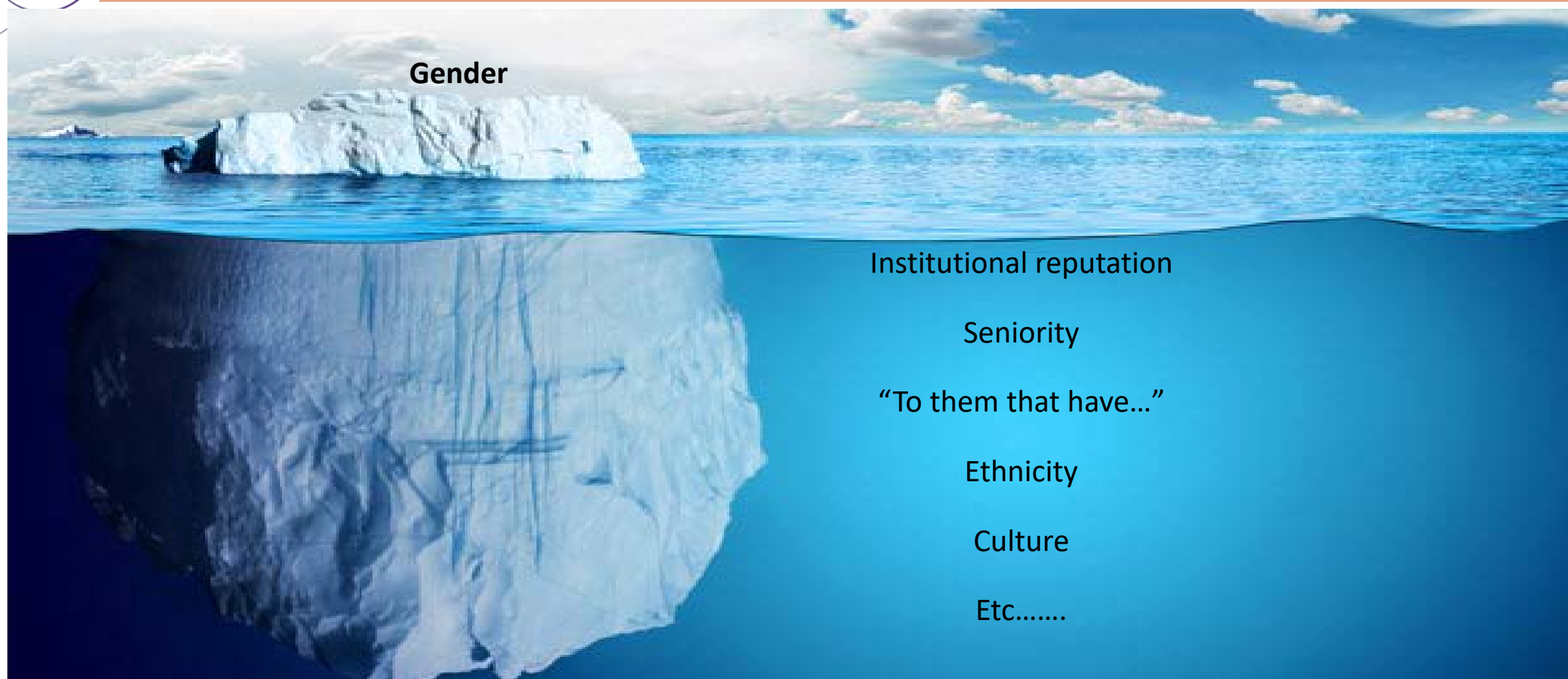
# PI gender and HST proposal selection statistics
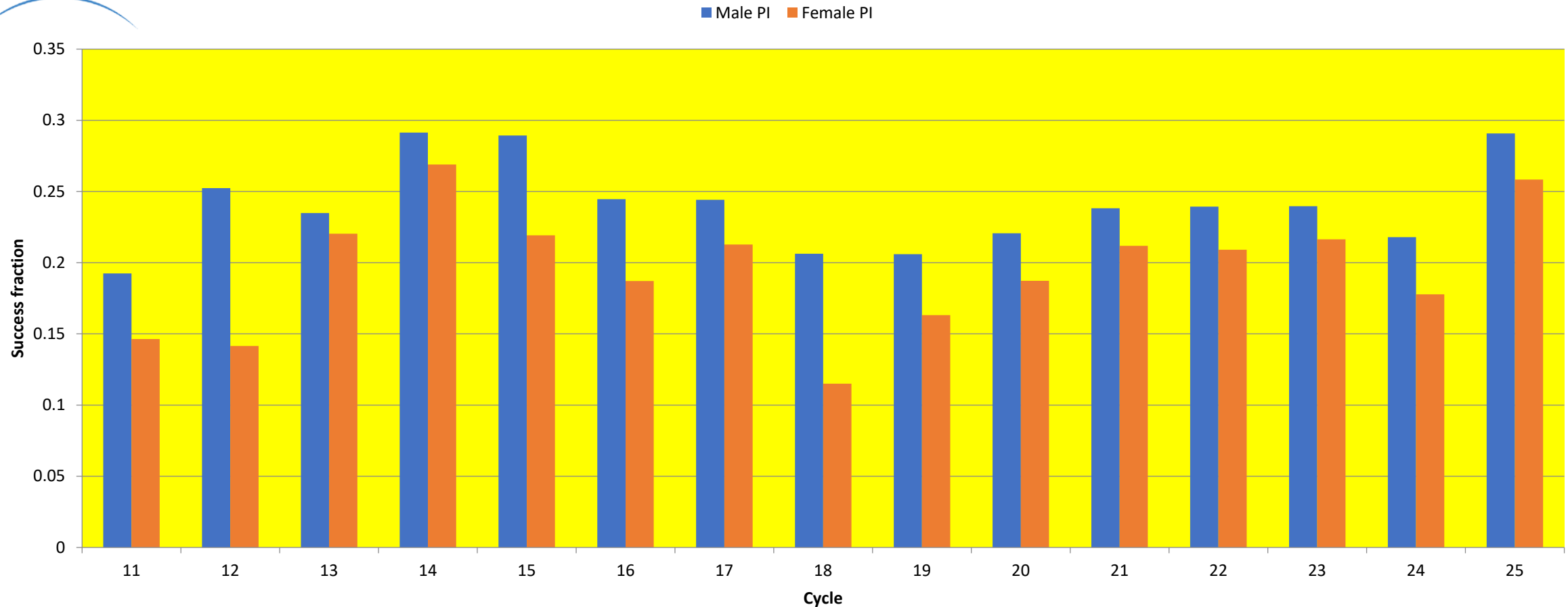


