Analysis of Panel Summaries of Proposals Submitted to the S-STEM Program

Ms. Samara R. Boyle, Rice University
Dr. Yvette E. Pearson P.E., Rice University

Dr. Yvette E. Pearson holds a B.S. in Civil Engineering and M.S. in Chemistry from Southern University and a Ph.D. in Engineering and Applied Science from the University of New Orleans. She is Associate Dean for Accreditation, Assessment, and Strategic Initiatives in the George R. Brown School of Engineering at Rice University, a Program Evaluator for the Engineering Accreditation Commission of ABET, a registered Professional Engineer in Louisiana, a former Program Director in the Division of Undergraduate Education at the National Science Foundation, and a Fellow of the American Society of Civil Engineers.

Dr. Margaret E. Beier, Rice University

Margaret Beier is a Professor of Psychology at Rice University in Houston, TX. She received her B.A. from Colby College, and her M.S. and Ph.D. degrees from the Georgia Institute of Technology. Margaret's research examines the predictors of performance in educational and occupational settings. In particular, she is interested in the effects of examining gender, age, ability, personality, motivation, and self-regulation on a range of outcomes. She is a member of the American Educational Research Association and a Fellow of the Society for Industrial and Organizational Psychologists.

Ms. Jacqueline Gilberto, Rice University
Prof. Stephen P. Mattingly, University of Texas at Arlington

STEPHEN MATTINGLY is a Professor in Civil Engineering at the University of Texas at Arlington. Previously, he worked at the Institute of Transportation Studies, University of California, Irvine and University of Alaska, Fairbanks. His most recent research projects address a variety of interdisciplinary topics including developing an app for crowd-sourcing bicycle and pedestrian conflict data, transportation public health performance measures, policy and infrastructure improvements resulting from bicycle and pedestrian fatality crashes, linking physical activity levels to travel modes, transportation mobility for the transportation disadvantaged, and the development of planning and transit performance measures for access to opportunities, integrating sustainability into the engineering curriculum and creating an engineering sustainability minor.

He has published several articles in the Transportation Research Record, other journals and conferences on these and other related topics. He is currently serving on the Transportation Research Board (TRB) Committee on Aircraft/Airport Compatibility and is a past member of the TRB Committees on Traffic Flow and Characteristics and Transportation Network Modeling. Stephen is also a member of the American Society for Engineering Education (ASEE).

Dr. Ann Saterbak, Duke University

Ann Saterbak is Professor of the Practice in the Biomedical Department and Director of First-Year Engineering at Duke University. Saterbak is the lead author of the textbook, Bioengineering Fundamentals. Saterbak's outstanding teaching was recognized through university-wide and departmental teaching awards. In 2013, Saterbak received the ASEE Biomedical Engineering Division Theo C. Pilkington Outstanding Educator Award. For her contribution to education within biomedical engineering, she was elected Fellow in the Biomedical Engineering Society and the American Society of Engineering Education.

Anila K. Shethia, Rice University

Anila K. Shethia holds a B.B.A. in Management Information Systems and an M.B.A. from University of Houston. She is currently the Education Research Manager in the George R. Brown School of Engineering at Rice University.

Analysis of Panel Summaries of Proposals Submitted to the NSF S-STEM Program

Abstract

This research paper describes a preliminary analysis of panel summaries of proposals submitted to the National Science Foundation's (NSF's) Scholarships for Science, Technology, Engineering and Mathematics (S-STEM) program. S-STEM provides awards to institutions to fund scholarships and to implement evidence-based strategies to recruit, retain, and graduate students from low-income backgrounds who have the academic potential to succeed in eligible STEM disciplines. The ultimate goal of the program is to build the US STEM workforce. In 2017, Rice University received funding from NSF to support teams of principal investigators and their co- investigators, who were experts in educational or related research areas, to attend a two-day workshop that was developed to help them prepare more competitive proposals to the S-STEM program. The emphasis was on investigators from predominantly undergraduate institutions, primarily those located in Established Program to Stimulate Competitive Research (EPSCoR) jurisdictions and/or designated as Minority-Serving Institutions. One of the workshop's aims was to investigate factors that impact the success (or lack thereof) of proposals to the S-STEM program. We began with examining the feedback participants received from review panels on their proposal submissions. In this case study, we compare panel summaries for five S-STEM proposals submitted from five different institutions, exploring the similarities and differences in the overall reviews, as well as the strengths and weaknesses cited for both awarded and declined proposals that were awarded and declined in the context of their alignment with NSF's merit review criteria. This is submitted for consideration as a traditional paper presentation.

