

Good and Band Proposals — Report of the trial in 2024 —

Yusuke Aso (KASI)

February 3, 2026

1.1. Background: Review of ALMA proposals

ALMA users submit proposals and receive review comments on them. The distributed peer review provides 10 comments per proposal.

You may find similar comments for your different proposals, like...

- Strengths: This observation aims to distinguish between two scenarios.
- Weaknesses: It is unclear how this goal contributes to the whole field.

1.2. Background: Strengths for a proposal

Rank

- 1: **Objectives** and setup are reasonably justified.
- 2: Good **goals**, **targets**, and **objectives**.
- 2: The **question** is important for **star formation**. ALMA is necessary.
- 3: The **objectives** are promising.
- 3: BD formation is now known well. The **targets** are promising.
- ~
- 5: **Multiple scenarios** are discussed.
Clear **demonstration** of the **goal** and analysis.
- 5: Well **plan** and clear **goal**.
- 5: Increasing the pre-BD sample is important.
- 7: **N₂H⁺** is a good idea. The **robbery case** is interesting.
- 8: The **question** seems essential.
The **filament** enables the **supply-robbery** check.

general words
specific in the proposal

General words for good ranks.
More specific words for bad ranks.
→ Were the reviewers careful when giving a bad rank?

1.3. Background: Weaknesses for a proposal

Rank

- 1: Can distinguish various **motions**? The analyses are for **isolated** cores.
- 2: Need more justification for the lower **resolution**.
- 2: Why pre-BD is important in star formation? The **S/N** seems borderline.
- 3: Other **regions** are $\sim 2x$ closer. Not really **unstable**. Dust **β** is low.
- 3: CARMA is **enough** for the filament. Fragmentation needs a better **res**.
- ~
- 5: Need the justification for the **C¹⁸O** line and **3 mm** continuum.
- 5: Nothing.
- 5: IR does not put a good **upper** limit. All CO would be **frozen out**.
No outflow may be because of the crowded **environment**.
- 7: **<0.01 M_J** in the background. **Why** a larger sample?
CO is **larger** and N₂H⁺ is **smaller** than the core scale.
- 8: Discussion about the objectives is **not well linked** to the mechanisms.

setup including target
science of interest
intention not conveyed

Critical but reasonable
for good ranks.
Misunderstanding in
the worst ranks.
→ Needs to be simpler?
Inevitable?

1.4. Background: Piled up comments

ALMA has reviewed users' proposals since Cycle 0, 2013.

The distributed peer review started in Cycle 8, 2021.

→ Numerous sets of proposals and review comments.

The big data may tell us possible tendencies of typical review comments.

2.1. ALMA Korea's trial: Prop. prep. WS

- In 2024, the ARC node Korea started a workshop dedicated to preparation for the ALMA proposal, mainly targeting graduate students.
- As a first trial, I led a 1-hour session to find and discuss such tendencies.

→ *This talk will review this trial.*



Program for the 1st Workshop

- Date: February 14th (Wednesday)
- Venue: KASI Jangyeongsil Hall, Room 331-2
- 1:00 PM - 1:50 PM: Guidelines for Proposal Writing (Seokho Lee) ([PDF](#))
- 2:00 PM - 2:30 PM: Successful Proposal Writer's Experience 1 (Yusuke Aso) ([PDF](#))
- 2:30 PM - 3:00 PM: Successful Proposal Writer's Experience 2 (Bumhyun Lee) ([PDF](#))
- 3:20 PM - 4:30 PM: Discussion about Reviewers' Tendencies (Yusuke Aso) ([PDF](#)) ([PDF](#))

Program for the 2nd Workshop

- Date: February 28th (Wednesday)
- Venue: KASI Jangyeongsil Hall, Room 331-2
- 10:30 AM - 12:30 AM: Presentations by the participants
- 2:00 PM - 5:30 PM: Presentations by the participants

2.2. ALMA Korea's trial: Goal

The original goal of this discussion session was to find the possible tendencies among multiple proposals and review comments on them.

