
Systematics in HST Proposal Allocations

I. Neill Reid

Science Mission Office

Space Telescope Science Institute



Purpose



We have analysed statistical data from past HST to investigate the potential for evidence of systematic trends and/or biases

- The results are complex
- We are informing the STUC of the results
- We will also brief the Cycle 21 TAC
- We are not drawing broad conclusions

Motivation



Why are we looking at these statistics?

- There is growing recognition in the community that unconscious biases can play an important role in all decision making processes, even those related to the “hard” sciences
- The STScI Hard Science/Soft Skills has raised local awareness of the potential importance of such issues
- *More specifically, we have received direct questions from several prominent community members on whether there is a difference in the success rate of HST proposals by gender*

Why don't we know this already?

- We don't ask for any gender input as part of the HST proposal submission, so teasing out that information requires a separate study

Unconscious bias - gender



Studies show that unconscious bias plays a role in

- Assessing job applications (Trix & Penka, 2002)
- Evaluating performance (DiTomaso et al, 2007)
- Ranking scientific capabilities (Winners & Wold, 1997; Moss-Racusin et al, 2012)

Biases are generally tied to expectations and pre-supposed stereotypes

Biases can be mitigated if we recognise that they are likely to be present, even in what may seem to be an objective process

References:

DiTomaso, N., Post, C., Smith, D.R., Farris, G.F., Cordero, R., 2007, *Administrative Science Quarterly*, 52, 175: Effects of structural position on allocation and evaluation decisions for scientists and engineers

Moss-Racuson, C.A., Dovidio, J.F., Brescoli, V.L., Graham, M.J., Handelsman, J. 2012, *PNAS*,

Trixis, F., Psenka, C., 2002, *Discourse & Society*, 14, 191: Exploring the color of glass: Letters of recommendation for female and male medical faculty

Wenneras, C., Wold, A., 1997, *Nature*, 387, 341: Nepotism and sexism in peer review

The peer review process



Peer review is frequently used as a selection process:

- Grant funding
- Telescope time
- Refereed publications
- High-level postdoctoral fellowships

Peer review is implemented in two ways:

- Individual reviews, usually written, submitted to a central source
- Panel reviews, conducted (at least partly) in an interactive environment

Peer Review is a SUBJECTIVE process

How effective is peer review, and is it subject to unconscious gender bias?

- Limited number of studies with contradictory results
 - Marsh, Jayasinghe & Bond (2008, American psychologist) –
 - analysed ~2,300 proposals submitted to Australian Research Council in 1996 [Overall success rate = 21%]
 - little evidence for systematic bias except from reviewers nominated by proposers
 - Bornmann, Mutz & Daniel (2008, Journal of Infometrics) –
 - analysed ~6,000 research grant proposals submitted to Swiss National Science Foundation 2004-2006 (~60% success rate)
 - 57.8% success rate for female PIs (645/1117), 65.4% for male PIs, (3248/4965)

Are there discernible systematics in HST proposal results?

The HST TAC process



HST proposals are reviewed by panels:

- Smaller proposals reviewed by specialist panels
 - Set up as mirror panels to minimise major conflicts
 - Proposals assigned based on self-identified keywords
- Large proposals reviewed by TAC supercommittee
 - Panel chairs + At-Large members
- TAC members are drawn from the astronomical community
 - **STScI staff do not serve on the TAC**

HST adopts a 3-phase proposal review

- Preliminary grades submitted by all unconflicted panelists before the meeting
 - Used to establish a triage list (bottom 35-40%)
- In-person panel discussion and re-grading for surviving proposals
- Review of final ranked list (with mirror panels) for topic balance

System has been subject to only minor changes over last 10 cycles

- Accumulated ~9500 proposals in that time period

Analyse for trends with respect to PI gender

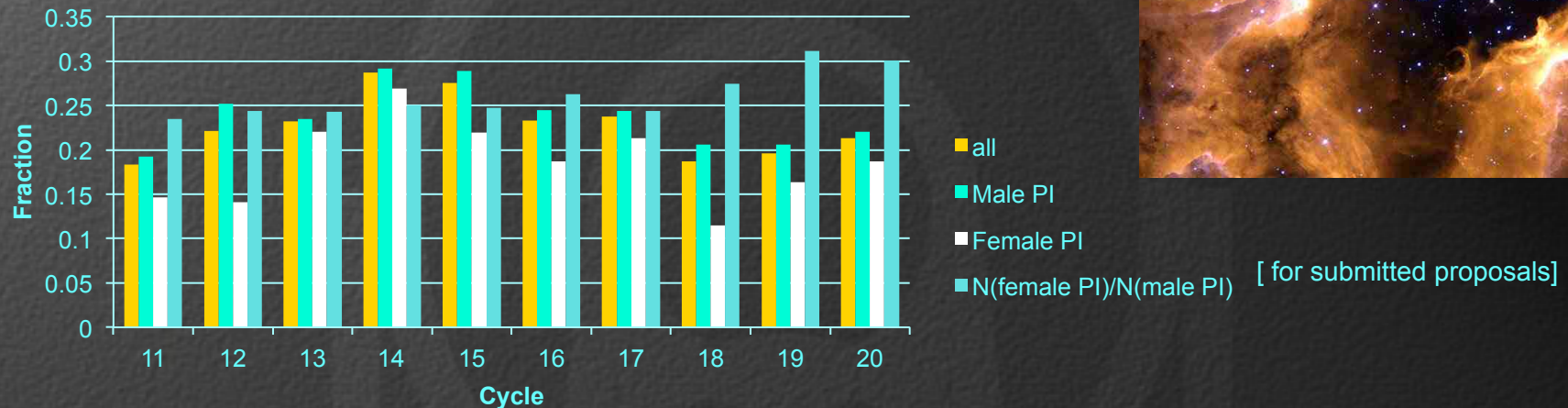
- PI gender identified on a proposal-by-proposal basis using the name + follow-up on the web
 - >99.5% of applicants have an unambiguous digital footprint
- Assumes that PI has most impact on the overall proposal structure & content