Introduction

Background

Through its Scholarships in Science, Technology, Engineering and Mathematics (S-STEM) program, the National Science Foundation (NSF) awards grants to institutions of higher education to support students from low-income backgrounds who have the academic talent and potential to succeed in STEM academic programs and ultimately, enter the STEM workforce or graduate programs. S-STEM funding supports scholarships of up to $10,000 per year per student, depending on demonstrated financial need. In addition, effective with the fiscal year 2016 (FY16) funding cycle, projects must include plans to implement and study evidence-based strategies to foster student success as articulated by S-STEM program goals, in order to generate knowledge that contributes to a greater understanding of how those strategies work among various educational settings, institutional contexts, and other parameters for students who meet S-STEM program requirements [1].

Rice University received funding from NSF in 2017 to host a series of workshops to help faculty members at predominantly undergraduate institutions (PUIs), with emphasis on those located in Established Program to Stimulate Competitive Research (EPSCoR) jurisdictions, to develop competitive proposals to the S-STEM program. To date, we have hosted three workshops, annually in 2017, 2018, and 2019. At the time of proposal submission, there had been no nationwide efforts that involved onsite proposal development activities focused on improving competitiveness of S-STEM proposals from institutions located in EPSCoR jurisdictions and/or PUIs, especially since the FY16 program changes.

The workshop presenters addressed the following aspects of S-STEM proposals:

- assuring they identified and addressed institutional and/or program needs that aligned with S-STEM program goals;
- adopting/adapting evidence-based strategies to address those needs;
- distinguishing between research and evaluation and developing a plan for investigating questions that would generate knowledge from the study of the strategies implemented;
- locating and working with a qualified evaluator to provide formative and summative feedback; and
- addressing STEM workforce development by building effective, mutually beneficial academic-industry partnerships [2].

Participants were asked to collect baseline data from prior S-STEM (if applicable) and/or other initiatives on campus to inform their proposals. As noted by Pearson, Crosby, et al, "... projects should seek to couch locally identified problems and needs in the larger context of educational research to help the broader STEM education community determine what interventions work best with scholars in their environments" [3]. Therefore, PIs were strongly encouraged to seek information on program and institutional challenges they could address and study as part of their S-STEM projects and share what they learned with the STEM education community. Before the workshop, participants were encouraged to speak with prior S-STEM PIs at their institutions, even if in other disciplines, to glean from their lessons learned (whether positive or negative) and utilize that information in developing their projects. During the workshop, they were introduced to tools such as logic models to help with project planning and to initiate conversations as they met with prospective project evaluators.

One of the novel elements of our project was that it not only provided training, but also included a research component designed to generate knowledge about the participants' experiences in developing and submitting S-STEM proposals in order to illuminate challenges they experienced, their responsiveness to merit review criteria, and how partnerships enhanced their development. We studied this, in part, by examining panel summaries from both successful and unsuccessful proposals to determine the relative strengths and weaknesses cited by reviewers.

Purpose

The purpose of this paper is multifaceted. First, it is intended to inform prospective S-STEM principal investigators (PIs) of some of the strengths and weaknesses that are common among proposals that were awarded as well as those that were declined. This will guide them in things to consider when developing proposals in the future. Second, it is intended to provide NSF with insights that will help them with outreach to the PI community and with preparing panelists to conduct reviews. Finally, this initial analysis of data from a single workshop cohort will help us frame a larger scale, comprehensive study of outcomes for all three cohorts of participants.