- Clear systematic trend for HST proposals led by male PIs to have a higher success rate
- Comparable analyses since conducted by other facilities & agencies, including NOAO, Chandra, ALMA, & ESA
  - Indications of similar systematics in results from other observatories and missions

# Bias is complex



Gender

Institutional reputation

Seniority

"To them that have…"

Ethnicity

Culture

Etc…….
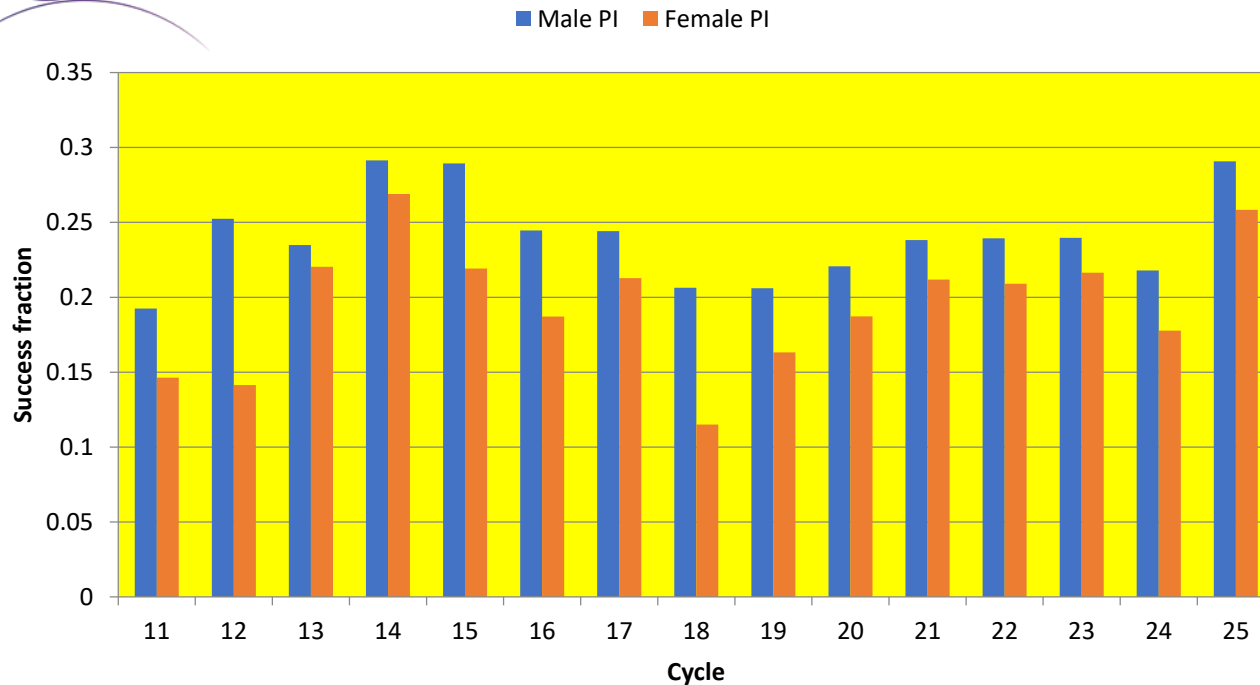
The gender-based offset is likely the tip of the iceberg – a measured effect that points to other inequities and biases that are harder to measure and quantify.

# PI gender and HST proposal selection statistics



HST proposal statistics show that, through 15 cycles, proposals led by male PIs have had a consistently higher success rate than those led by female PIs.