Can we find some similarities in any category?

- Accepted / declined
- Proposals / comments
- Strengths / weaknesses

Can we find other similarities that are not mentioned in individual comments?

What I imagined...

- “This is the best target.” makes reviewers unnecessarily uncomfortable.
- A schematic figure makes a good impression.

2.3. ALMA Korea's trial: Methodology

- The participants should read many pairs of proposals and comments.
→ The table below.
- The sample should not be biased on either ranks or topics.
→ I set Galactic and extragalactic categories.
- The findings should be checked and discussed among participants.
→ The sample was delivered in advance.

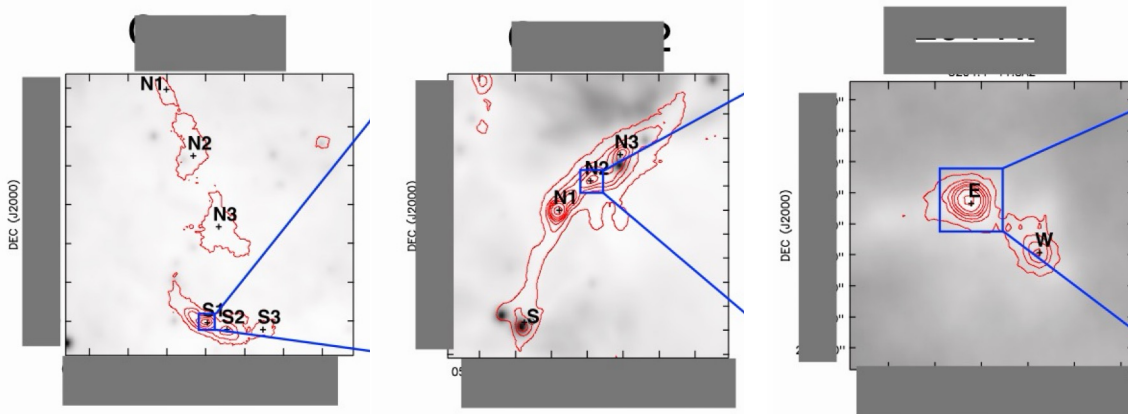
I asked 7 anonymous volunteers to provide their pairs for the workshop.
They kindly provided 12 proposals in total.

	Accepted	Declined
In our galaxy	3	3
Extragalaxies	3	3

Written 2 cycles ago or later.
“Accepted” means working now.
“Declined” means likely retrying.
→ **Anonymity** is necessary.

2.4. ALMA Korea's trial: Preparation

- All the KASI ALMA members looked over some of the 24 documents and hid information that could be used to identify the objects.
- Hiding them is fine because our goal is to find tendencies beyond individual targets.
- We delivered each set to the other category group for anonymity.



Review 1

Rank: 7/10

Comments: The proposal presents a clearly outlined strategy to distinguish polarization due to self-scattering or magnetically aligned grains in the Class 0 disk by obtaining polarization observations at longer wavelength than existing observations. The source is well chosen as it has one of the largest disks among Class 0 sources and Band 7 polarization observations already exist. The proposal could be strengthened by a discussion on whether the requested resolution of $0.35''$, matching the previous Band 7 observations but only resolving the disk major axis, is enough to constrain the polarization origin in the disk, or whether important features could remain hidden at this resolution.

3.1. Results: Group 1

I showed 9 key questions to both groups as examples. This group followed those questions.

1. Did the proposal get good and bad ranks together? What do you feel are the reasons?
 - Differences in reviewers' familiarity.
 - Broad & specific, easy & accurate are necessary.
2. Are there any common good or bad comments? Within a proposal or over some proposals.
 - One says an insufficient resolution; another says sufficient.
 - Some criticisms are based on misunderstanding.
 - Clear justification and explanation are necessary.
3. What expression/factor triggered good or bad impressions? Particularly, a proposer should avoid expressions that trigger a bad impression but are not directly related to the main science.
Too long captions could reduce reviewers' concentration.
4. What fraction of comments is about the idea itself? How about targets or technical settings?

