Bulletin of the AAS • Vol. 54, Issue 1 (Obituaries, News & Commentaries, Community Reports)

Report on the Review of ATP Inclusion Plans by DEI Expert and Science Expert Panels

Tim Sacco¹, Dara Norman¹

¹NOIRLab

Published on: Mar 03, 2022

DOI: 10.3847/25c2cfeb.19262acc

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

ABSTRACT

In this report, our team analyzes the reviews of the inclusion plans recently required by NASA's Astrophysics Theory Program (ATP). Specifically, we identify and assess themes presented by PIs submitting research proposals, and how these themes were evaluated by both science and expert inclusion review panels. We highlight the types of inclusion plans that PIs develop when given only general guidance, as well as show how science and expert inclusion panel reviewers evaluate proposed inclusion plans. While there is significant overlap in how the science and expert inclusion panel reviewers evaluate inclusion plans, we find the expert inclusion panel reviews were typically more detailed in both their evaluations of proposal strengths and weaknesses than the reviews provided by the science panels.

EXECUTIVE SUMMARY

This report stems from an analysis of research inclusion plans recently required by NASA's <u>Astrophysics Theory Program</u> (ATP). In support of NASA's core value of Inclusion, NASA's ATP piloted the addition of a one- to two-page inclusion plan to the proposals in which PIs were asked to address: (1) plans for creating and sustaining a positive and inclusive working environment for those carrying out the proposed investigation; and (2) contributions the proposed investigation will make to the training and development of a diverse and inclusive scientific workforce.

NASA staff charged us to review the distribution of key themes that were presented by PIs, as well as how these themes were assessed by a panel of diversity, equity and inclusion (inclusion) experts for 120 proposals. We compare a subsample of 32 expert inclusion reviews with reviews provided by the panel of science experts reviewing the same proposals. We have future plans to compare the full set of 120 proposal reviews made by science expert panels.

1 DATA AND METHODS

NASA Astrophysics Theory Program (ATP) provided our team with 120 Inclusion Plans from science proposals. These Inclusion Plans were reviewed by: (1) a science panel of astronomers, and (2) an inclusion panel of experts on diversity, equity, and inclusion (DEI) initiatives. ATP provided all of the expert inclusion panel reviews, and also a spreadsheet that detailed the distribution of adjectival ratings for all 120 Inclusion proposals provided by the DEI expert reviewers. The adjectival score categories are as

follows: Excellent (E), Excellent/Very Good (E/VG), Very Good (VG), Very Good/Good (VG/G), Good/Fair (G/F), Fair (F), Fair/Poor (F/P), and Poor (P).

Recognizing that, due to available resources, extracting the astronomer comments from all 120 proposals would be impractical for this initial comparison, we requested a subsample of science review comments for 32 of the proposals that roughly aligned with the distribution of the expert scores. (We plan to revisit a comparison of the full set of science review comments in a later paper.) *Figure 1* shows a comparison of the distribution of 120 inclusion proposal scores with the 32 proposals from which we sample. The focus of this distribution of ratings (G, G/F, F) reflects our interest in understanding how the community of astronomers assessed inclusion plans that DEI experts found to be marginal or inadequate. Proposals with an expert rating of poor (P), were in nearly all cases, unquestionably inadequate and it is likely that astronomers' reviews would also reflect this. Likewise, plans rated very highly are likely to have to have been recognized as such by the astronomers. Thus, while we examine the few highly ranked proposals, we focus attention on more marginal plans. We used NVivo qualitative analysis software to identify key themes in the inclusion plans. We do not have adjectival ratings from the science expert reviewers.

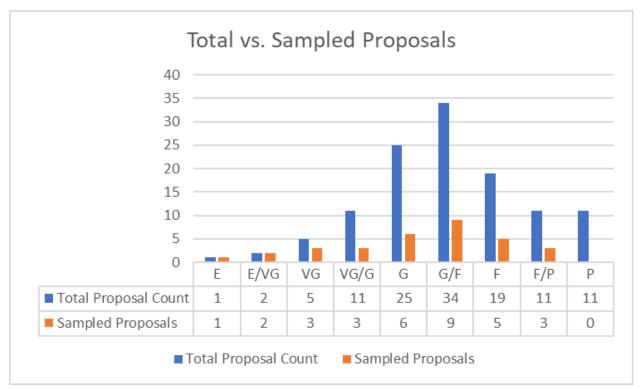


Figure 1: The distribution of expert inclusion review panel grades for the 120 Inclusion Proposals (blue) compared to the 32 subsampled proposals (orange) that include comments from the science expertise reviewers. The adjectival ratings are all from the DEI inclusion experts since these ratings were not provided by the science panel members.