Working in consultation with Space Telescope User Committee & HST Project, proposal format was adjusted to de-emphasise PI information:

- Cycles 22/23: PI name removed from front page of proposal

- Cycle 24: initials replaced forenames

- Cycle 25: alphabetical listing, PI not identified

- Changing the proposal format did not appear to change the outcomes
  - Decided to obtain professional help
  - Decision endorsed by STUC & HST Project

# External advice

- Consultants
  - Prof. Stefanie Johnson, Leeds School of Business, University of Colorado
    - Expert on mitigations for unconscious bias
    - https://www.colorado.edu/business/stefanie-johnson
  - Prof. Jessica Kirk, now at University of Memphis

- Analysed (anonymized) grades from Cycles 21 & 24
  - Preliminary grades from Cycle 21 (PI known) show clear evidence for bias
    - Male reviewers show statistical preference for proposals led by male Pis
  - Preliminary grades from Cycle 24 (PI not known ) show no bias
  - Final outcomes from both cycles

- Johnson & Kirk sat in as observers on the Cycle 25 review (6/2017)
  - They noted that ~60% of the discussion focused on scientists (PI & team members) rather than the proposed science
  - They recommended moving to a fully anonymous review process
  - Their recommendation was endorsed by the STUC and accepted by the STScI Director

Implementing a dual anonymous proposal process

# Working Group on Anonymising Proposal Reviews

**Charge:** The working group is charged with developing an implementation plan for anonymous proposal reviews. The goal is to start implementing the plan with the Cycle 26 HST ΔTAC process. The WGAPR should consider the broad implications of adopting a fully-anonymous proposal review system.

The working group will
- Identify the appropriate process for rendering proposals anonymous, including modifications to the current proposal format
- Review and, as necessary, modify the current proposal evaluation criteria and grading system
- Engage in dialogue with the community to solicit input and identify, and mitigate, concerns
- Provide instructions and guidelines to the community for writing proposals
- Provide instructions and guidelines to the reviewers for reviewing proposals

**Membership:**
Louis Strolger (Chair – STScI), Peter Garnavich (Notre Dame), Stefanie Johnson (Colorado), Mercedes Lopez-Morales (CfA, STUC), Christina Richey (JPL), Paule Sonnentrucker (STScI, ESA), Michael Strauss (Princeton), Brian Williams (STScI)
Ex officio: Tom Brown & Neill Reid (STScI)

# Actions & recommendation

- WGAPR actions:
  - Polled the community for comment: ~60 responses, equally divided for/neutral/against
    - Primary concern was mitigating the potential for exaggeration & false statements
  - Considered implementation mechanisms, documentation & appropriate safeguards

- WGAPR recommendation:
  - "Based on the available literature, feedback from the community, and the discussions of the Working Group, it is our recommendation that the Institute move toward a dual-anonymous proposal process beginning with Cycle 26 HST in late 2018. We understand that a fully anonymous process requires active participation from community, and that there is notable apprehension as to what the effect of anonymizing will do to the scientific productivity of the observatory. We therefore recommend a phased approach, in which most of review is done anonymously with a sensibility check done at the very end of the review."

- Recommendation endorsed by the STUC and accepted by the STScI Director (4/18)

# Adopted changes to the proposal submission process

- Proposers craft their PDFs (scientific justification and description of observations) to be anonymous.
  - Exclude names and affiliations of the proposing team, including in figures and references to personal websites.
  - Do not claim ownership of past work, e.g., "my successful HST program (GO-######)…" or "Our analysis shown in Strolger et al. 2012…"
  - Rather, cite references in passive third person, e.g., "The HST program GO-##### did…", or "Analysis shown by Strolger et al. 2012…". This includes references to proprietary data and software (personal communication).
  - Cite relevant ancillary observations & datasets, but do not explain access to private facilities (telescopes, laboratories) in the main proposal text
  - **Do** describe the work proposed, e.g., "We propose to do the following…" or "We will measure the effects of…"

- Proposers *can* provide reviewers with all the relevant information
  - The primary intent is to eliminate "the team" as a topic for discussion, not make it impossible to guess who might be on that team

# Additions to the proposal format

- Proposers must submit a Team Expertise and Background exposition with their Phase I submission. This section is separated from the main body of the proposal, not anonymous, and will be used in a final stage of the review after the scientific ranking is completed.
  - The Team Expertise and Background section should identify team members with particular skills crucial to the execution of the proposal and clarify access to any private facilities that support the science program.

- Proposers are no longer required to submit detailed Management Plans for Large, Treasury, or Archival programs at Phase I. These will be required and reviewed in budget proposal process.

- *Instructions for proposal writers and reviewers, including extensive examples, are linked from the Call for Proposals for each cycle.*

# Instructions to TAC members

- Consider proposals **solely** on the scientific merit of what's proposed.

- Do not spend any time attempting to identify the PI or the team. Even if you think you know, discuss the science and not the people.

- In the panel discussions leading up to the scientific ranking, do not make guesses on identities, insinuate the likely identities, or instigate discussion on a possible team's past work.
  - Levelers are present in each panel room to ensure this doesn't happen

- Keep in mind that language can be very important in discussing proposals. Utilize the appropriately neutral pronouns (e.g.,"what they propose", or "the team has evaluated data from a C25 program").

- Levelers are present in every panel in addition to panel support staff

- Their role is to ensure that the panel discussions focus on scientific merit. Unlike the chairs, they are not listening for issues pertaining to the science, rather they are focused on the discussion itself.

- If the discussion veers to comments on the proposing team, their past work, their validity, or their identities, the leveler's job is to refocus that discussion.

- **They have the authority to stop the discussion on a proposal**.

- If, in the deliberation of a given proposal, an investigator's self-revealed identity becomes impossible to ignore, and that identity has a clear impact on the discussion, the proposal should be flagged for potential disqualification. The levelers may bring this to the attention of the panel if they feel this threshold has been crossed.

