



STScI | SPACE TELESCOPE
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Updates to the TAC grading scheme

Lou Strolger

10.15.2020



Issues with the Grading Schema

Based on feedback from TAC panelists and SPG observations of the process, we note the following shortcomings with the current grading scheme:

- **Grade Reversal:** the scale (5=poor to 1=excellent) runs in the opposite direction than some expect.
- **Restricted Grading:** the full grade range could be more broadly used, i.e., there are few “poor” proposals, Reviewers may unnecessarily restrict themselves to a narrow grading range (adjectival equivalencies).
- **Grade Compression:** the grade range is less useful after the triage round, as all proposals at that point are nominally “good” to “excellent” (adjectival equivalencies).
- **Multiple Grade Categories:** We asked for grading in along 3 (+1) dimensions, *Scientific Merit* (scientifically sound?), *Importance to Astronomy* (urgent?), and *Unique Capabilities* (taking advantage of HST’s unique capabilities?). We included a 4th dimension called Complexity that is utilized by Grants to evaluate a funding model for JWST (in Cycles 2+).

However descriptions were often misunderstood and unclear, especially how they apply to all types of proposals, e.g., how to score and AR Theory proposals in *Unique Capabilities*? Also not made clear how the grades would be used or combined, as a straight average or otherwise.



Proposed Solutions

SPG evaluated and discussed several options for addressing these issues, including changing the numerical scale, improving the adjectival equivalents, etc.

However, given the limited time to integrate changes into the TAC management tool (SPIRIT) in advance of JWST Cycle 1 & HST Cycle 29, we opted to focus on **critical changes for now**, and address more transformative **changes later** after those reviews.



Critical Changes

Grade renormalization: Our thoughts are that renumbering the scale (broadening, e.g., to 1-10, and/or reversing, 10=excellent and 1=poor) would cause further confusion with some of our more experienced reviewers.

Dr. Cecilia Reyes

| Propo sal ID | Impact within sub-field | Broader Impact | Suitability |
|-----------------|-------------------------------|-------------------|-------------|
| 0103 | 1 | 1.5 | 3 |
| 0104 | 2 | 1.8 | 3 |
| 0105 | 2.5 | 2 | 1 |
| 0106 | 1.2 | 1.5 | 2.2 |
| 0107 | 2.1 | 1.2 | 1 |
| 0119 | 4 | 4 | 2.5 |
| 0120 | 2 | 2 | 2 |

We will standardize (or renormalize) grades from each reviewer across their assignments *within the same grading category*. Renormalized scores will have a mean of 3 and a deviation of 1.



Critical Changes

Grade renormalization: Our thoughts are that renumbering the scale (broadening, e.g., to 1-10, and/or reversing, 10=excellent and 1=poor) would cause further confusion with some of our more experienced reviewers.

Dr. Cecilia Reyes Normalized

| Proposal ID | Scientific Merit | Importance to Astronomy | Unique Capabilities | Average Normalized grade |
|-------------|------------------|-------------------------|---------------------|--------------------------|
| 0103 | 1.9 | 2.5 | 4.1 | 2.8 |
| 0104 | 2.9 | 2.8 | 4.1 | 3.2 |
| 0105 | 3.4 | 3.0 | 1.7 | 2.7 |
| 0106 | 2.1 | 2.5 | 3.1 | 2.6 |
| 0107 | 3.0 | 2.1 | 1.7 | 2.3 |
| 0119 | 4.9 | 5.2 | 3.5 | 4.5 |
| 0120 | 2.9 | 3.0 | 2.9 | 2.9 |

We will standardize (or renormalize) grades from each reviewer across their assignments *within the same grading category*.

Renormalized scores will have a mean of 3 and a deviation of 1.

Renormalized grades from a reviewer on a given proposal will be averaged across 3 categories to produce **the average standardized reviewer grade**.



Critical Changes

Grade renormalization: The average of *average normalized grades* determines proposal grade and preliminary ranking

| Proposal ID | Dr. Cecilia Reyes | Dr. Victor von Doom | Dr. Jean Grey | Prof. Charles Xavier | Dr. Otto Octavius | Grade | Ranking |
|-------------|-------------------|---------------------|---------------|----------------------|-------------------|-------|---------|
| 0103 | 2.8 | 3.8 | 3.8 | 1.8 | 1.8 | 2.8 | 3 |
| 0104 | 3.2 | 4.2 | 3.2 | 4.2 | 4.2 | 3.8 | 6 |
| 0105 | 2.7 | 1.7 | 1.7 | 2.7 | 1.7 | 2.1 | 1 |
| 0106 | 2.6 | 3.6 | 3.6 | 3.6 | 2.6 | 3.2 | 5 |
| 0107 | 2.3 | 1.3 | 3.3 | 1.3 | 3.3 | 2.3 | 2 |
| 0119 | 4.5 | 5.5 | 5.5 | 5.5 | 4.5 | 5.1 | 7 |
| 0120 | 2.9 | 2.9 | 3.9 | 3.9 | 1.9 | 3.1 | 4 |



Critical Changes

Grading Categories and Criteria: We modify the grading categories (within the same framework) and provide new criteria for both the preliminary grading and virtual panel grading phases:

Preliminary grading criteria

| Grade | Impact within the sub-field | Broader impact | Suitability |
|-------|---|--|---|
| 1 | Potential for transformative results | Transformative implications for one or more other sub-fields | Can only be done with HST |
| 2 | Potential for major advancement | Major implications for one or more other sub-fields | Major advantages in using HST over other facilities |
| 3 | Potential for moderate advancement | Some implications for one or more other sub-fields | Some advantages in using HST over other facilities |
| 4 | Potential for minor advancement | Minor impacts on other sub-fields | Minor advantages in using HST over other facilities |
| 5 | Limited potential for advancing the field | Little or no impact for other sub-fields | HST offers little or no advantage over other facilities or advantages of using HST is not clear |

* decimal grades ok



Critical Changes

Grading Categories and Criteria: We will keep *Complexity* grades to provide feedback to JWST FRC modeling, with a focus on programmatic complexity, i.e., in the reviewer's opinion and relative to other programs of similar scope (proposal category, orbit size, etc.), how difficult will it be to carry through the program from award to dissemination of results?



Future Changes

In future cycles, may move to **semantic differential scale** grading—
Reviewers select along a sliding scale with generic polar markers

Possibly not show numeric backbone,
reducing potential for grade confusion

May revisit grading categories and
criteria, allowing for more weight to be
given to categories of specific
emphasis.

0044 Final Grading

Grading Status **Completed**

Scientific Merit
Moderate degree of merit
Breaks some new ground scientifically/reinforces existing hypotheses

Importance to Astronomy
Low importance to astronomy
Multiple teams working in this area of study or similar

Unique Capabilities
HST provides only means to observe
HST provides the only means to effectively complete proposal

Observational Complexity
Medium complexity
Average technical complexity, some possible challenges