The analysis described here focuses on what, not why

STUC Meeting, April 25 2013

Overall success rate by PI gender

HST proposal success rate



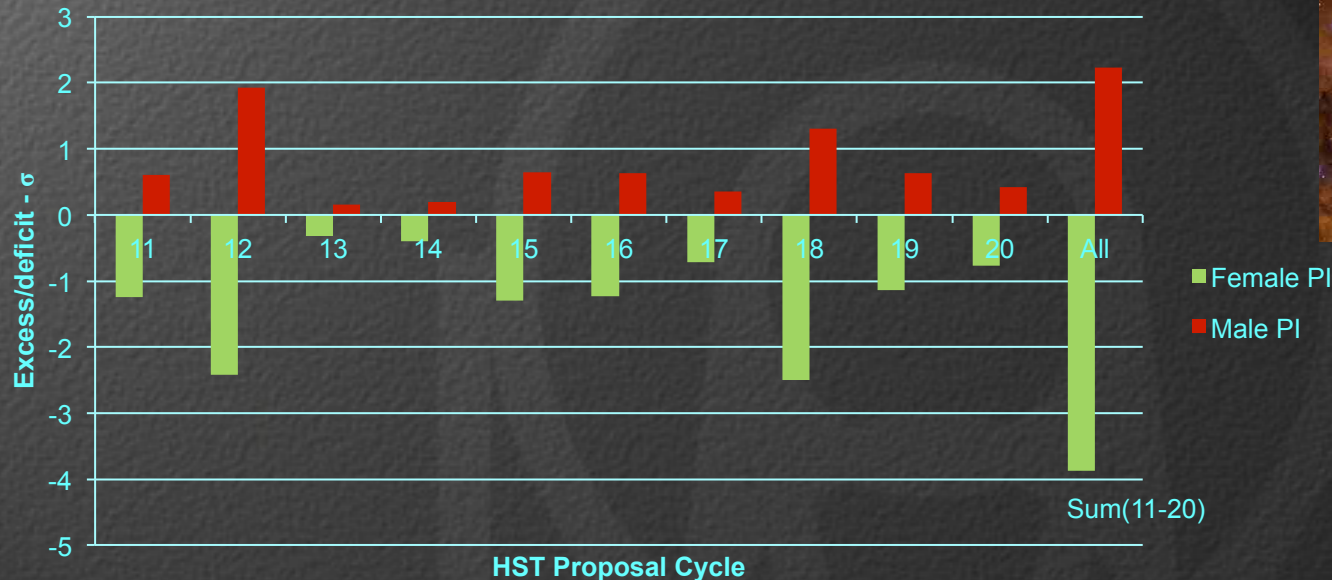
11 - 20	All	Male	Female
Submitted	9411	7457	1954
Accepted	2095	1750	354
Rate	22.26%	23.47%	18.12%

Female PIs account for 20.8% of submitted proposals (1954/9411), but only 16.9% (354/2095) of accepted proposals

Submission demographics – gender:

- Female:Male PI ratio increases from 1:4 in Cycle 11 to 1:3 in Cycle 20
- But the F:M PI ratio for large proposals remains low, 1:5 in Cycle 20