# Compliance with guidelines

- TAC members are asked to flag proposals that may have violated the dual anonymous rules. Those proposals should be brought to the attention of STScI Science Policy Group members.

- Proposals that have egregiously violated the dual anonymous rules may be disqualified from consideration. Ideally, such cases should be identified prior to the meeting.

- Less serious cases (a stray "we" or "our") should be also be pointed out. Panelists should attempt to ignore these less flagrant errors whenever possible, and keep focused on the scientific merits.

- Cases that are too difficult to ignore, or not sufficiently anonymized, should be commented on in the recommendations to the Director, and may be disqualified.

- Panelists should provide specific feedback in their comments to proposers if a proposal was not sufficiently made anonymous.

# Panelist/proposal conflicts

Conflicts are identified solely on a personal basis

- Personal involvement in a proposal
- Involvement of a close collaborator or competitor on a proposal

- Prior to the TAC, panelists submit names of close collaborators and competitors

- Panelists can also flag proposals as potential conflicts during the TAC discussions

- Institutional conflicts are not considered

- All conflicts require the panelist to leave the room

- The source of the conflict is **not** identified

- Significantly reduces the impact of conflicts and increases panelist participation in discussion

# A final check

- HST time is available to any scientist who presents a highly compelling scientific case. However that time is a highly valued resource that must be used responsibly.

- After the scientific ranking is complete, the panel is given the list of investigators (alphabetized) and the Team Expertise and Background sections for proposals above their nominal orbit-allocation line.

- Panelists should raise specific proposals for discussion. If there are clear, compelling deficiencies in the expertise required to see through the goals of the proposal, panel must decide by consensus to flag the submission for potential disqualification, and provide a detailed justification in their comments to the Director.

- The criteria for sufficient expertise is left to the panels in order to evaluate cases as necessary (e.g., particularly difficult observations, difficult analyses)

- General inexperience with HST data is, in itself, a disqualifier. Nor should the failure to publish past datasets, unless there's an extraordinary issue with the team's publication history.

- **Proposals can only be eliminated in this final review**. It will not be used to re-evaluate or upgrade programs below the nominal allocation line.

Outcomes from the review process

# Implementation

- Dual anonymous proposal review has been implemented for Cycles 26 & 27
  - Cycle 26 TAC review (10/18) – medium & large proposals for Cycle 26
  - Cycle 27 TAC review (6/19) – full range of proposals for Cycle 27
  - Cycle 26 & 27 mid-cycle proposals & Director's Discretionary proposals

- High rate of compliance with the dual anonymous guidelines
  - Vast majority of proposals can be adapted in a straightforward manner
  - Only 5 egregious cases requiring disqualification from ~1800 proposals

- Team Expertise
  - A few (larger-scale) proposals were flagged for inspection, either as complex or requiring specific expertise
  - No proposals were subsequently recommended for elimination

- Documentation & communication
  - STScI has identified some areas where additional documentation would be useful
  - Cycle 28 preparations will include additional training & an opportunity for levelers and chairs to tag up prior to the meeting

