



STScI | SPACE TELESCOPE
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

JWST/HST Proposal Grading System

Neill Reid

JSTUC: 18 March 2024



Why do we have multiple grades?

“Imagine going to your physician for a medical examination. During the exam, the physician records data on your height, weight, blood pressure, and heart rate, and asks you questions about your lifestyle and how you are feeling. After gathering all this information, imagine the physician entering the data into a computer that uses a mathematical algorithm to calculate a single number describing your physical condition. The physician then reports the number to you, suggests how you might improve it, and sends you on your way.

Would you be satisfied with such an examination? Would you have faith in a physician who analyzed information about you in this way? Would you trust a computer algorithm to tally the information and offer an accurate assessment of your health.?...

One single grade is inadequate....

Clear rubrics matter”

from Berkowicz & Myers, Education Weekly, Feb 2018



Why did we select these three parameters for grading?

- Impact within the sub-field
 - The results derived from the proposed scientific program should be widely recognized as having an important impact for colleagues working on similar problems
- Out-of-field impact [“Why should I care?”]
 - JWST is a general purpose observatory
 - The results derived from science programs should be recognized as having relevance beyond the immediate field of study
- Suitability
 - JWST time is heavily over-subscribed
 - Science programs should require JWST observations as opposed to using other facilities
 - But there’s some fuzziness; some science questions might be addressed by other facilities (e.g. NIRCам vs WC3IR) or by alternative techniques (e.g. IR vs. UV)

Currently, the grades are normalized (separately, for each reviewer) and given equal weight

1. Equal weighting may not be the best approach
2. Normalising suitability may not be appropriate

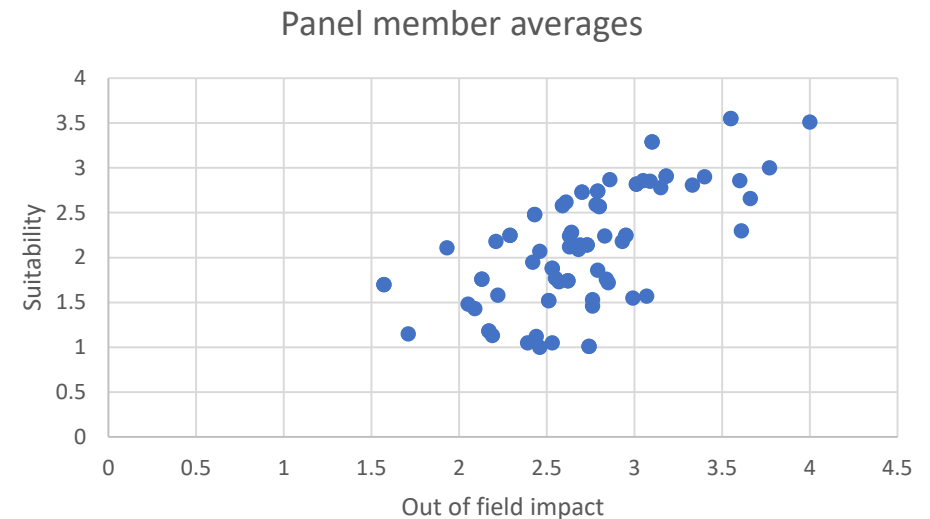
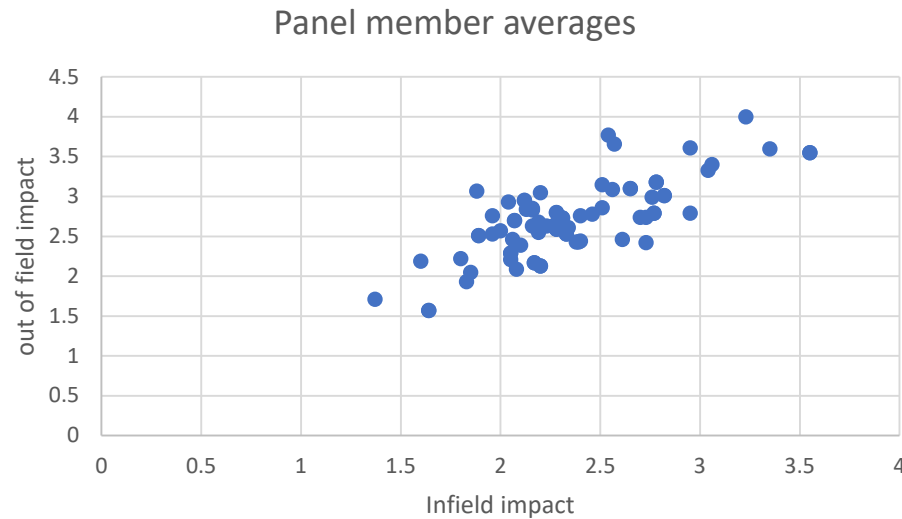


Are the three grades used in different ways?

Compile average preliminary grades (over all proposals reviewed in the panel and not normalized) for 81 JWST Cycle 3 panelists

Averaged grades for individual panelists are correlated, but with some dispersion

This suggest that reviewers are considering different factors when they assess proposals





Individual proposal grading

Match grades for individual proposals – preliminary grades, not normalized

- 30 panelists from 7 panels – Gal, Exo, STPhys, SMBH, STPops, LSS

Grades show correlation, but in most cases with some dispersion

- A few panelists give identical grades for 2 or all 3 parameters

Again, suggests that, in general, reviewers are considering different factors for each grade