Overall success rate by PI gender



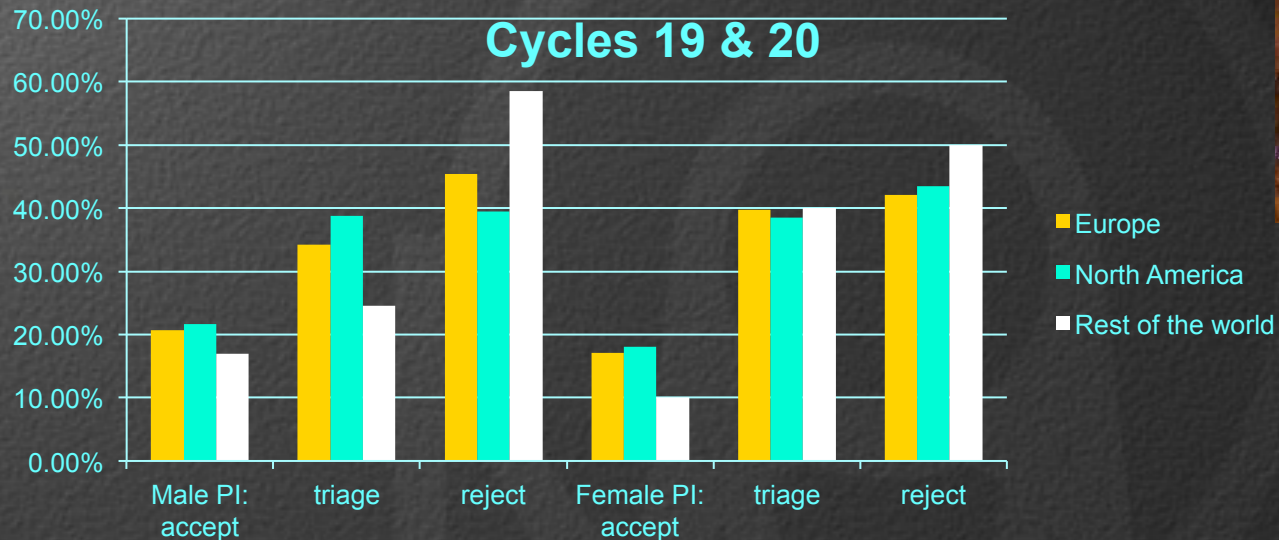
Example: Cycle 18
Av success rate = 18.5%
Female PI = 11.5%
(26/226)
Expected = 41.8
Difference = 15.8
 $\sigma = 15.8 / \sqrt{41.8} = 2.4$

Proposal selection is *not* a Poissonian process – but some (many?) scientists tend to think in those terms.

What do the results look like in this context?

- We calculate the difference between the male/female proposal success rate and the average success rate
- Those differences are normalised by dividing by \sqrt{N}
 - see example for Cycle 18
- Viewed in \sqrt{N} terms, no single cycle reaches as 3σ deviation for either male or female PIs
- Viewed as a multi-year dataset, there is a consistent pattern of results

Statistics by geographical regions



Proposals grouped by region for Cycles 19 & 20:

- North America – USA + Canada - ~80%
- Europe – plus Israel & Russia - ~17%
- Rest of the World - ~3%

Results:

- Rest of the world has the lowest overall success rates
- Europe/North America have similar success

Triage selection shows a smaller male/female offset than approved statistics (37.5% male vs 38.8% female)

Science categories

HST proposals are reviewed by 6 sets of panels:
AGN/QSOs (2), Cosmology(2), Galaxies(3)
Planets (2), Stars (3), Stellar Populations (2)

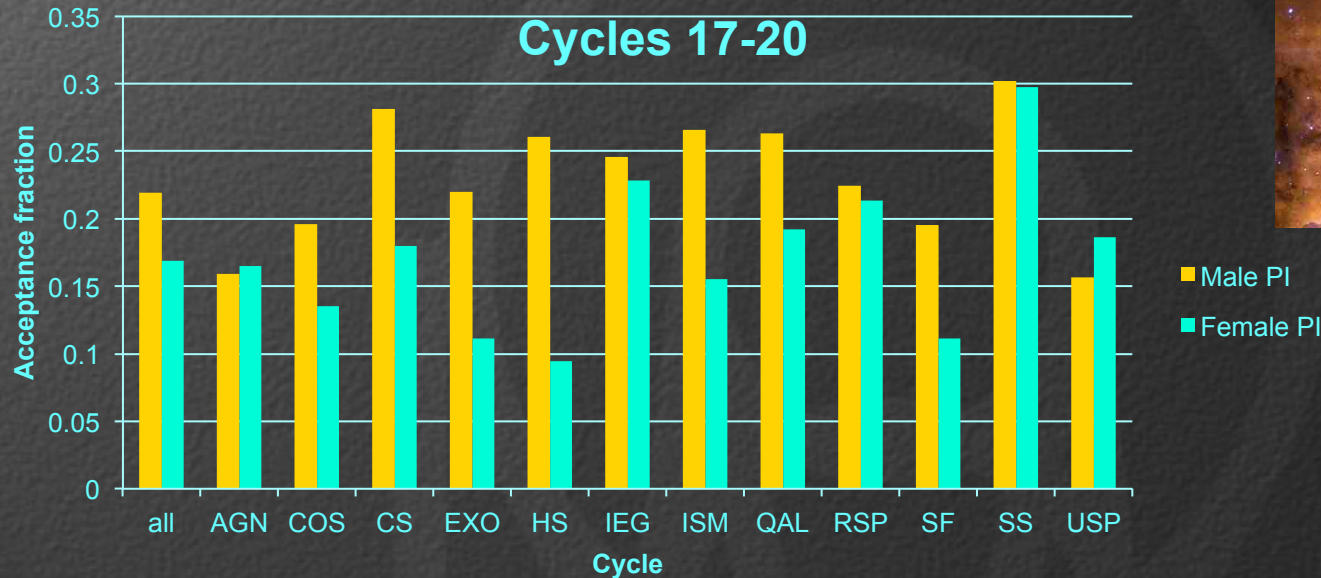
Proposals are self-categorised by applicants:

- AGN – active galactic nuclei (AGN/QSOs)
- COS – cosmology (Cosmology)
- CS – cool stars (Stars)
- EXO – exoplanets (Planets)
- HS – hot stars (Stars)
- IEG – interstellar medium in external galaxies (Galaxies)
- ISM – interstellar medium within the Milky Way or nearby galaxies (Stellar populations)
- QAL – quasar absorption lines (AGN/QSOs)
- RSP – resolved stellar populations in nearby galaxies (Stellar populations)
- SF – star formation (Planets/Stars)
- SS – solar system objects (Planets)
- USP – unresolved stellar populations in distant galaxies (Galaxies)

A handful of proposals in some categories may be directed to different panels
Categories can change, but have been constant since SM4 (Cycles 17-20)



Statistics by science categories



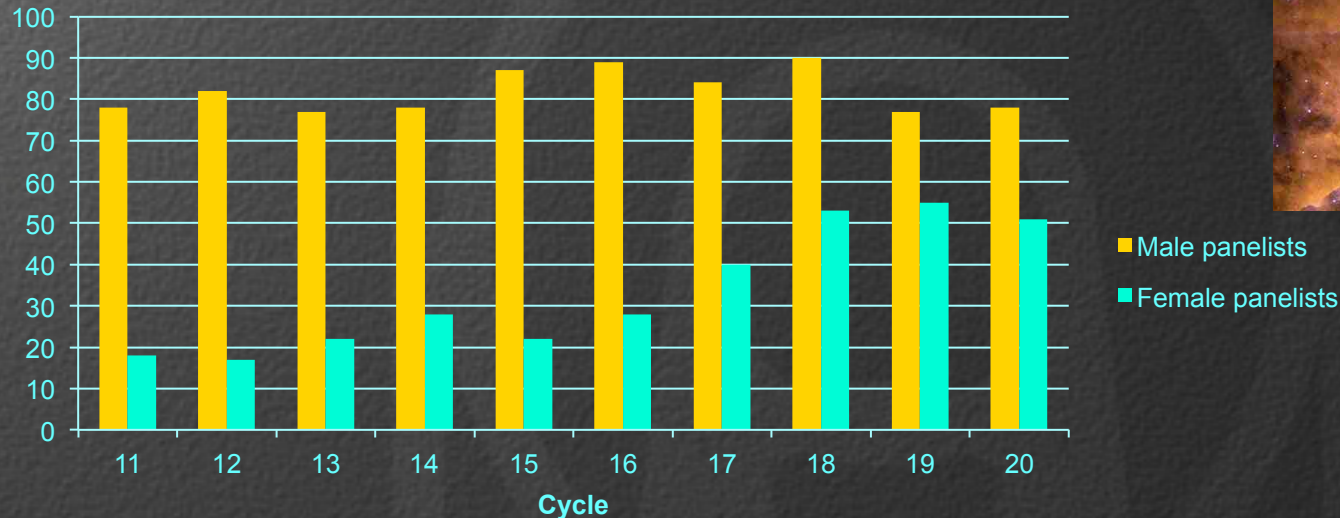
Combined statistics for Cycles 17-20:

- Same panel system, same science categories

Significant differences in gender-based success rates by science topic

- But note that panels cover more than one topic
 - Planets includes both SS and EXO plus some SF
 - Stellar Pops includes both ISM and RSP
 - AGN & QSOs includes AGN & QAL