- Proposal discussions were characterized as more collegial and efficient
  - Focus was squarely on the science rather than the scientists
  - "There was a noticeable shift in the depth of discussions as well. It was clear that reviewers had read the proposals very diligently, and that without the distraction of names and institutions, there was no recourse but to focus on the proposed science." (P. Natarajan, chair of the Cycle 26 TAC)
  - "Overall, I felt that the double anonymous system as implemented in this cycle by STScI worked quite well and has more beneficial aspects than negative ones. I also heard many more positive comments about the system from TAC members than negative ones. Some of the feedback about the double anonymous system is as follows:
    - 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 proposals.
    - 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."
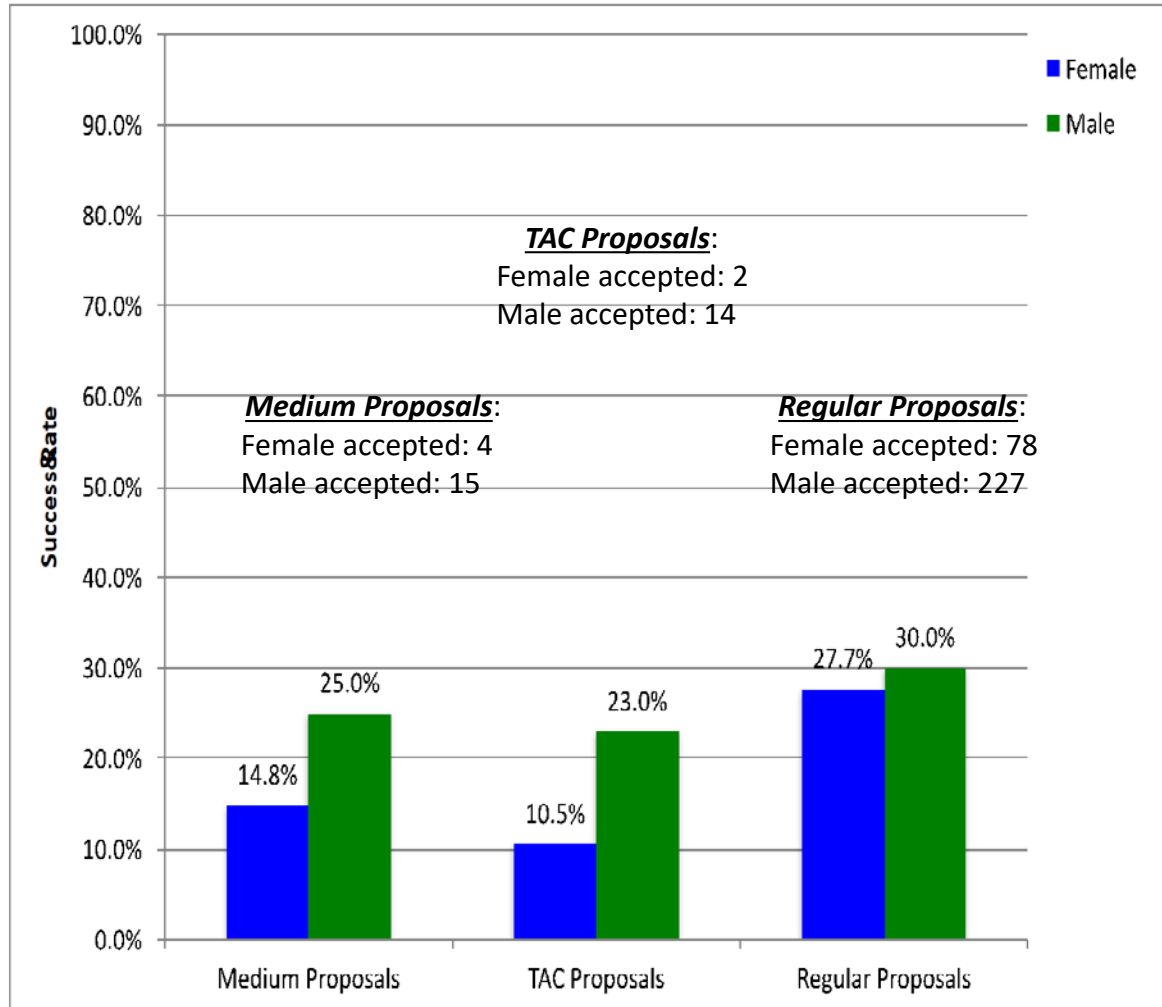      (R. Somerville, chair of the Cycle 27 TAC)
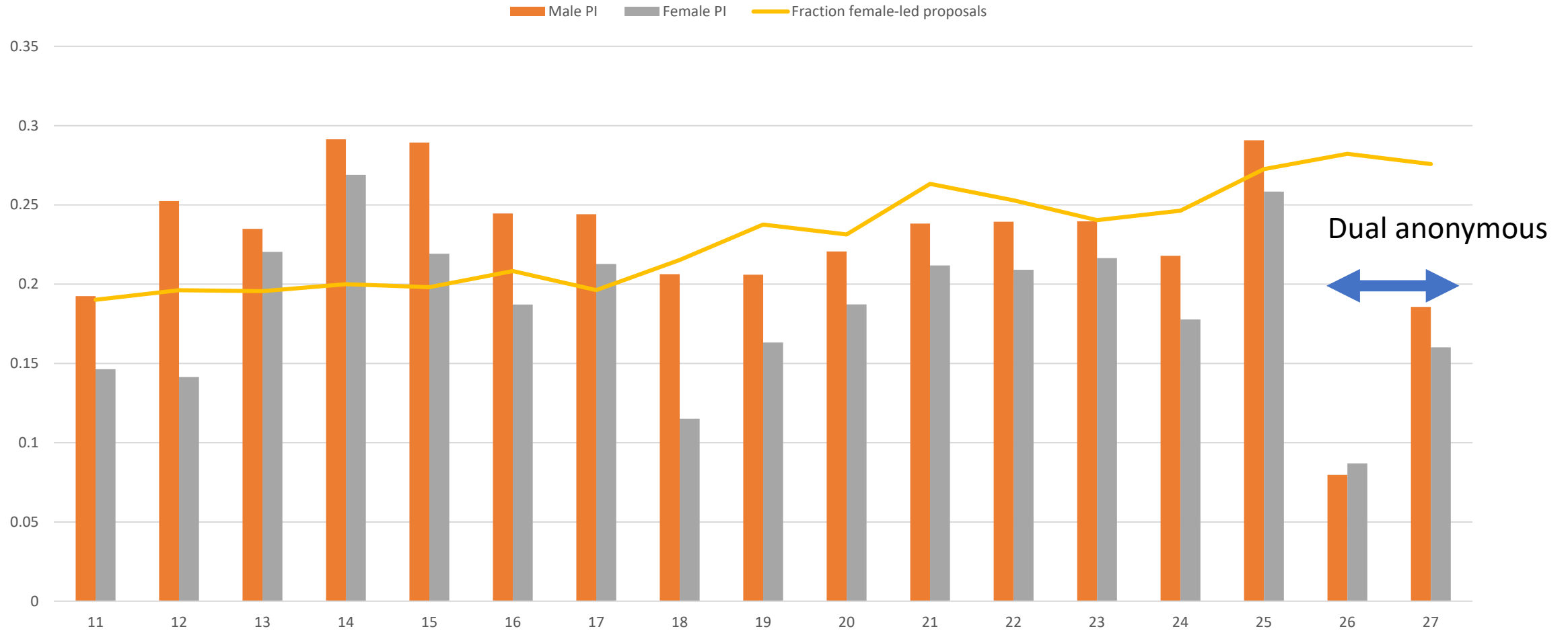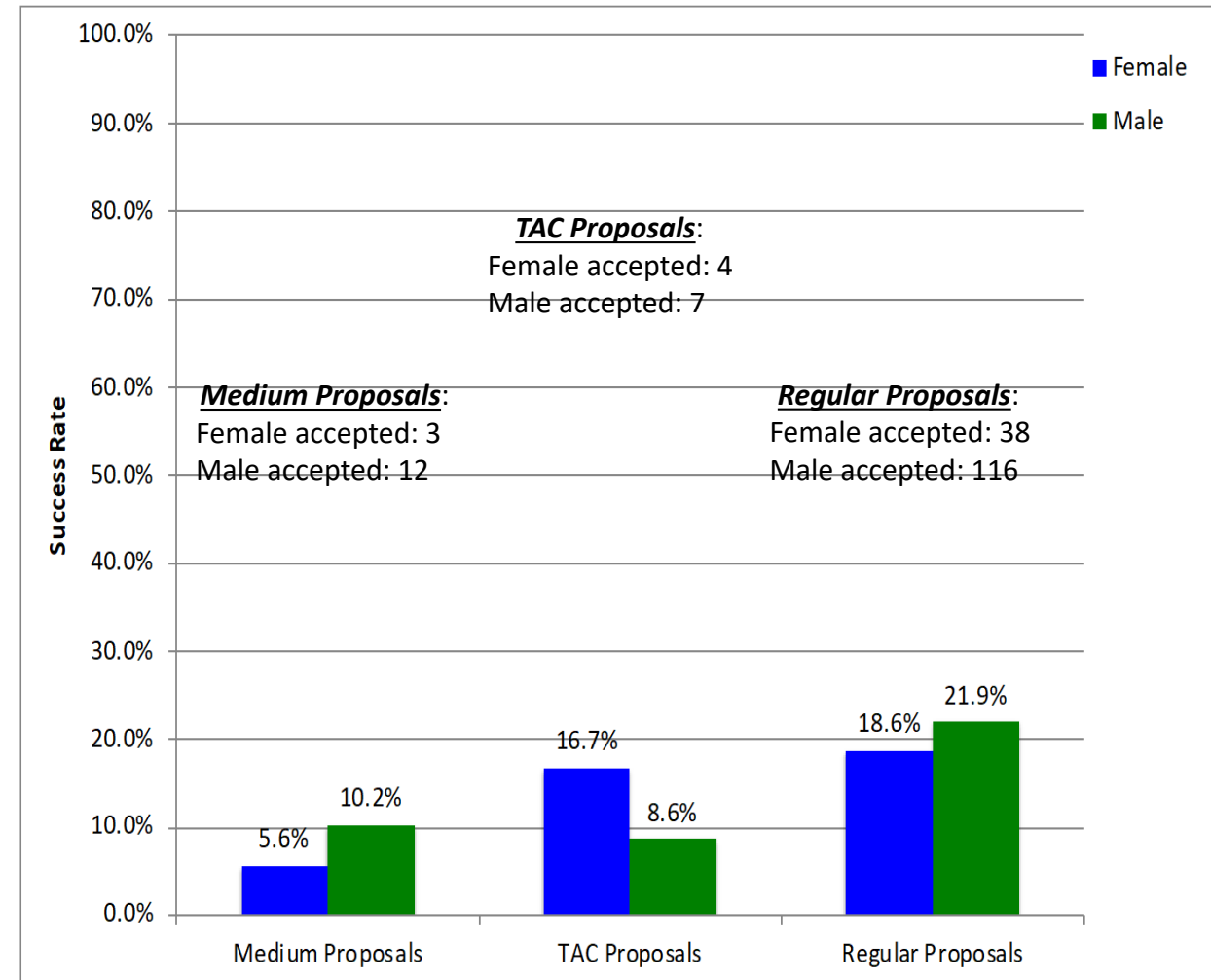
# Success by seniority