2 ANALYSIS

Below, we examine several themes that PIs cover in the Inclusion Plans for the 120 proposals. First, we focus on the theme of 'DEI credentialing,' in which project PIs outline their personal history or the history of the institutions for which they work with equity and inclusion efforts. Then, we discuss the themes of leveraging institutional resources, fostering an inclusive environment, and cross institutional partnerships. Within these larger themes, we will review key aspects that contributed to the ranking of the proposals. See Table 1 for a summary. We will also use our sample of 32 proposals to compare how science expert reviewers assessed proposals compared to expert DEI inclusion reviewers.

		# of	% of Total
Theme	Sub-theme	proposals	proposals
DEI Credentialling		64	53%
	History with DEI work	44	37%
	Institurion DEI Credentials	30	25%
	Department DEI Credntials	22	18%
Leveraging Institutional Resources		68	57%
	Bridge programs	9	8%
Fostering an inclusive Environment/Climate		84	70%
	Code of Conduct	22	18%
	Normal mentoring practice	27	23%
	Evaluation of inclusivity	12	10%
	Substantial Evaluation plans	4	3%
	Open Communication	15	13%
	Regular Meetings	23	19%
	Equitable Recruitment	13	11%
	DEI training	13	11%
Cross Institutional Partnerships		13	11%
	Credentialling	8	7%
	Leverage partnership for DEI	5	4%

Table 1: PI Inclusion Plan main themes discussed in this report with frequency of use in the 120 proposed plans. Each proposal may include multiple themes and subthemes.

2.1 DEI Credentialing

More than half of the inclusion proposals (64 out of 120) attempted to outline the DEI credentials of a project, whether that focus on the PI's past credentials, or efforts taken in the host department or university. Forty four out of 120 proposals highlighted the PI's history with DEI to signal that the inclusion plan would likely work. Often, these proposals outlined PIs' past DEI activities, but varied in how the descriptions of activities were evaluated by the inclusion expert panels. For instance, one proposal (G) detailed the PI's role in establishing a 'Women in Physics' working group in their department. Another proposal (F/P) reports the PI's history of lecturing in summer schools with diverse students, and another (F) reports the PI's history of including undergraduate and graduate students on research projects. It was also common for proposals to outline the institutional credentials of the environment in which projects are embedded. Thirty out of 120 outlined institutional efforts at the university level to credential the project, while 22 out of 120 outlined department level credentials. For instance, one proposal (F) reports its host department eliminated the GRE

requirements for graduate admissions, which had broadened the pool of graduate applicants. Another proposal (G) details a Diversity and Inclusion working group in the host department, which has created means for anonymous feedback on inclusion efforts and works to secure grants to support URM students. Comments suggest that the differences in scores relate to the panel's ability to tie the departmental or institutional actions to those of work done by the PI.

The above types of DEI credentialing shaped how both sets of reviewers evaluated proposals. In some highly ranked proposals, those with grades E/VG to VG/G, panel reviewers did not mention the fact that proposers had spent time detailing their DEI credentials, and instead used the review space to highlight the strengths and weaknesses of concrete plans put forth by the PIs. Panel reviewers often noted when slightly lower ranked proposals (most of those scored VG, G or G/F) reported their DEI credentials as a positive. For instance, science panel reviewers lauded one proposal (VG) because the PI had had a history mentoring women and people of color as a major strength of the project, while the expert inclusion panel reviewers did not comment on the DEI credentials of the PI. Similarly, both science and expert inclusion panels celebrated another proposal (G) because the PI had worked on their department's graduate recruitment committee, which would in turn shape their inclusivity practices.