In this study we conducted a coarse, preliminary analysis to answer the following questions: (1) What are the most commonly noted strengths and weaknesses of S-STEM proposals by review panels? (2) How do the relative strengths and weaknesses of awarded and declined proposals compare? To do this, we performed a case study of the panel summaries from the reviews of proposals submitted by 2017 workshop participants. Specifically, we systematically reviewed comments received on panel summaries for five proposals (two awarded and three declined, each from a different institution) to understand the differences in the nature and number of comments associated with both types of proposals. This qualitative approach is intended to inform the direction of future research rather than to provide conclusive evidence of trends in reviewer comments [4].

Methods

Data Collection

The workshops supported two-person teams, comprised of the PI (who, in accordance with S-STEM program requirements was a faculty member teaching in an S-STEM eligible discipline) and a team member with expertise in educational or related areas of research (referred to in this project as researcher participants). A total of 21 teams (42 participants) comprised the 2017 cohort. We administered a follow-up survey to all participants one year after the March 2017 submission date to request information on the status of their proposals, the degree to which they benefitted from the workshop, and the workshop elements they applied in developing their proposals. Regarding proposal status, we asked if it was awarded, declined, pending and in negotiations with the program officer, or if it was not submitted. At least one member of 19 of 21 teams responded to the survey, giving a 90.5% response rate based on team representation (not individual participants). Of the 19 teams, 12 (63.2%) indicated they submitted a proposal in 2017. Table 1 summarizes the statuses of those submissions as of the date of the follow-up survey.

Table 1. Summary of 2017 Cohort Proposal Status after One Year

Proposal Status	Quantity	Percent
Awarded	2	10.5%
Declined	5	26.3%
In Negotiations	3	15.9%
No Response	2	10.5%
Not Submitted	7	36.8%
TOTAL	**19**	**100.0%**

In the survey, we asked respondents to copy and paste text from their panel summaries under five separate headings:

- Intellectual Merit Strengths
- Intellectual Merit Weaknesses
- Broader Impacts Strengths
- Broader Impacts Weaknesses
- Summary Statement

We also asked them to provide counts of each individual rating the proposal received (i.e., excellent, very good, good, fair, and poor). We converted those ratings to numeric scores (excellent = 5, very good = 4, good = 3, fair = 2, poor = 1) and computed average numeric scores for each proposal used in this analysis.

Of the twelve teams who indicated they submitted a proposal in 2017, a total of five had records with all of the aforementioned data complete. Three were for proposals that were declined; two were for proposals that were awarded. We used those five projects in this study. Table 2 provides descriptions of the awardee institutions' characteristics to add context to our analysis.

Table 2. Characteristics of Awardee Institutions Used in Case Study

Institution ID	Basic Carnegie Classification and Control	EPSCoR?	Status
INST50	Master's Colleges & Universities: Larger Programs Public Institution	N	Awarded
INST60	Doctoral/Professional Universities Public Institution	N	Declined
INST11	Master's Colleges & Universities: Medium Programs Public Institution	Y	Declined
INST21	Master's Colleges & Universities: Larger Programs Public University	N	Declined
INST51	Master's Colleges & Universities: Larger Programs Public University	N	Awarded

Data Analysis

NSF uses two merit review criteria to evaluate proposals submitted to all of its funding opportunities; intellectual merit refers to a project's potential to advance knowledge and broader impacts refers to its potential to benefit society. The Proposal & Award Policies and Procedures Guide (PAPPG) presents five questions that program officers and reviewers consider in evaluating proposals relative to the two merit review criteria [5], [6]:

1. What is the potential for the proposed activity to:
    a. advance knowledge and understanding within its own field or across different fields (Intellectual Merit); and
    b. benefit society or advance desired societal outcomes (Broader Impacts)?
2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?
3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?
4. How well qualified is the individual, team, or organization to conduct the proposed activities?
5. Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?