| Cycle 23 | F | M |
|----------|-----|-----|
| Ph.d. up to 1999 | 17.6% | 20.8% |
| Ph.d. from 2000 | 23.4% | 26.2% |

| Cycle 24 | F | M |
|----------|-----|-----|
| Ph.d. up to 1999 | 16.2% | 22.3% |
| Ph.d. from 2000 | 18.4% | 21.4% |

| Cycle 25 | F | M |
|----------|-----------|-----------------|
| Ph.d. up to 1999 | 23.3% 17/73 | 30.5% 103/337 |
| Ph.d. from 2000 | 26.6% 68/266 | 28.3% 153/540 |

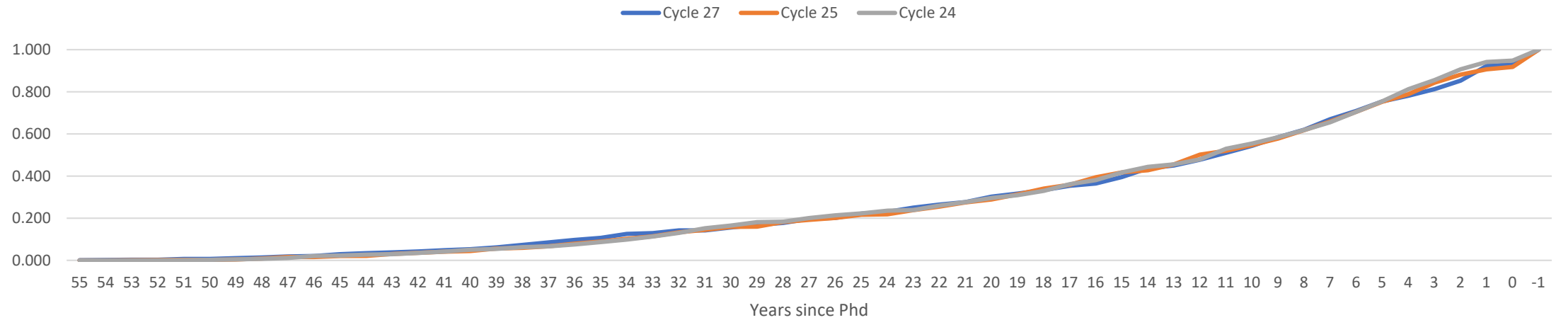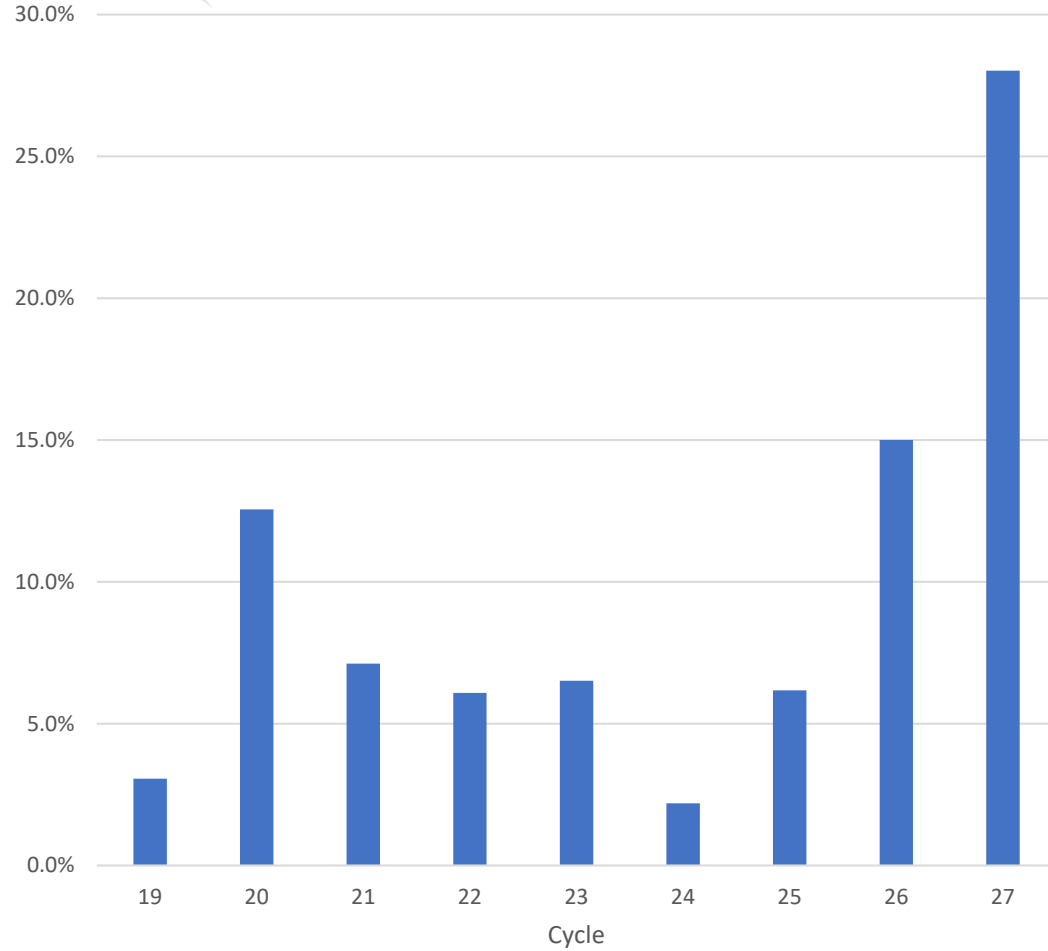| Cycle 27 | F | M |
|----------|-----------|----------------|
| Ph.d. up to 1999 | 6.6% 4/61 | 16.2% 41/250 |
| Ph.d. since 2000 | 18.6% 41/221 | 19.6% 96/489 |
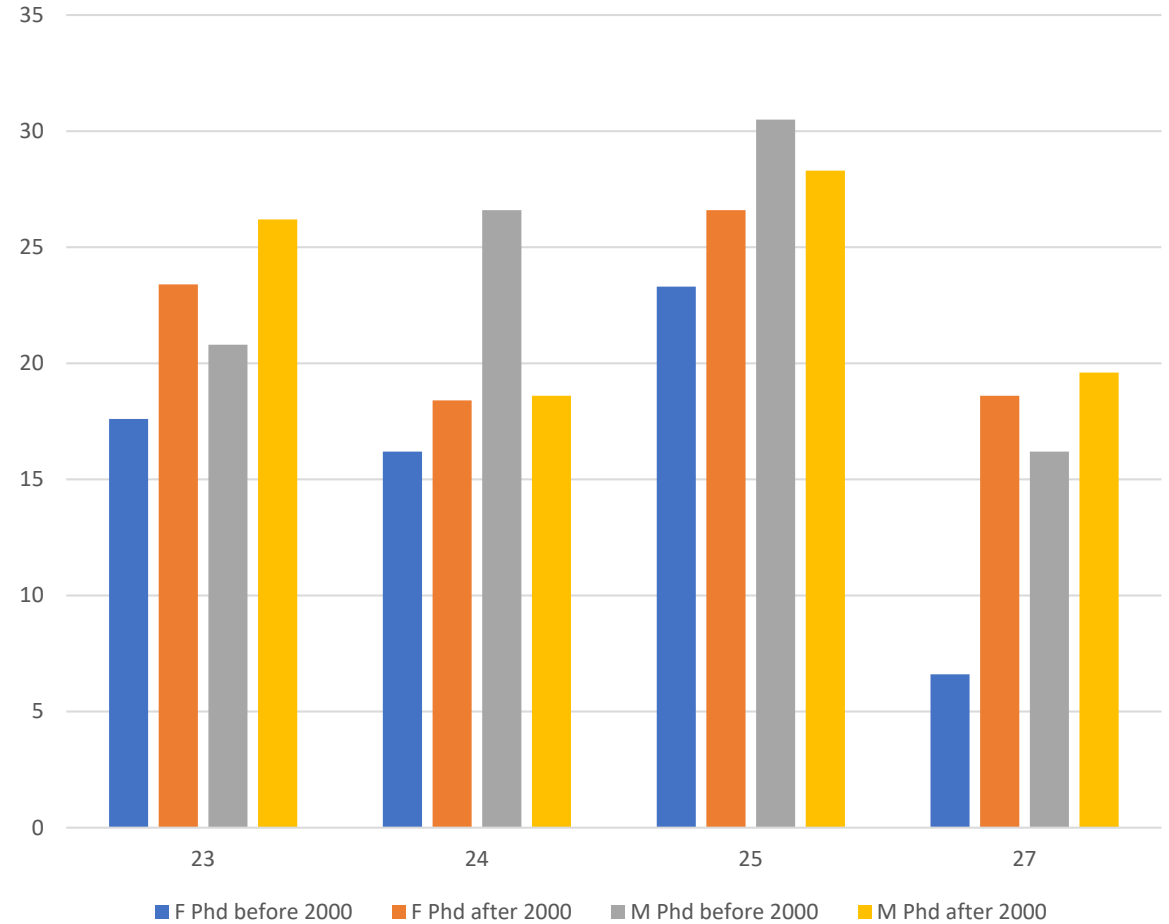
# PI Seniority