In some lower ranked proposals, how PI's outlined their DEI credentials were listed as both strengths and weaknesses by inclusion expert panels. For instance, in one proposal (F), expert inclusion panel reviewers noted the PI's history with DEI as a strength of the project, but also wrote "the plan...did not clearly state how [the diversity and inclusion activities of the PIs] would be an integral part of the way in which the proposed research would be accomplished." In one proposal (G/F), expert inclusion panel reviewers listed the PIs past experience with DEI as a strength, but also critiqued the proposal because it gave no details about the university or department contexts that would influence the inclusion activities of the project.

Both science and expert inclusion panel reviewers often critiqued proposals for not adequately tying credentials to the broader inclusion plan of a project. In one proposal (G/F), expert inclusion panel reviewers described the PIs involvement in a Women in Physics group as a strength, but critiqued this credentialing suggesting that the PI did not seem to have experience with other axes of diversity. Expert inclusion panel reviewers looked favorably on a (G/F) proposal because of the PI's history with the APS-Inclusion, Diversity, and Equity Alliance (APS-IDEA), but critiqued the proposal for a lack of "discussion about how the resources from this network will be used to

help guide programming efforts." These examples also highlight a common approach to how expert inclusion panel reviewers structured feedback: they would note the credentials as a broad strength while digging into the weaknesses of how it was being leveraged was a common approach for reviewers to organize their reviews by. (See additional commentary below in section 2.5.)

How PIs described their work with underrepresented populations sometimes undermined their DEI credentials as well. For instance, in one (G/F) proposal, the PI wrote that, while they were unsure if the doctoral student that would work on the project would be from an underrepresented group, they were open to working with minority students and had a history of working with students on past projects "that can be identified as minorities." The expert inclusion panel reviewers wrote that this specific section of the proposal "demonstrated a difference in tone...where the discussion of minorities could be perceived as exclusionary." While the PI is attempting to lay out their past work mentoring underrepresented students or postdocs, expert inclusion panel reviewers also critiqued a proposal (G) for "[demonstrating] a very limited understanding of the real challenges/barriers to equity and inclusion relevant to the field" based on how this topic of diversity is discussed.

Looking at how different panel reviewers evaluated inclusivity in credentialing proved to be a fruitful point of comparison. For instance, expert inclusion panel reviewers of a (G/F) proposal noted that the PIs description of working with students with "emotional or mental disabled students" compromised the proposal's goals of inclusivity. Science panel reviewers also took note of this language, writing: "The proposal mentions "... (emotional or mental[sic] disabled) students' as being possibly disfavored by an informal weekly meeting but does not provide any evidence to support this statement. Many other types of students may also be disfavored by these informal weekly meetings." Here, we see the science panel and expert inclusion panel reviewers latching onto the same phrase, both listing it as a weakness, but there are differences: science panel reviewers wanted more inclusive language in that informal meetings may be detrimental to more students than those singled out in the proposal while the expert inclusion panel reviewers raised concerns over the language and its implications for inclusivity more broadly.

This point is highlighted in another example. The expert inclusion panel reviewers write that, in one proposal (G/F), the PI's discussion of a current student of the project was "not indicative of an inclusive or welcome environment." Specifically, the proposal revealed an "inappropriate amount of detail about this student's family situation...

[and] did not address any efforts to understand and mitigate the barriers to participation in the project for this student." The science panel did not focus on these aspects in their review. The only weaknesses of the proposal noted by the science panel reviewers were that "the proposal does not provide sufficient information on how the PI will create an inclusive environment" and "some of the terms used in the Inclusion Plan are not spelled out." While both panels express similar concerns, the comments from inclusion expert panel members are much more direct, providing much less ambiguity about the concerns as feedback to the PI.

2.2 Leveraging institutional resources

Sixty eight out of 120 made some reference to leveraging institutional resources available to them through the host university or department. Both science panel and expert inclusion panel reviewers viewed the ability to leverage resources from host departments or institutions as a strength to proposed projects. However, inclusion reviewers wanted more concrete detail on how institutional resources would be leveraged or connected to the project. For instance, expert inclusion reviewers wrote that one (F) proposal did not "adequately describe how institutional resources...available external to the project would be strategically integrated into the project...to foster inclusion." Expert inclusion panel reviewers sought more information for how a proposal (G/F) would draw on department resources in strengthening the proposed project. Conversely, for one (G) proposal, expert inclusion panel reviewers wrote that the PI's plan to draw on the various training programs offered by the host institution as a way to train graduate students on the project as a major strength of the project.

