CORE QUESTIONS and REPORT TEMPLATE
for
FY 2020 NSF COMMITTEE OF VISITOR (COV) REVIEWS

Guidance to NSF Staff: This document includes the FY 2020 set of Core Questions and the COV Report Template for use by NSF staff when preparing and conducting COVs during FY 2020. Specific guidance for NSF staff describing the COV review process is described in the "COV Reviews" section of NSF's Administrative Policies and Procedures which can be obtained at https://inside.nsf.gov/tools/toolsdocuments/Inside%20NSF%20Documents/Policy,%20Procedures,%20Roles%20and%20Responsibilities%20for%20COV%20Reviews%20and%20Program%20Portfolio%20Reviews.pdf[1].

NSF relies on the judgment of external experts to maintain high standards of program management, to provide advice for continuous improvement of NSF performance, and to ensure openness to the research and education community served by the Foundation. COV reviews provide NSF with external expert judgments in two areas: (1) assessments of the quality and integrity of program operations; and (2) program-level technical and managerial matters pertaining to proposal decisions.

The program(s) under review may include several sub-activities as well as NSF-wide activities. The directorate or division may instruct the COV to provide answers addressing a cluster or group of programs – a portfolio of activities integrated as a whole – or to provide answers specific to the sub-activities of the program, with the latter requiring more time but providing more detailed information.

The Division or Directorate may add questions relevant to the activities under review. Copies of the report template and the charge to the COV should be provided to OIA prior to forwarding to the COV. In order to provide COV members adequate time to read and consider the COV materials, including proposal jackets, COV members should be given access to the materials in the eJacket COV module approximately four weeks before the scheduled face-to-face meeting of the COV members. Before providing access to jackets, the Conflict of Interest and Confidentiality briefing for COV members should be conducted by webinar, during which, NSF staff should also summarize the scope of the program(s) under review and answer COV questions about the template.

Suggested sources of information for COVs to consider are provided for each item. As indicated, a resource for NSF staff preparing data for COVs is the Enterprise Information System (EIS) –Web COV module, which can be accessed by NSF staff only at http://budg-eis-01/eisportal/default.aspx. In addition, NSF staff preparing for the COV should consider other sources of information, as appropriate for the programs under review.

For programs using section IV (addressing portfolio balance), the program should provide the COV with a statement of the program's portfolio goals and ask specific questions about the program under review. Some suggestions regarding portfolio dimensions are given on the template. These suggestions will not be appropriate for all programs.

Guidance to the COV: The COV report should provide a balanced assessment of NSF's performance in the integrity and efficiency of the ***processes*** related to proposal review. Discussions leading to answers of the Core Questions will require study of confidential material such as declined proposals and reviewer comments. ***COV reports should not contain confidential material or specific information about declined proposals.*** The reports generated by COVs are made available to the public.

We encourage COV members to provide comments to NSF on how to improve in all areas, as well as suggestions for the COV process, format, and questions. For past COV reports, please see http://www.nsf.gov/od/oia/activities/cov/.

[1] This document has three parts: (1) Policy, (2) Procedures, and (3) Roles & Responsibilities.

FY 2020 REPORT TEMPLATE FOR
NSF COMMITTEES OF VISITORS (COVs)

The table below should be completed by program staff.

Date of COV: September 22-23, 2020

Program/Cluster/Section: All

Division: Engineering Education and Centers (EEC)

Directorate: Engineering

Number of actions reviewed: 175

Awards: 80 awards (AWD)

Declinations: 75 full proposal declines (DECL)

Other: 15 Return Without Review (RTNR) proposals and 5 not invited pre-proposals (NIVT)

Total number of actions within Program/Cluster/Division during period under review: 3013

Awards: 846 AWDs

Declinations: 1758 DECLs

Other: 63 invited pre-proposals (INVT), 312 NIVT, 26 RTNR, and 8 withdrawn (WTH) proposals

Manner in which reviewed actions were selected: Stratified Random Sample
Jackets were randomly selected to include the desired distribution of awards, declinations, and returned without review proposals within each program within each cluster across the four fiscal years under review. The Centers subcommittee also received NIVT pre-proposals.

Each COV member in the Workforce Development/Broadening Participation in Engineering (WFD/BPE) and Engineering Education subcommittees received 15 jackets and each COV member in the Centers subcommittee received 5 jackets due to the size and nature of the Centers proposals versus other proposals. The total number of jackets reviewed was 175, 5.8% of the 3013 actions in FY 2016- FY 2019. The number of proposals in each program was chosen to best reflect the overall portfolio while ensuring that the COV received proposals from across all programs. The tables below show the number of proposals assigned to COV members for each of the programs in each cluster.

WD/BPE			
Program	AWD	DECL	RTNR
BP	2	2	1*
INCLUDES	1	1	
RET	1	1	
REU	3	3	

Eng Ed **			
Program	AWD	DECL	RTNR
CAREER	1	1	1*
Other Eng Ed	2	2	
RIEF	1	1	
RED	1	1	
RFE	2	2	

*Due to the small overall number of RTNRs, the RTNRs from all programs within a subcommittee were combined and each COV member received 1, resulting in a variety of programs represented across a subcommittees sample of RTNRs.