3.1. Results: Group 1

5. What kinds of justification are requested in the comments?

- A lack of data analysis plans results in a bad score.
- Lack of sample size. → “Pilot” can be a solution? → Too competitive for pilots. Start with other instruments?

6. What fraction of comments ask for details more than in proposals? (E.g., the proposal does not mention this possibility; the model/analysis needs to be described in more detail.)

The following 6 comments on a proposal request additional explanations, saying that the goal is ambitious, but the setup is not well justified.

- More careful justification for target selection.
- Figure of spatial distribution.
- Optical thickness is not addressed.
- Clarification about whether the settings are proper for the examination.
- No reference for a feature mentioned in the proposal.
- Clarification about what observables depend on the quantity of interest.

3.1. Results: Group 1

7. Are there any comments about figures or tables?

A comment that the explanation for a figure is too insufficient.

→ Intuitive, uncomplicated figures with good explanations are necessary.

8. Are the weaknesses in the comments based on the scientific discussion in the proposal?

Middle-rank comments are moderate.

It would be difficult to give extreme ranks to unfamiliar fields.

9. What fraction of reviewers do you feel are experts in the proposed science?

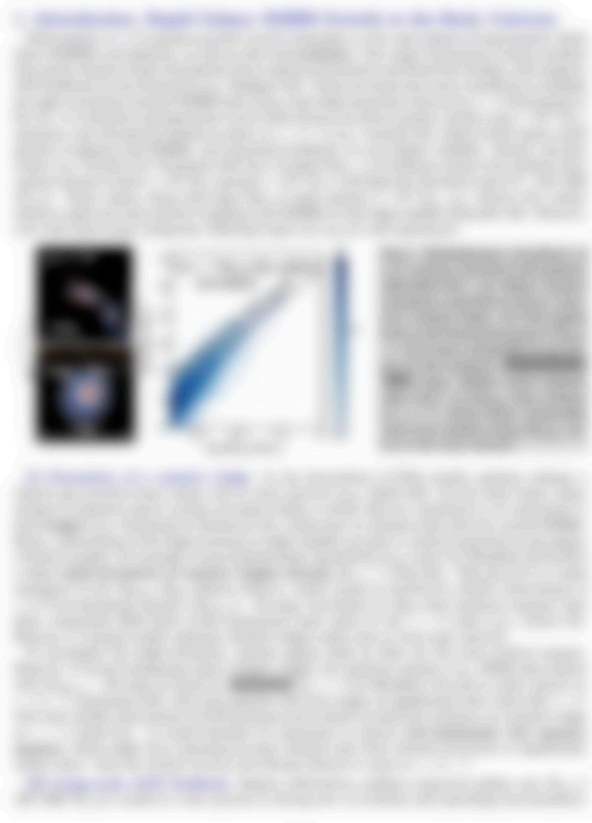
30%–50%, based on the fraction of extreme comments.

Unfamiliar reviewers would check calculations to find critical points.

3.2. Results: Group 2

ID in WS	Rank	Writing structure	Note
1, 2, 3	Overall good	Paragraphs are well-divided. Sentences are structurally firm/solid.	Proposals explains techniques in detail.
4, 5, 6	Good and bad	Connected paragraphs look like chunks.	Comments say the target is not special.