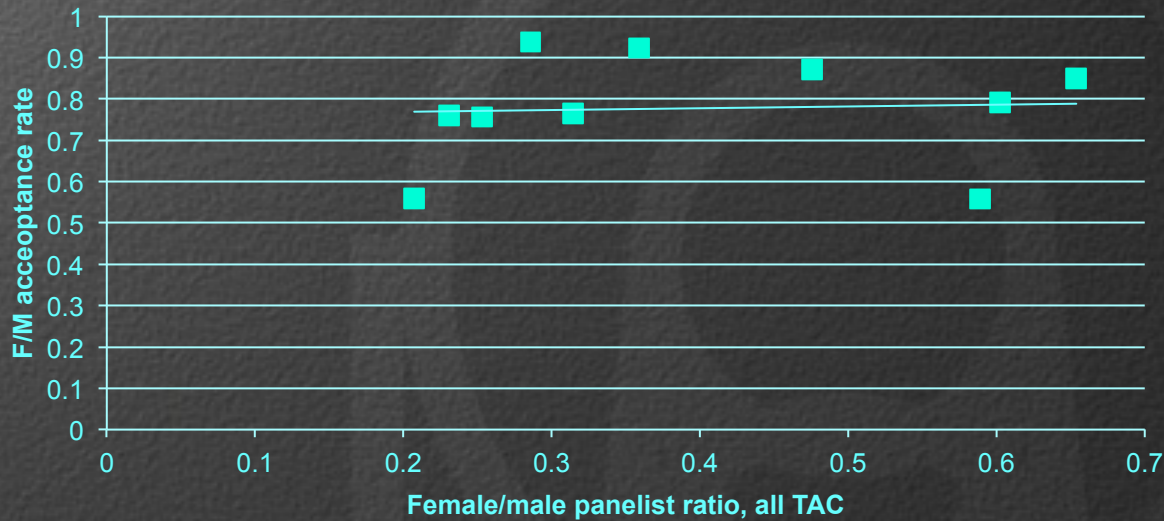
Gender distribution on the TAC



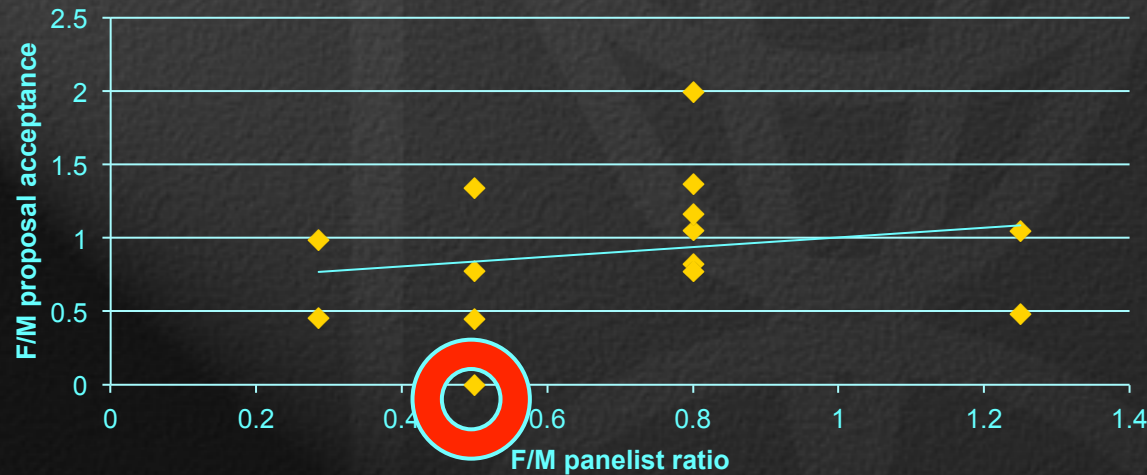
Concerted effort to broaden participation in the TAC in recent years

- Aim for greater diversity, including gender diversity
- Aim for greater participation by active, less senior researchers
- Participation by female researchers has increased from ~19% in Cycle 11 to ~40% in Cycle 20

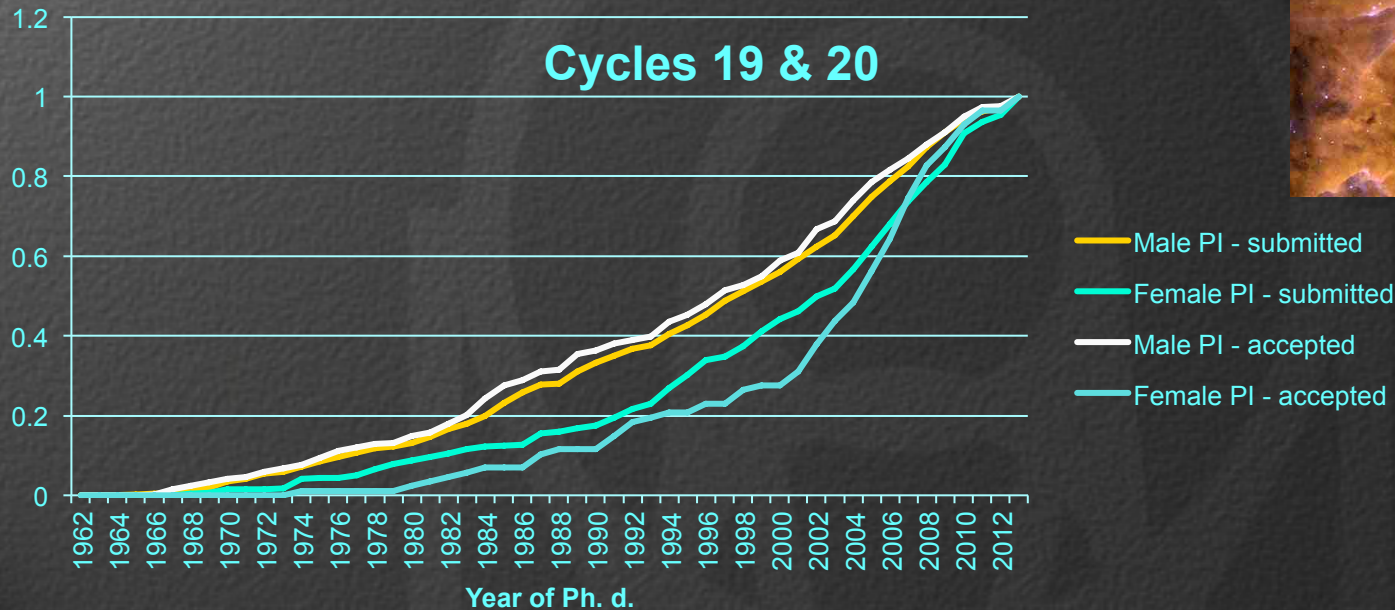
TAC gender distribution and success rates