Question 1a specifically addresses intellectual merit and 1b addresses broader impacts; questions 2 through 5 pertain to both criteria.

Content from the five completed panel summaries was analyzed via deductive coding, where codes were derived from the merit review questions above. Table 3a through Table 7 summarize data relative to each of the merit review questions: Table 3a and 3b cover intellectual merit and broader impacts, respectively; Table 4 covers question 2 (creative, original, or potentially transformative concepts); Table 5 covers question 3 (well-reasoned, well-organized, sound plan); Table 6 covers question 4 (team qualifications); and Table 7 covers question 5 (adequate resources). For example, comments describing proposed activities or a proposal's foundational concepts as "unique" or "interesting" were coded as pertaining to question 2 and compiled in Table 4. Remarks on the way proposed activities were to be carried out, including the way they were organized and assessed, or the underlying reasoning or rationale supporting those activities, were coded to question 3 and compiled in Table 6. Each summary was initially coded by two researchers, and each coding disagreement resolved through discussion.

All content in each summary was analyzed for connection to one of the merit review questions. Content not found to be relevant to any of the five questions was excluded from the tables, as was any significantly detailed or identifying descriptions of institutions or proposal activities. Excerpts within the following tables were edited for clarity and brevity, with program-specific information and institution names omitted.

Results and Discussion

Tables 3a and 3b contain excerpts from panel summaries of both awarded and declined proposals, delineating strengths and weaknesses relative to both intellectual merit (Table 3a) and broader impacts (Table 3b). Due to the nature of the S-STEM program, in particular, its emphasis on STEM workforce development, it is difficult to differentiate between some aspects of intellectual merit and broader impacts. In the tables that follow, we assigned comments based on the headings to which they were ascribed in the panel summaries. The numbers in parentheses indicate the average numeric scores for the proposals' ratings. The two awarded proposals had scores of 5.0 (four "excellents") and 4.3 (one "excellent" and two "very goods"); the declined proposals had scores of 4.0 (one "excellent", two "very goods", one "good"), 3.7 (two "very goods", one "good"), and 2.5 (two "goods", two "fairs").

We acknowledge limitations in our analysis due to a few factors. First, our sample size is small ($N_{total} = 5$; $N_{awarded} = 2$; $N_{declined} = 3$). Second, we recognize that each panel is different, and consistency among panels is difficult, if not impossible, to achieve and to measure. Further, we did not have access to all the information on which panelists and program officers based their evaluations, so we could not determine the rationale for why some proposals were recommended for award and others for decline, nor was that our aim. Thus, the results presented here are not intended to be generalizable conclusions; rather, they are intended to form the basis of future research and to provide general insights for prospective PIs, reviewers, and NSF program officers to consider in their work.

Intellectual Merit: Not surprisingly, Table 3a shows many more comments related to intellectual merit strengths for awarded versus declined proposals. Building on lessons learned from past projects was a common strength for both. Additional strengths for awarded proposals were generally noted in areas related to knowledge generation (well-defined research questions, appropriate sample sizes, and application of evidence-based strategies) and in the rationale for the project as substantiated by institutional data (disparities in retention rates). Reviewers noted weaknesses in both declined and awarded proposals, some related to uncertainty about how the proposed project would be distinct from existing projects and how specific investigations would be conducted (declined proposals). The weakness in the awarded proposal (see bottom left table entry) points to potential disconnects between the program elements and the attrition points to be addressed by the project.

Table 3a. Summary of Strengths and Weaknesses Relative to Intellectual Merit

	What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	"The program builds on past success and lessons learned." (5.0) "… this proposal has two well-defined research questions and uses mixed methods" (5.0) "Power analysis demonstrates the N is large enough to detect an effect" (5.0) "The proposal makes a clear case for the focus in the computer science/mathematics area given the retention rates of Pell-eligible and non-Pell eligible students in computer science and mathematics appear significant." (5.0) "This proposal will provide course-based undergraduate research experiences … in the students' first year" (4.3) "They use the [framework] as a model to build upon." (4.3)	"The success and lessons learned from a previous funded NSF project … allowed the PIs to incorporate improvements into the new [program]." (3.7)
WEAKNESSES	"Much of the proposal … seems predicated upon the idea that students need additional support in the first [one or two years]. However, … underrepresented minority students generally achieve rates comparable to all students in terms of first and second year retention, but have substantially lower four-year and six-year graduation rates." (4.3)	"It was not apparent how the [project] is different from the [previous project] because similar terminology was used to describe the rationale and goals of both projects." (3.7) "How will intersectionality be studied?" (4.0)