Several proposals mentioned specific programs, like Bridge programs, that could be leveraged in recruiting more diverse applicants for positions on a research project. Both science and expert inclusion panels saw access to Bridge programs or similar programs as a potential strength of projects. But inclusion expert panel reviewers often sought more clarity regarding the specific links between these programs and the proposed projects. Expert inclusion panel reviewers wrote that a (G/F) proposal mentioned "a student team member from a Bridge program...[but] the proposal does not sufficiently detail adoption of recruiting or selection practices that are aligned with the goal of diversity." Similarly, in a review for a (G) proposal that states it will draw on an affiliated Bridge program to foster a more inclusive environment, expert inclusion panel reviewers wrote that the proposal lacked "detail as to what the flaws are with current practices and what solutions the PI will propose to rectify them." Just

mentioning the Bridge program is not sufficient; inclusion reviewers wanted to know how resources were going to be used.

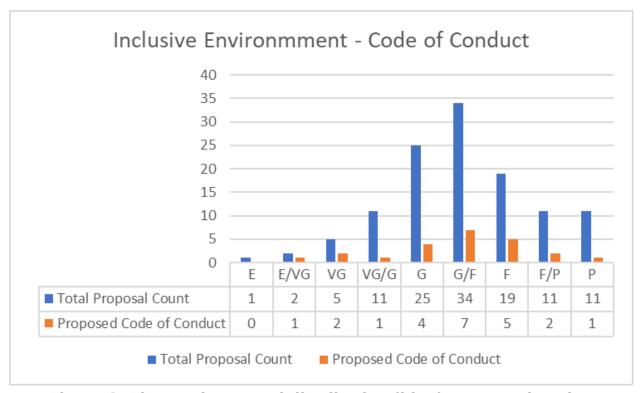


Figure 2: The total proposal distribution (blue) compared to the proposals that include a code of conduct. The plot shows that the Codes vary in quality of proposed execution and cover the full range of scores rankings given by expert inclusion panels.

2.3 Fostering an inclusive environment

A key component of NASA ATP's call was how proposals would create and sustain a positive and inclusive working environment for the research team. Eighty four out of 120 proposals attempted to describe how they would foster an inclusive environment for the research team. One of the most common strategies was to adopt a formal Code of Conduct that would guide team meetings and interactions. Twenty two out of 84 (26%) proposals sought to establish a formal Code of Conduct. Both science and expert inclusion panels noted that the planned Code of Conduct in one (VG) proposal is a strength to fostering an inclusive environment, but the expert inclusion panel wrote a weakness of the proposal was a lack of explaining *how* a Code of Conduct would create a more inclusive environment. Again, science and expert inclusion panel reviewers have different sets of expertise; inclusion reviewers are more concerned with the concrete details of the inclusion plans and the effects of the proposed activities.

Figure 2 highlights the distribution of proposals that put forth some sort of Code of Conduct in relation to the overall score distribution of inclusion proposals. This figure shows that proposed codes of conduct are pretty evenly distributed across scores, mirroring the broader score distributions of proposals. In terms of inclusive practices, Codes of Conduct are relatively easy to adopt with little formal commitment from a team. Low stakes inclusion practices like Codes of Conduct are what proposers will do when they are given little guidance about what kinds of inclusion strategies are effective or acceptable. Similarly, 27 out of 84 proposals (32%) that put forth a plan for fostering an inclusive environment proposes achieving these goals through mentoring, most of these were "normal mentoring practices" that would be expected of PI/student relationships even when an inclusion proposal was not required. What these numbers show is that, without guidance, proposers will take the path of least resistance. As a result the strategies put forth are not intentionally focused on racial or ethnic inclusion.