No indication of a correlation between TAC/panel gender composition and gender-based proposal success rate



Gender demographics - proposals



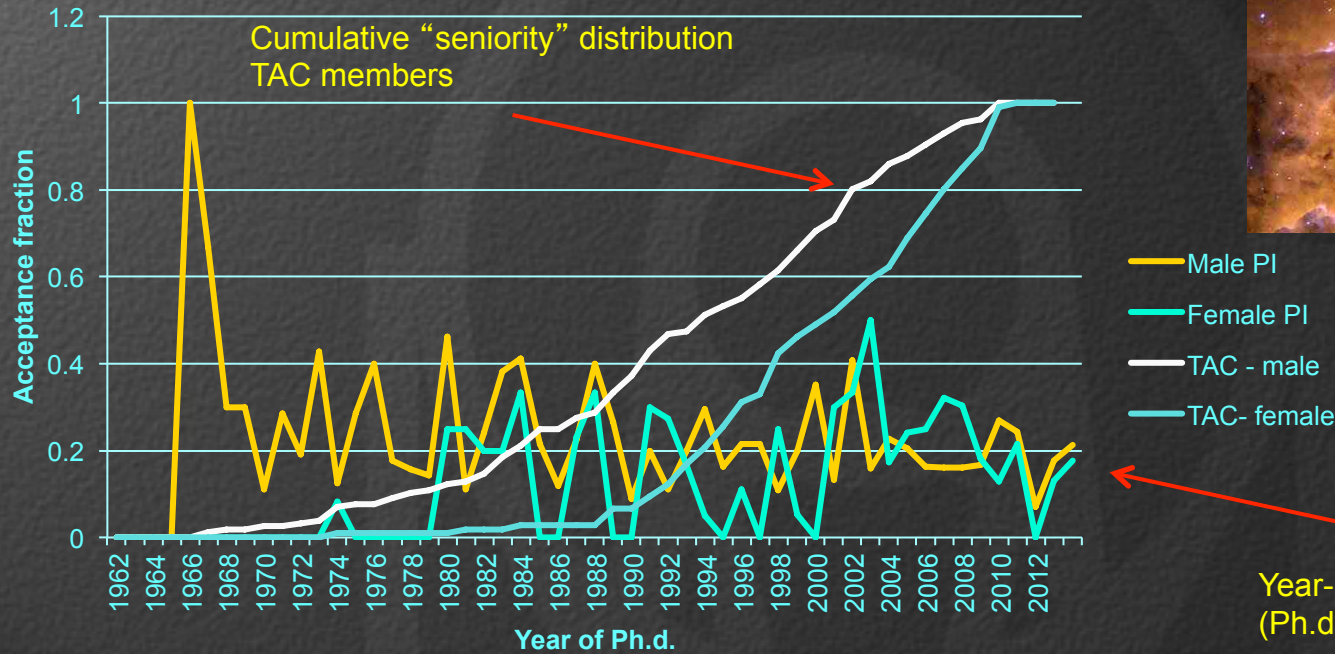
Seniority for Cycle 19 & 20 PIs:

- Year of Ph.d. identified (thanks Jill Lagerstom & STScI library staff)
- Cumulative distributions for male/female submitted/accepted

Results:

- Male PIs – submitted & accepted distributions identical
- Female PIs – significantly higher proportion of accepted proposals from recent (post ~2000-2002) Ph.d.

Gender demographics – success rates & TAC



Cycles 19 & 20

Year-by-year acceptance rate (Ph.d. of PI)