Broader Impacts: Upon first glance, Table 3b shows a fairly balanced reporting of strengths and weaknesses for broader impacts in terms of the number of comments. However, some of the strengths of declined proposals actually read as weaknesses (e.g., dissemination approach was not clear), or were neutral in tone (e.g., the approach was "standard"). By contrast, reviewers tended to note unique aspects of dissemination approaches (e.g., will reach out to others) and broader populations (e.g., students, community, other researchers) that would be affected by the research in awarded proposals. Specifically, for the awarded proposals, broader impact strengths were primarily in the areas of the project's reach beyond the program and/or institution involved, with others related to overcoming disparities to improve the success of students from underrepresented groups and using career mentoring to enhance academic-industry partnerships. Industry partnerships were noted as particularly strong in one of the declined proposals as well. Though the declined proposals had some strengths that were, in general, of the same nature as the awarded proposals, they were not always communicated as strongly. For example, the panel summary in an awarded proposal states, "The proposal aims to close the gap and increase the retention of women and underrepresented students in STEM" whereas nearly the same strength is cited in a declined proposal as, "The broader impacts statement of this proposal is standard (increase student participation, improve retention, etc.)." The difference in overall review scores for these two proposals (5.0 versus 2.5) may indicate why the latter was stated much less enthusiastically. We reiterate the aforementioned limitation that because each panel is different and it is unlikely the summaries we analyzed were generated by the same panel, it is not possible to make direct comparisons here.

Table 3b. Summary of Strengths and Weaknesses Relative to Broader Impacts

	What is the potential for the proposed activity to benefit society or advance desired societal outcomes?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	"The proposal aims to close the gap and increase the retention of women and underrepresented students in STEM." (5.0) "… the partnerships between industry and academia will be enhanced by using the early career professional mentor model." (5.0) "The strong educational methods will likely lead to dissemination of results in several STEM education publications and conferences." (3.0) "The PIs also plan to reach out to several others in the region to share the results." (5.0) "… with benefits not only to scholars, but also to other students at the institution." (4.3) "Projects … are beneficial for both the students and the community." (4.3)	"The university and PIs have strong partnerships with industry" (3.7) "The broader impacts statement of this proposal is standard (increase student participation, improve retention, etc.)" (2.5) "… the more significant broader impact as to whether or not the [approach] will be successful and how results would be disseminated were not clear." (2.5) "This is a minority-dominated institution, so the program will not serve the underrepresented population at the college, but will serve the STEM career field well by increasing diversity." (4.0)
WEAKNESSES	"The proposers may want to consider if there are unintended implications of using phrases about race when the program is open to all students who meet the criteria." (4.3)	"The broader impacts may be limited by the proposal's potential to advance knowledge." (2.5)

The remaining tables summarize panel comments that we found to address the remaining four merit review questions. As stated previously, the expectation is that these questions be considered for both intellectual merit and broader impacts, thus there are elements of both that appear in Tables 4 through 7 that did not appear in Tables 3a and 3b.