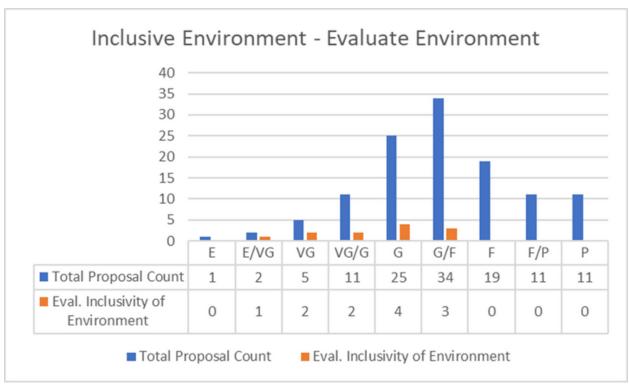


Figure 3: Proposals that plan to evaluate the inclusivity of their environment (orange) compared with the overall distribution of research inclusion plans. The lowest ranked proposals tend to lack proposals for these kinds of evaluations.

Fewer proposals, only 12/84 (14%), put forth a plan to evaluate the inclusivity of the environment. *Figure 3* shows the distribution of proposals that proposed evaluations of the working environment compared to the total distribution of inclusion proposals by

rank. The plot shows that proposals that received the lowest rank (F, F/P, and P categories) of grades did not propose any kind of evaluations. Looking more closely at these 12 proposals that did propose evaluations, one third of them (4/12) are actually substantial plans. However, with different qualities as demonstrated by the spread of rangings, indicating that some PIs were more successful in explaining how they would implement these plans compared to others. One (E/VG) proposal planned to assemble an Inclusion Advisory/Evaluation Committee for the project that would inform the project over time. "The PI will work with this committee in developing recruiting strategies...The committee will meet the project team [with and without] the PI once a semester, conduct a written survey of all team members once a semester and write an annual inclusion evaluation report to be included in the annual progress report to NASA." Both science and expert inclusion panel reviewers noted the evaluation plan was a major strength of the proposal. The only weakness given in regards to the plan was by the expert inclusion review panel, who wanted more information about whether there was precedent for such a committee at the PIs home institution, and how the committee would be "established in practice."

Less substantive evaluation plans still received reasonable scores from reviewers. For instance, one proposal (G) wrote that inclusion goals "will be reviewed and revised...at least once a year. This is done to evaluate research progress, avoid misunderstandings, and identify any possible improvements in the working relationship." The PIs propose an evaluation system that facilitates better relationships between mentors and mentees. The expert inclusion panel reviewers critiqued another proposal (G) for failing to connect evaluative practices to broader DEI goals. Another proposal (G/F), put forth a plan in which the team leadership would "ensure progress...over diversity, equity, and inclusion...[by developing] a layered understanding of how to evaluate progress, success, and ongoing challenges." Both science and expert inclusion panels critiqued the plan. Science panel reviewers were critical that the plan "does not mention any specific actions related to participants from minority groups." Expert inclusion panel reviewers wrote a major weakness was that the plan did not adequately discuss "how climate and activities would be evaluated." In another instance, science reviewers critiqued a (VG) proposal "much of [the inclusion plan] relies explicitly on the PI - for example to evaluate how things are working in the yearly reviews. Some mechanism for auditing by someone else...would be useful." Similarly, the expert inclusion panel reviewers wrote: The proposed plan lacks an adequate description of metrics of success and a related assessment plan...and lacks a

detailed description of the [proposed] rubric. It could be beneficial to seek the advice of experts in the rubric's development, as this is non-trivial."

In contrast with the adoption of Codes of Conduct, it is difficult to enact a proper evaluation of a project or program. The fact that fewer proposals tried to enact evaluations suggested this is the case and not all that planned an evaluation were successful. It is worth noting, however, that these plans still show a more narrow range of rankings than those of the plans that only planned to adopt a code of conduct.

2.4 Cross-institutional partnerships

The authors of this report are part of the US Extremely Large Telescope Program's (US-ELTP) Research Inclusion Initiative (RII) team. The RII is designed to support research participation of the broadest astronomical community by specifically addressing the concerns of researchers at small and under-resourced institutions interested in participation in US-ELTP. One key strategy for broadening participation of underrepresented researchers is fostering cross-institutional partnerships between researchers at small or under-resourced institutions and research university astronomers that are more traditionally represented in proposals for telescope time. In recent years, federal agencies have championed cross-institutional partnerships as an effective strategy for increasing the participation of underrepresented groups across science, technology, engineering, and math disciplines. 12

Out of 120 proposals, only 13 discuss cross-institutional partnerships in relation to the proposed project, eight of which discussed cross-institutional partnerships in a DEI credentialing way. For instance, one proposal (G/F) notes that the proposed project would take place in a department that is working with a Bridge program. A (VG/G) proposal notes that Co-Is on the project (located at another institution) have a history of recruiting URM students. Another proposal (F/P) states the PIs will "do [their] best" to recruit students from a high school partnership and REU summer interns program.