3.2. Results: Group 2



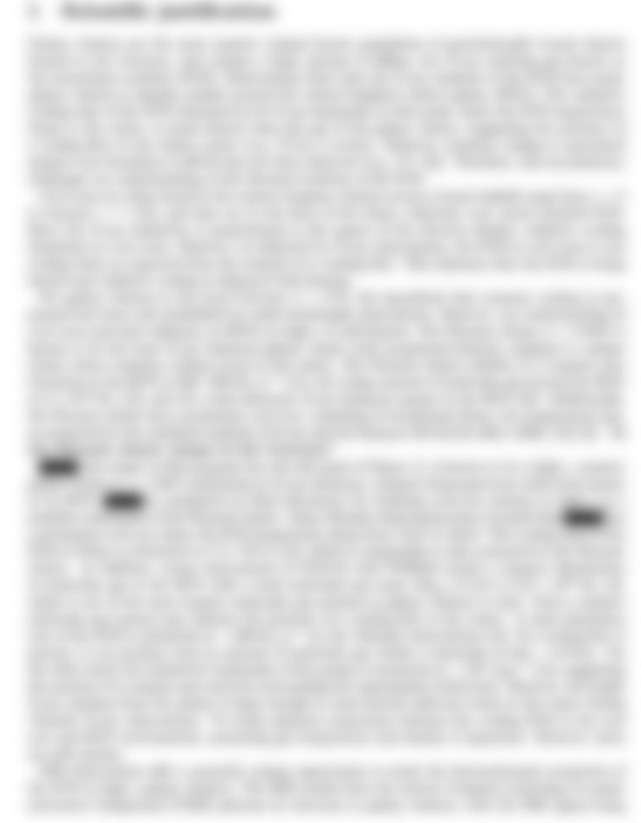
The first pages of two proposals, both of which show the introduction.

Left:

- Subsection
- Figure
- Color (letters, caption)

Right:

The paragraphs longer than one page have no subsection.



3.2. Results: Group 2

The reviewers seem to be divided into two groups.

- Unfamiliar with the field – They prefer connectivity/consistency in the context.
- Familiar with the field – They prefer detailed explanations and robust feasibility.

Note. There is no actual information on how familiar each reviewer is.

Composition of comments

- Few comments criticize the technical parts.
- Not many comments criticize ideas themselves.

4.1. Discussion: Common results

Both groups found that

- The variation in reviewers' familiarity with the topic.
- Lack of explanations results in a bad rank.
- The feasibility, including calculations, seems to be checked carefully.

4.2. Discussion: Outcome

The participants found typical and rare comments.

- Request for additional explanations.
- Feasibility check.
- Good and bad ranks for declined proposals.

The participants also **imagined** how a reviewer would feel.

- Reviewers' concentration.
- Pilot proposals would not be attractive to reviewers.
- Chunks of paragraphs make a bad impression on reviewers.
- *Reviewers may get tired after reviewing many proposals.
Catching early IDs might be advantageous.*

4.3. Discussion: Limitation

- Statistical analysis was limited by the small sample size and the lack of computer use.
→ Interesting, but sensitive to IP or privacy.
- No hypothesis was made before the delivery or discussion.
→ A potential direction, although this could limit the outcome.
- No format of discussion, albeit with example questions.
→ With a format, the participants would not have speculated about how a reviewer feels, which was not the main focus.
- Graduate students, different sciences.

4.4. Discussion: Survey after WS

In the survey after the 2024 workshop

- 4 out of 10 participants selected the discussion as a useful session.
- “The discussion session was good for experienced participants.”
- “It was useful to know tips that are not in the official guide.”

The 2025 workshop did not have the discussion session.

In the survey after the 2025 workshop, 3 out of 12 participants mentioned that the discussion session would be useful to know why a proposal was accepted/declined.

5. Summary

- KASI started the proposal preparation workshop in 2024.
- There, I led a 1-hour session where the participants discussed tendencies in review, using 12 pairs of comments and proposals.
- The participants were divided into two groups, and both found:
 - ✓ Variation in reviewers' familiarity with the topic
 - ✓ Bad rank due to lack of explanations
 - ✓ Importance of feasibility
- The participants also imagined reviewers' feelings.
- To improve this trial, we should think about the statistical analysis, setting a hypothesis/format, and the target/purpose.