Table 4 includes comments that were made regarding the creative, original, or potentially transformative aspects of the proposed activities. Interestingly, comments about proposal creativity, originality, and potentially transformative concepts were limited to strengths (i.e., the lack of creativity was not cited as a weakness), and they were only present for declined proposals (i.e., no summaries for awarded proposals cited creativity as a strength). As indicated in the S-STEM solicitation, projects are expected to contribute to the knowledge base on strategies to recruit, retain, and graduate students from low-income backgrounds and have them enter the STEM workforce or graduate school, and creative, potentially transformative ideas should be inherent in the generation of knowledge from these projects. However, because the solicitation is very prescriptive and focused on implementing and studying evidence-based strategies, PIs may feel limited to a large degree on being able to present novel ideas. And it is possible that strengths are not evident for this category in awarded proposals because those that are the strongest and most competitive adhere more strictly to the prescriptive adoption/adaptation of evidence-based practices without addressing nuances that could potentially be a novel contribution to the knowledge base.

Table 4. Summary of Strengths and Weaknesses Relative to Merit Review Question 2

	To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	None Coded	"Use of the software developed by the PI under the [grant] … as the computer-based test for the student selection is an interesting idea." (2.5) "Interesting combination of [disciplines]." (2.5) "One of the improvements was the design of two levels of scholarships … [giving] motivated students the ability to become part of the cohort and to become academically eligible for a full scholarship." (3.7) "The out-of-classroom learning experiences are unique and a great connection with their mentors and peers." (4.0) "The [program] is promising as a way to bring other students into STEM, perhaps even beyond the ones receiving the scholarship within the [program]." (4.0)
WEAKNESSES	None Coded	None Coded

Tables 5a and 5b show the comments about the strengths and weaknesses of the proposed project plan, respectively. Reviewers found many positive aspects of awarded and declined proposals relative to the project plans. Notably, cohort building and logic models were common strengths for both groups. Panel summaries of declined proposals featured very strongly positive comments that did not appear in summaries for awarded proposals (e.g., job placement and graduate school as metrics of success, extensive assessment, and mentoring by faculty, peers and industry representatives). By contrast, reviewers tended to focus on weaknesses of proposed plans for declined proposals; there were no items stated in weaknesses for awarded proposals that we coded in this category. Interestingly, in some cases, the summaries included statements of weaknesses for items that were presented as strengths which, in some cases, might confuse PIs. Two examples from the proposal with an average score of 3.7 are provided below:

- Strength: "The program includes strong and diverse engagement and cohort building that involves faculty, peer and industry mentoring and numerous campus and community activities." Weakness: "The process in which faculty, peer, and professional mentors would be chosen and trained … was not provided."
- Strength: "The assessment plan is extensive". Weaknesses: "[The] engagement assessment instrument has not been developed. It does not seem feasible that this instrument can be developed, tested, and validated before using it to assess the impact … on students." "There were no metrics in place to measure the success of the general public's awareness of STEM and NSF-funded scholarships." "… there were no proposed longitudinal studies tracking career placement and graduate school enrollment…"

Table 5a. Summary of Strengths Relative to Merit Review Question 3

	Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	"Notably, they have considered how to ensure that the projects are appropriate for the skill level of first-year students." (4.3) "Cohort formation is done well both in and out of the classroom" (4.3) "… there is also a well-articulated logic model that outlines and aligns the inputs, activities, outputs, and outcomes of the project." (5.0) "The panel agreed there is exceptionally strong cohort building in Year 1" (5.0)	"[The] plan is very long, but is very supportive and looks to do great things." (4.0) "It's positive that success is measured by looking at both matriculation into graduate programs or job placement." (4.0) "The logic model laid out is very clear." (4.0) "The program includes strong and diverse engagement and cohort building that involves faculty, peer and industry mentoring and numerous campus and community activities." (3.7) "The assessment plan is extensive" (3.7) "[The] program contained a strong engagement and cohort model that will be sustainable beyond the funding period." (3.7)

In the case of the first statement, the strength reads as a broad statement that is complimentary of the overall plans for mentoring and cohort building, followed by a weakness that points out details that would further strengthen the proposal. The strength statement would be clearer if it included a brief explanation of what the reviewers found to be strong. The second examples seems a bit more contradictory. Stating the assessment plan is extensive conveys a sense of its comprehensiveness/completeness; however, pointing out numerous elements that are missing – instrument(s), metrics, tracking progress into STEM careers or graduate school – leaves major questions about the elements that were actually strong. We also noted that multiple strengths for the proposal with an average numeric score of 2.5 were included by reviewers in the panel summary as weaknesses (Table 5b).