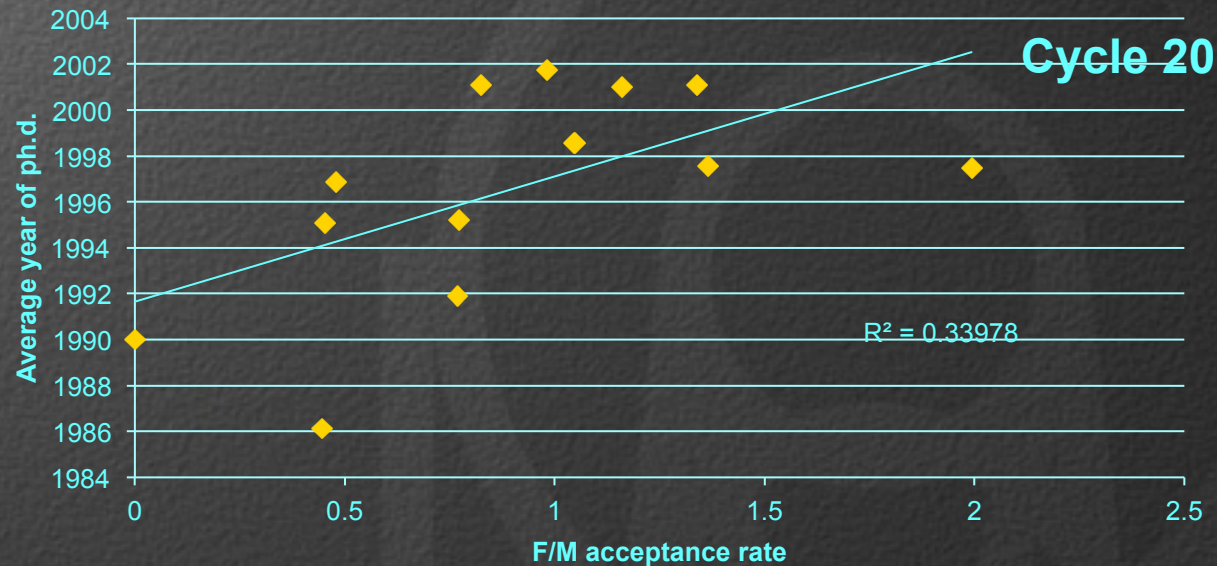
Year-by-year success rates for male & female PIs:

- Approximately constant success rate by seniority for male PIs
- Significant upturn in the success rate for female PIs post-2000;
 - Female PIs who graduated since 2000 are slightly more successful than male contemporaries: $63/276 = 22.9\%$ vs $141/705 = 20\%$

TAC membership:

- 50% of female TAC members, 25% of male TAC members Ph.d. post-2000

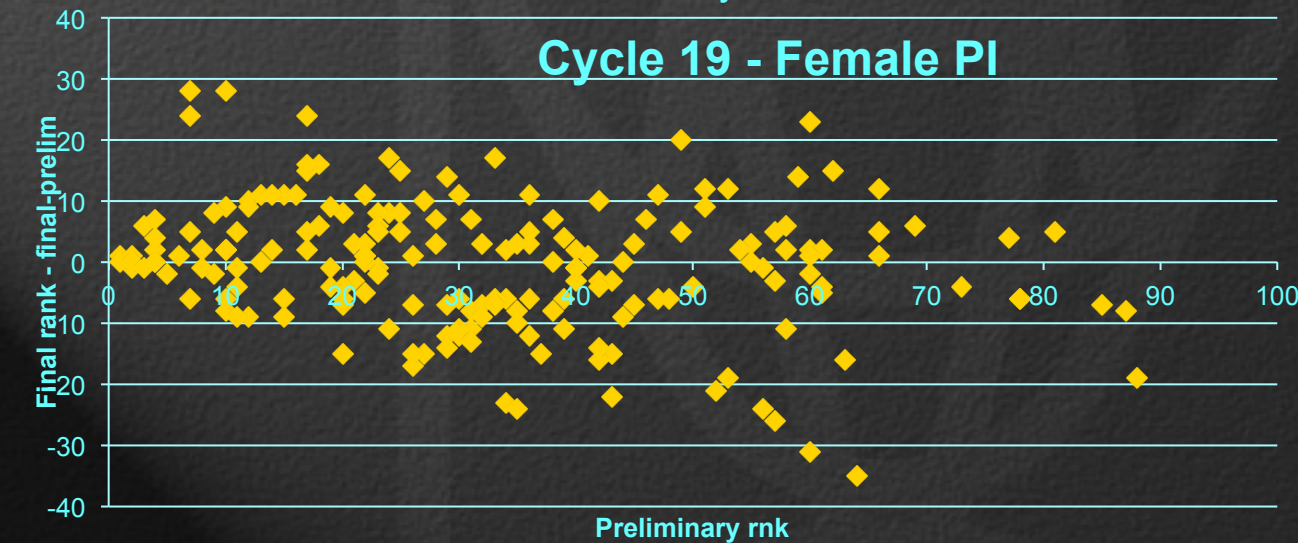
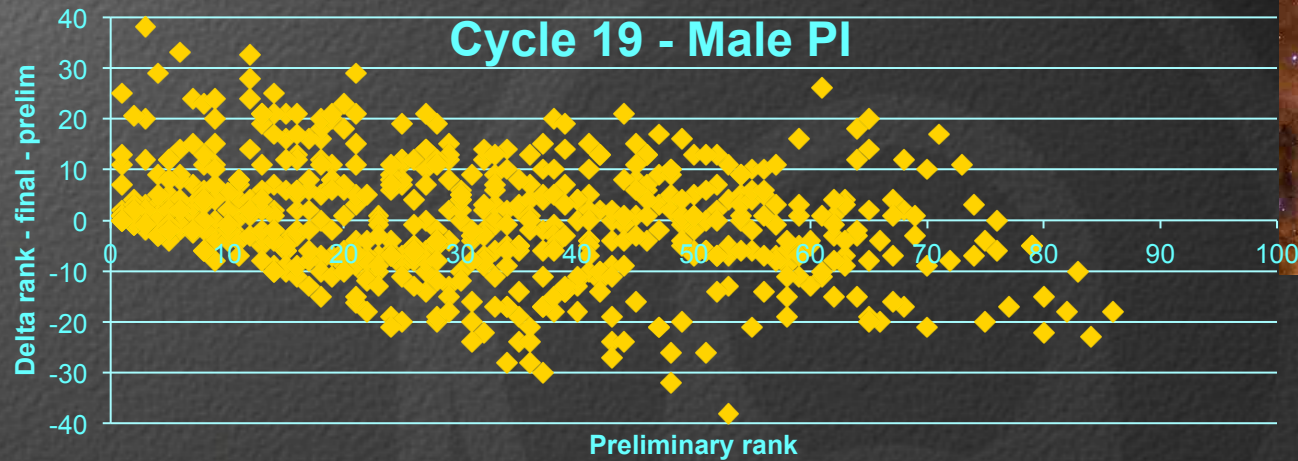
Panel demographics & success rates