PI seniority: submitted proposals

Cycle 27   Cycle 25   Cycle 24

Years since Phd

PI Seniority: accepted proposals

Cycle 27   Cycle 25   Cycle 24

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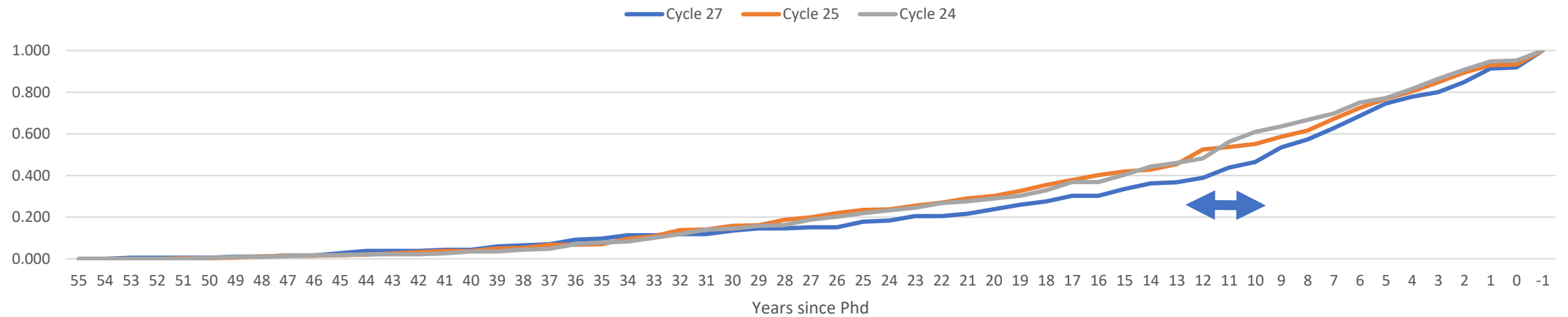
| Cycle | New PIs | Total accepted proposals | Fraction |
|-------|---------|--------------------------|----------|
| 27 | 51 | 182 | 28% |
| 26 | 6 | 40 | 15% |
| 25 | 21 | 340 | 6% |
| 24 | 5 | 228 | 2% |
| 23 | 17 | 261 | 7% |
| 22 | 16 | 263 | 6% |
| 21 | 18 | 253 | 7% |
| 20 | 29 | 231 | 13% |
|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

- Statistics show a systematic trend with PI gender over many cycles
  - We made a number of proposal format adjustments with little impact
  - We moved to a dual anonymous process based on advice from external experts who highlighted the discussion focus on personnel, rather than science

- The process has been implemented through Cycles 26 & 27
  - The community has adapted fairly smoothly to the new requirements
  - Overwhelming majority of proposals are compliant
  - Team expertise section is proving useful in verifying expertise and facility access
  - Reviewers generally find the process more collegial and efficient

- Introducing dual anonymous proposal review is not a magic bullet
  - But the substantial increase in new (to HST) PIs is very interesting
  - We will continue to monitor the results from future TACs