We differentiate these credentialing proposals from the five of the 13 proposals to discuss cross-institutional partnerships with plans to leverage the partnership as a strategy to achieving diversity, equity, and inclusion. *Figure 4* highlights the distribution of all proposals to discuss cross-institutional partnerships with those that had a plan to leverage them. For instance, a (G/F) proposal commits to recruiting the students that will work on the project from a minority-serving institution via a Bridge program. Another proposal (VG/G) plans to leverage the fact that the project spans multiple institutions for mentorship by requiring that students on the project have an

external mentor at one of the other institutions. A proposal ranked (F/P) for being inadequately short said it would recruit minority students for the project through the 'Minority University Research and Education Project.' A (G) proposal reports that the proposed project is part of a multi-institution collaboration that includes a HBCU and is affiliated with the PAARE program, writing that "funding the research proposed here will provide opportunities for underrepresented minorities to collaborate with the project."

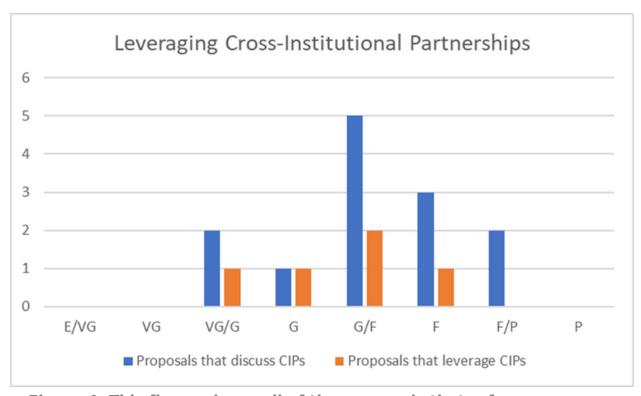


Figure 4: This figure shows all of the proposals that references crossinstitutional partnerships as part of their plan. The blue histogram indicates proposals that only mention such partnerships, whereas the orange histogram shows those fewer proposals that leveraged crossinstitutional partnerships as part of DEI.

As *Figure 4* highlights, a team did not automatically do well just because they attempted to leverage cross-institutional partnerships. As the above figure highlights, one proposal that tried to leverage a cross-institutional partnership still received an F score. There are teams out there attempting to do the right thing, but not uniformly doing a good job with it.

2.5 Assessing students

Forty eight of the 120 inclusion proposals made reference to some aspects of assessment within or of the projects. More than half of these (26/48) discussed ways in which students would be evaluated within the proposed projects. Often student assessment was through some formal tool like Individual Development Plans, a STEM training tool that has become increasingly common in recent years. Seven proposals planned to use formal student management tools. Others proposed less formal systems to assess student or postdoc progress on a project. For instance, one proposal (VG/G) outlines a plan to provide students and postdocs on the project with written reviews once per semester. "This will highlight important milestones (a) towards completion of their project, (b) along the career path, and (c) any other issues that they feel need attention." Similarly, another proposal (G) states: "the PI will provide [students] timely feedback on their progress, results, and presentations." Social scientists of STEM mentorship and student assessment have found that these practices do not benefit all students equally. More guidance on equitable ways to mentor and assess students on a project would be beneficial for future research inclusion panels.

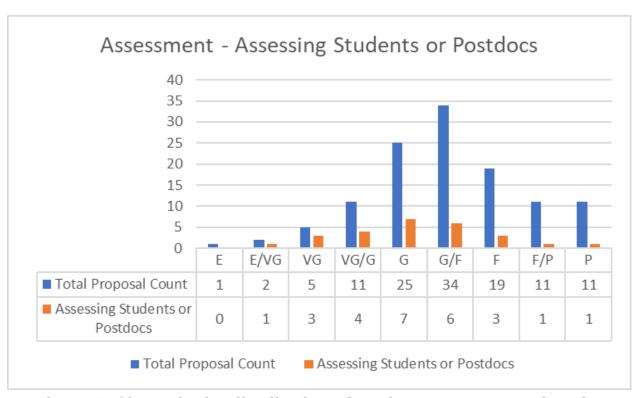


Figure 5: Shown is the distribution of student assessment plans in relation to the overall distribution of inclusion panel scores. This figure shows that the distribution of plans with student assessment components maps with the overall distribution of inclusion proposals. Student assessment does not correlate with higher scores on their own, and PIs should use guidance about good student assessment practices.