Does the acceptance rate of male & female PI' d proposals vary with panel seniority?

- Apparent trend towards higher female/male acceptance ratios in panels with more recent average Ph.d. year

The discussion phase



-ve delta = improved rank by panel/TAC

Are female PIs proposals ranked down more than male PIs in the discussion

- Not clear there's a significant difference

STUC Meeting, April 25 2013

Joe Friday Summary: Just the facts



1. Proposals with female PIs have consistently had a lower overall success rate than male PIs over the last 10 HST cycles
2. There is little evidence for significant variation in gender success rates as a function of geographical origin
3. F/M ratios are closer for triage than for proposal success rates
4. F/M success rates show significant variation with science topic *but* the same panels can produce different results for different topics
5. There is little indication of a correlation between gender success rates and the gender distribution on the selection committees
6. There is an indication that more senior panels tend to generate lower F/M success rates
7. There is less evidence for gender differences for proposals where the PI obtained his/her Ph.d. after 2000; indeed, in this range female PIs have a slight advantage in Cycles 19 & 20

Comments



Scientists tend to assign causality where they see correlation:

BUT

Post hoc is not necessarily propter hoc

- Many factors may underlie the statistical results
 - Unconscious (conscious?) bias w.r.t. institution, reputation, gender, seniority, generational camaraderie, self marketing, etc etc

We have a historical dataset

At this juncture, we have no means of identifying causation

- 792 researchers from the world-wide community have participated in HST TACs through Cycles 11- 20
 - **These results are likely to be indicative of broader community trends in the same time-frame**

Summary & Actions



We have been asked about gender success rates on HST:
The results are complex, with some indications of systematic trends related to gender & seniority.

We should not try to fix apparent issues wrt HST (and JWST) proposal selection by setting quotas – particularly when we don't understand the underlying cause(s), and consequently cannot recommend a simple solution.

Cycle 21 TAC members will be briefed on these results
They will be reminded that

- **Peer review is a subjective process.**
 - In reviewing a proposal, panelists consider factors that go beyond the written word, but they should consider whether those other factors might be unduly influencing their decisions and be aware of the potential for unconscious bias.
- **The prime focus is still identifying the “best” science**
 - But in reviewing the final ranking, panels should consider not only the subject balance, but also the gender balance of top-ranked proposals. Panel chairs will be asked to discuss the final disposition as part of the panel report.

Backup



Proposal acceptance fractions, Cycles 11-20



Cycle	All	Accepted	Male PI	Accepted	Female PI	Accepted
11	1078	198	873	168	205	30
12	1045	231	840	212	205	29
13	905	210	728	171	177	39
14	725	208	580	168	145	39
15	737	203	591	171	146	32
16	821	191	650	159	171	32
17	958	228	770	188	188	40
18	1050	196	824	170	226	26
19	1007	199	768	160	239	39
20	1085	231	832	183	253	48
all	9411	2095	7457	1750	1954	354

STUC Meeting, April 25 2013

Statistics by geographical regions



	Male PI: all	Accept	Triage	Reject	Female PI: all	Accept	Triage	Reject	F/M ratio
Cycle 19									
Europe	143	27	49	67	45	6	16	23	0.32
North America	595	129	205	261	190	33	67	90	0.32
Rest	30	3	6	21	4	0	2	2	0.13
Cycle 20									
Europe	162	37	55	70	43	9	19	15	0.26
North America	635	138	272	225	203	38	84	81	0.32
Rest	35	8	10	17	6	1	2	3	0.17

Statistics by scientific topic



	Male PI: all	Accept	Female PI: all	Accept	F/M ratio
Cycles 17-20					
All	3194	701	905	153	0.28
AGN	358	57	103	17	0.28
COS	464	91	111	15	0.24
CS	167	47	50	9	0.30
EXO	191	43	45	5	0.24
HS	369	96	74	7	0.20
IEG	118	29	57	13	0.48
ISM	267	71	58	9	0.22
QAL	167	44	52	10	0.31
RSP	393	88	89	19	0.23
SF	128	25	72	8	0.56
SS	139	42	37	11	0.27
USP	421	66	161	30	0.38