** One COV member from this subcommittee could no longer participate, so jackets originally assigned to this individual were evenly redistributed to the remaining members to review.

Centers ***	
AWD ERC or ERC-related full proposal	2
DECL ERC or ERC-related full proposal or INVT ERC pre-proposal (not AWD)	1
NIVT ERC pre-proposal	1
RTNR ERC or ERC-related full proposal	1

*** One COV member from this subcommittee could no longer participate, and due to time constraints, jackets were unable to be redistributed for review. This subcommittee was only able to review 20 out of the 25 jackets originally assigned due to these events.

COV Membership

	Name	Affiliation
COV Chair or Co-Chairs:	Lance C. Pérez	University of Nebraska-Lincoln and NSF ENG Advisory Committee Representative
	Darryll J. Pines	University of Maryland and NSF ENG Advisory Committee Representative
COV Members:	Mary Besterfield-Sacre	University of Pittsburgh
	Adam Carberry	Arizona State University
	Ning Fang	Utah State University
	Stephanie Farrell	Rowan University
	Bonnie Ferri	Georgia Tech
	George Johnson	University of California, Berkeley
	Michelle Jordan	Arizona State University
	Vikram Kapila	New York University
	Adam Kirn*	University of Nevada-Reno
	Dimitris Lagoudas	Texas A&M University
	Susan Lord	University of San Diego
	Christopher Rahn	Pennsylvania State University
	Sonya Smith	Howard University
	Susan Smyth*	SME Board of Directors
	Alice White	Boston University
	*Indicates persons who were originally part of the COV but had to withdraw due to extenuating circumstances.	

MERIT REVIEW CRITERIA

An understanding of NSF's merit review criteria is important in order to answer some of the questions on the template. Reproduced below is the information provided to proposers in the Grant Proposal Guide about the merit review criteria and the principles associated with them. Also included is a description of some examples of broader impacts, provided by the National Science Board

1. Merit Review Principles

These principles are to be given due diligence by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether or not to recommend proposals for funding and while overseeing awards. Given that NSF is the primary federal agency charged with nurturing and supporting excellence in basic research and education, the following three principles apply:

- All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.
- NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.
- Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.

With respect to the third principle, even if assessment of Broader Impacts outcomes for particular projects is done at an aggregated level, PIs are expected to be accountable for carrying out the activities described in the funded project. Thus, individual projects should include clearly stated goals, specific descriptions of the activities that the PI intends to do, and a plan in place to document the outputs of those activities. These three merit review principles provide the basis for the merit review criteria, as well as a context within which the users of the criteria can better understand their intent.

2. Merit Review Criteria

All NSF proposals are evaluated through use of two National Science Board approved merit review criteria. In some instances, however, NSF will employ additional criteria as required to highlight the specific objectives of certain programs and activities.

The two merit review criteria are listed below. Both criteria are to be given full consideration during the review and decision-making processes; each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers must fully address both criteria. (PAPPG Chapter II.C.2.d.(i) contains additional information for use by proposers in development of the Project Description section of the proposal.) Reviewers are strongly encouraged to review the criteria, including PAPPG Chapter II.C.2.d.(i), prior to the review of a proposal.

When evaluating NSF proposals, reviewers will be asked to consider what the proposers want to do, why they want to do it, how they plan to do it, how they will know if they succeed, and what benefits could accrue if the project is successful. These issues apply both to the technical aspects of the proposal and the way in which the project may make broader contributions. To that end, reviewers will be asked to evaluate all proposals against two criteria:

- **Intellectual Merit**: The Intellectual Merit criterion encompasses the potential to advance knowledge; and

- **Broader Impacts**: The Broader Impacts criterion encompasses the potential to benefit society and contribute to the achievement of specific, desired societal outcomes.

The following elements should be considered in the review for both criteria:

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?

3. Examples of Broader Impacts

The National Science Board described some examples of broader impacts of research, beyond the intrinsic importance of advancing knowledge.[2] "These outcomes include (but are not limited to) increased participation of women, persons with disabilities, and underrepresented minorities in science, technology, engineering, and mathematics (STEM); improved STEM education at all levels; increased public scientific literacy and public engagement with science and technology; improved well-being of individuals in society; development of a globally competitive STEM workforce; increased partnerships between academia, industry, and others; increased national security; increased economic competitiveness of the United States; and enhanced infrastructure for research and education. These examples of societally relevant outcomes should not be considered either comprehensive or prescriptive. Investigators may include appropriate outcomes not covered by these examples."