2.6 Differences in how the science panels and expert inclusion panels ranked a topic

There were some ways in which the reviews from the science panels and expert inclusion panels overlapped, however, there were also some differences. First, the expert inclusion panel reviews were typically more detailed in both their evaluations of proposal strengths and weaknesses than the reviews provided by the science panels. There were instances in which the science panel reviewers did not list any strengths or weaknesses of a proposal while the expert inclusion panels identified strengths and weaknesses in these same proposals. Some of this difference may be due to instructions given to each group, however it is also likely that this disconnect is because science and expert inclusion panel reviewers are trained to examine different things. Take, for instance, cross-institutional partnerships, a topic in which we are particularly interested. *Figure 6* highlights how science panel reviews compared to expert inclusion panel reviewers in their approach to cross-institutional partnerships.

For instance, expert inclusion panel reviewers of a F/P proposal wrote that a strength of the proposal was the plan to recruit summer interns from MSIs, but a weakness of the proposal was that it did not describe the specific training activities that would be offered to summer interns. Science panel reviewers did not list the partnership as a positive, but did note that a weakness of the proposal was that it was unclear if the PI had access to programs that could help recruit MSI students. In this instance, the science reviewers are thinking about the activity as whether the project has access to a specific resource, while the expert inclusion panel reviewers approach the same proposal but ask a different question: is the inclusion plan developed enough that it is actually going to benefit the minority students it is designed to serve in a meaningful way.

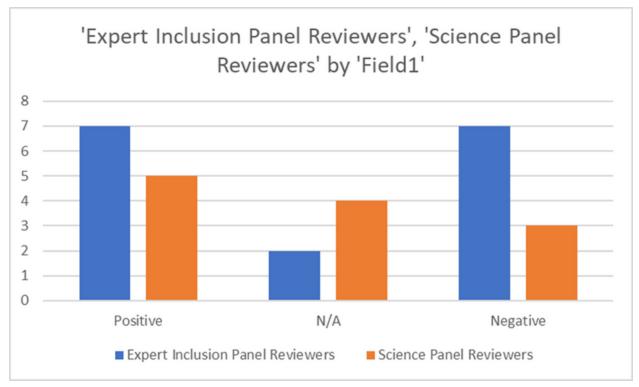


Figure 6: Only 12 proposals (out of 120) discussed some kind of crossinstitutional partnership (CIP). Figure 6 represents how reviewers
engaged with mentions of CIPs. There were instances when a review
noted CIP as both a strength and a weakness, and each instance is
plotted in the figure above. N/A represents instances in which a proposal
discussed a cross-institutional partnership, but the science or expert
inclusion panel reviewers didn't mention it as either a strength or a
weakness in the review. The plot shows that science reviewers comment
on these CIPs with less frequency, perhaps suggesting some
inexperience with the theme. However, when a critique is made, the
types of comments seem to track those of the inclusion expert panels
fairly well.

3 SUMMARY AND CONCLUSIONS

The proposed Inclusion Plans submitted by PIs were found to be mostly marginal or inadequate as reviewed by DEI expert panels, with the distribution of scores peaking at G/F (between good and fair). More proposals received scores on the lower end of the rankings, with very few in the top rankings. This demonstrates the need in the astronomy community for resources or training to help PIs understand what makes for a Very Good or Excellent proposal.

Given the very general nature of the NASA proposal instructions, PIs tended to focus the themes of the activities in their plans around topics that are the easiest to execute (e.g., codes of conduct), not necessarily those themes that would lead to the best advances in inclusion for their teams or departments. Showing credentials (i.e., credentialing) of one's own individual experience with DEI, or that of one's department or university in various ways was a popular theme. However, there was significant breath in the quality of such proposals. Generally, credentialing alone was not sufficient for the plan to be ranked highly and proposers needed to demonstrate how past experience qualified the proposer to do well with the current plan. The best proposals connect the PI's past DEI experiences to how they had thought out the details of their research inclusion plan.