Table 5b Summary of Weaknesses Relative to Merit Review Question 3

	Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	
	Awarded Proposals	**Declined Proposals**
WEAKNESSES	None Coded	"Many of the activities present in this proposal are likely to help students." (2.5) "The proposal clearly makes the case that many of their students have unmet financial need, and it is likely that giving these students financial aid would help them progress through school." (2.5) "There is a reference to having a certain number of minority students, but that cannot be a selection criterion." (4.0) "The scholarships are going to be awarded over 3.5 years for a 4 to 5 year degree program. There is concern over whether this will fully support those students." (4.0) "The process in which faculty, peer, and professional mentors would be chosen and trained … was not provided." (3.7) "[The] engagement assessment instrument has not been developed. It does not seem feasible that this instrument can be developed, tested, and validated before using it to assess the impact … on students." (3.7) "The extent of the industry mentoring and internships was not fully explained." (3.7) "There were no metrics in place to measure the success of the general public's awareness of STEM and NSF-funded scholarships." (3.7) "… there were no proposed longitudinal studies tracking career placement and graduate school enrollment…" (3.7) "This proposal also seems fairly ambitious in this institutional context." (2.5) "Although the programming will only begin in the second year of funding, even a year seems too short to accomplish all these tasks." (2.5)

Table 6 shows comments relative to the teams assembled to complete the research. Notably, only strengths were mentioned in these reviews, but the strength listed for a declined proposal (e.g., team is large and members have roles within that team) is a relatively neutral comment. This harkens to research on letters of recommendation that points out that letter writers are likely to provide faint praise over negative information about others in recommendation letters [7].

Table 6. Summary of Strengths and Weaknesses Relative to Merit Review Question 4

	How well qualified is the individual, team, or organization to conduct the proposed activities?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	"…this proposal … uses mixed methods by a highly-qualified educational researcher" (5.0) "The panel believes the project management team members are very strong and have clearly defined roles." (5.0) "The panel believes this will lead to a robust and diverse team to implement the project" (5.0) "The team is generally strong and meets the S-STEM solicitation." (4.3)	"The team is large, but all of the members appear to have roles within the project." (4.0)
WEAKNESSES	None Coded	None Coded

Table 7 provides reviews of the institutional resources in the proposals. Unlike the reviews of teams, reviewers were more likely to focus on weaknesses related to resources for both declined and awarded proposals. Generally, weaknesses were related to omitted information rather than a critique of the resources described in the proposals.

Table 7. Summary of Strengths and Weaknesses Relative to Merit Review Question 5

	Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	
	Awarded Proposals	**Declined Proposals**
STRENGTHS	None Coded	None Coded
WEAKNESSES	"Details about involvement by the Financial Aid and Admissions staff are lacking and should be planned within the institution." (4.3) "... [faculty] have agreed to serve as mentors but there is no budget for their participation." (5.0)	"… it was not apparent how the [scholarship] conversion was budgeted" (3.7)