As is evident from Table 1 and in other parts of this analysis, proposers rarely tackle the more difficult issues surrounding these few themes that we analyzed in this report. For example, while fostering an inclusive environment is discussed in 84% of plans only about 3% of proposers include plans to evaluate their working environment. Similarly, cross-institutional partnerships are rarely mentioned (11%) and even when they are, in only 4% of cases are those partnerships being leveraged to support DEI work put forth by proposals. The rest were mentioned primarily to outline the DEI credentials of a project. These findings suggest that without guidance, the astronomy and astrophysics community is not, generally, prepared to think about what makes significant change with respect to research inclusion means for their own work. The community needs guidance to better identify and craft plans that will achieve the goals of NASA's core value of inclusion.

Similarly, cross-institutional partnerships can be hard to make work, but may offer some of the best opportunities for inclusion along the difficult axes of ethnicity/race for addressing underrepresentation in science. However, because it is difficult to build these partnerships, few PIs are able to credibly add these to their proposal, and indeed few attempted to do so. More information could be provided to proposers regarding how cross-institutional partnerships can broaden participation for underrepresented groups and ways they can find opportunities for cross-institutional partnerships.

We also note that we had anticipated greater differences between the science and expert inclusion panel reviews, but the science panel reviewers often matched the overall sentiments expressed in reviews provided by the DEI expert panel. However, each group honed in on different elements of the proposal to question and critique, with the DEI expert panel probing questions and concerns with more detail. For instance, while both science and expert inclusion reviewers celebrated proposals that leveraged institutional resources for DEI goals, the expert inclusion reviewers sought

more clarity on how institutional resources would be leveraged or connected to the project. And generally, the expert inclusion panel reviews were typically more detailed in both their evaluations of proposal strengths and weaknesses than the reviews provided by the science panels. These findings suggest that we need interventions aimed at ensuring expertise on inclusion is available to the reviewers of inclusion plans, otherwise, reviewers may be asking the wrong questions in determining the potential of an inclusion plan to promote equity on the team. NASA could appoint inclusion experts to science panels or provide training/best practices/making sure that expertise is conveyed or available to the community.

Having DEI expertise in the review of inclusion plans has important consequences for how the proposals are reviewed. While this initial data suggests that in some instances science expert proposers adequately identify good vs bad proposals, each group is looking at the proposal with a different view of what concerns need to be addressed by these plans.

Restriction of the assessment of inclusion plans to a purely DEI expert committee could be counter-productive to goals on inclusion. It seems, from the analysis here, that the general Astrophysics community is not currently at the same level of awareness and understanding of the issues around DEI as an inclusion expert committee. However, if the goal for NASA by making inclusion a core value of the agency, is the normalization of inclusive practice in all parts of its mission, then science PIs and reviewers alike will need to learn more about and embrace the issues of inclusion as part of how science research is conducted and how collaboration and research groups operate. Segregation of the review of inclusion plans into expert panels is not likely to achieve this goal in the near term. Normalization of the best practices around inclusion will require that the full community embraces these goals as part of the culture of astronomy and astrophysics. The use of rubrics and guidance will be important to bringing the Astrophysics community along in their fuller understanding of the issues, concerns and best practices of making inclusion part of their research work. DEI experts should continue to be consulted in building rubrics and giving guidance. Workshop training opportunities for science reviewers that demonstrate what they should be looking for in the inclusion plans could also be helpful.

Footnotes

1. National Science Foundation. 2015. CEOSE Biennial Report to Congress 2013-2014: Broadening Participation in STEM.

https://www.nsf.gov/od/oia/activities/ceose/documents/2013-2014%20CEOSE%20Biennial%20Report%20to%20Congress_Final%20Version_09-08-2015.pdf-

2. Hardcastle, Valerie Gray, Stacie Furst-Holloway, Rachel Kallen, and Farrah Jacquez. 2017. "It's complicated: A multi-method approach to broadening participation in STEM." *Equality, Diversity, and Inclusion: An International Journal* 38 (3): 349-361. <u>—</u>