Conclusion and Recommendations

Our review of panel summaries of declined and awarded proposals to the NSF S-STEM program suggests some similarities among the two categories, and some differences. For intellectual merit and for teams, reviewers tended to be much more effusive about the benefits of awarded proposals than declined proposals. For broader impacts, about the same number of comments were received for awarded and declined proposals, but the tenor of the positive comments was much different; that is, comments in the strengths category for declined reviews tended to be somewhat negative or neutral in tone. Comments about the approach and project plan also tended to differentiate awarded from declined proposals; although positive aspects of both awarded and declined proposals were noted, reviewers pointed out a number of disadvantages for declined proposals (and none for awarded proposals). While for many NSF programs, reviewers tend to more strongly emphasize intellectual merit than broader impacts, indicating a better understanding of or connection to that criterion [8], the panel summaries we examined attended more strongly to broader impacts. We believe this is due, in part, to the nature of the S-STEM program and its emphasis on improving outcomes for students from low-income backgrounds. Inherent in the solicitation is a degree of "messiness" that, in some instances, blurs the lines between intellectual merit and broader impacts. We also note that the proposals that were part of the 2017 cohort's submissions were reviewed when the knowledge generation requirement was still new; neither program officers nor reviewers had become well-acquainted with what this aspect of the proposals should reflect. All of these factors likely contribute to difficulties the reviewers showed in properly delineating strengths and weaknesses under the appropriate merit review criteria. And although our findings point to the need for more research, this study clearly indicates that more useful feedback for research teams would be helpful, particularly in the area of broader impacts. Specifically, the feedback provided to research teams for broader impacts did not seem to vary depending on whether the proposal was awarded or declined.

Interestingly, only panel summaries for declined proposals contained strengths for creative, original, and transformative ideas; no panel summaries of awarded proposals identified strengths in this area. Additionally, in some cases, awarded and declined proposals shared very similar, if not the same, weaknesses. This raises questions for proposals that are declined with relatively high scores (i.e., 4.0 or above). Specifically, we recommend that NSF program officers explore more deeply how decisions are made with regards to making recommendations for awards or declines for proposals of similar quality with similar strengths and weaknesses.

Our future research will more systematically analyze panel summaries for all three cohorts of workshop participants (2017, 2018, and 2019), which will enable us to further examine the preliminary findings from this paper. We hope this work prompts additional research on questions beyond the nature and scope of this analysis. One recommendation is to explore how knowledge generation from the adoption/adaptation and study of evidence-based practices can be creative and potentially transformative.

Acknowledgment

Support for this work was provided by the National Science Foundation under award number 1708329. The opinions, findings, and conclusions or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] National Science Foundation, “NSF Scholarships in Science, Technology, Engineering, and Mathematics (S-STEM) Program Solicitation,” Arlington, VA, NSF 17-527, 2016. [Online] Available: https://www.nsf.gov/pubs/2017/nsf17527/nsf17527.htm

[2] Pearson, Y.E. and Phillips, C.M.L. and Beier, M. and Gilberto, J. and Mattingly, S.P. and Saterbak, A. and Sheng, Y. and Shethia, A.K. and Sun, R., "Evaluating the Quality of Project Summaries for S-STEM Proposals," 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, 2018, June. ASEE Conferences, 2018. https://peer.asee.org/30448

[3] Y.E. Pearson, K.E. Crosby, E.R. Blevins, B.R. Isbell, A.P. Kruzic, S.P. Mattingly, S. P., L.L. Peterson, and J.C. Tiernan, “Challenges, Opportunities, and Impacts of S-STEM Projects: Insights for Institutional Capacity Building at Minority-Serving Institutions” Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. 10.18260/p.26472.

[4] J.W. Creswell, *Qualitative Inquiry & Research Design: Choosing among Five Approaches*, 3rd ed. Thousand Oaks: SAGE, 2013.

[5] National Science Foundation, “Proposal & Award Policies & Procedures Guide,” Arlington, VA, NSF 17-1, 2017. [Online] Available: https://www.nsf.gov/pubs/policydocs/pappg17_1/index.jsp

[6] National Science Foundation, “Proposal & Award Policies & Procedures Guide,” Arlington, VA, NSF 20-1, 2020. [Online] Available: https://nsf.gov/pubs/policydocs/pappg20_1/index.jsp

[7] J.M. Madera, M.R. Hebl, and R.C. Martin, (2009), “Gender and Letters of Recommendation for Academia: Agentic and Communal Differences,” *J. App. Psych. 94*(6), 1591–1599. https://doi.org/10.1037/a0016539

[8] D. Verdín, “Quantifying and Assessing Trends on the National Science Foundation's Broader Impact Criterion,” paper presented at 2017 ASEE Annual Conference & Exposition, Columbus, OH. https://peer.asee.org/28778