[2] NSB-MR-11-22

INTEGRITY AND EFFICIENCY OF THE PROGRAM'S PROCESSES AND MANAGEMENT

Briefly discuss and provide comments for *each* relevant aspect of the program's review process and management. Comments should be based on a review of proposal actions (awards, declinations, returns without review, and withdrawals) that were *completed within the past four fiscal years*. Provide comments for *each* program being reviewed and for those questions that are relevant to the program(s) under review. Quantitative information may be required for some questions. Constructive comments noting areas in need of improvement are encouraged.

I. Questions about the quality and effectiveness of the program's use of merit review process. Please answer the following questions about the effectiveness of the merit review process and provide comments or concerns in the space below the question**.**

QUALITY AND EFFECTIVENESS OF MERIT REVIEW PROCESS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE
1. Are the review methods (for example, panel, ad hoc, site visits) appropriate? Comments: The consensus across proposals submitted and reviewed in the EEC division is that the methods were appropriate for all program categories. Consistency between panel reviews and individual reviews should be maintained. ERCs: The ERC review process is comprehensive, lasting almost two years from solicitation release to award announcement. Preliminary proposals, full proposals, and site visits each require reviews from diverse stakeholders in both panel and site visit formats. At each step in the process, the proposing team can respond to the reviewer comments, improving the proposal and the eventual awarded centers. The finalists undergo extensive review through a blue-ribbon panel, engineering review board, and director's review board. The number of reviews for planning grants were not always consistent, i.e., at times there were more reviews for an approved planning grant than for those that were not approved. Criteria for triage was not clear, especially for the planning grants and preproposals. Transparency with respect to the guidelines for the blue-ribbon panel composition would be helpful. WFD/BPE: The review methods were appropriate and complete for the vast majority of eJackets in the portfolio. Most proposals were reviewed by panels, or in some cases through ad hoc review. In rare cases, a waiver of external review was granted, and these instances were deemed appropriate by the COV subcommittee. In these cases, the waiver was documented, and the proposal was reviewed internally. In one case, an external review was requested but not	YES

received; because of the importance of the proposed workshop a review waiver was granted. This was deemed appropriate by the COV subcommittee. Occasionally, a proposal was returned without review for reasons that were documented. ENGR Education: For the panel proposals (CAREER, RED, RIEF, RFE), the panel approach seemed to follow the process delineated in the information provided; and the methods (e.g., flow chart) are appropriate. In most cases, 4 individuals reviewed proposals. There was one jacket where there were 3 reviewers, and a 4th panelist was sought ad hoc. It might be good to provide in the Program Officer (PO) summary why a 4th panelist was sought (e.g., the first three reviews were inconsistent). There were questions about EAGER/RAPID programs, as the review process was not consistent across the various proposals. For example, some proposals were panel reviewed and others were reviewed by a single PO. In such cases, it was not always clearly delineated as to why a panel was sought or not. Further, in three instances proposals came in as EAGERs but were awarded as regular grants, but with no explanation as to why. We recommend that in cases like this, that the reason for the actual review process be documented in the summary, as well as how they were eventually awarded.	
2. Are both merit review criteria addressed a) In individual reviews? b) In panel summaries? c) In Program Officer review analyses? Comments: Both merit review criteria are addressed with some variation in individual reviews. ERC: The individual reviews, panel summaries, and PO review analyses, for the most part, do address both review criteria. With regards to the Engineering Research Centers proposal reviews, the individual reviews, panel summaries, and PO review analyses all addressed the merit review criteria. Individual reviews, in some cases, did not address the broader impacts of the proposal. In one case the lone female member of the panel was the only member to notice that the proposal lacked a clear vision towards an ERC and that the proposal made no mention of the inclusion of underrepresented minorities or women. Proposals that receive a positive recommendation should address both criteria. WFD/BPE: With regards to WFD/BPE proposal reviews, both criteria were called out and addressed in the individual reviews, panel summaries and review analyses. Panels offered another level of refinement over individual reviews, offering a synthesis of ideas, observations and perspectives. The Program Officer review analyses were different in	YES

style but always appropriate. They captured the panel summary well, and the PO took appropriate next steps in requesting additional information from the PI as needed for a competitive proposal. ENGR Education: Individual reviews addressed both merit criteria. With that said, there is often variation in the individual reviews, as some panelists go into considerable detail while others provide more minimalist statements. The panel summaries do pull together the individual reviews and address the merit criteria in a more complete manner. The PO summaries do address the merit criteria as well as the conversations of how the panel got to the merit criteria decision. Further, the PO summary indicated a rational for the decline. A particular proposal had six reviewers and two panel summaries, one by SES and one by EEC. The SES panel summary did not explicitly address these criteria while the EEC summary did. A recommendation for the future would be to provide more clarity of what IM and BI are and what the individual panelists should be providing in their review. This benefits individual panelists, but also facilitates the review process and feedback to the PI regardless if they are rewarded or declined. Further, it improves the consistency of the individual reviews and panel summaries. Note, these criteria often depend on the nature of the RFP.	

3. Do the individual reviewers giving written reviews provide substantive comments to explain their assessment of the proposals? Comments: In general, the reviewer comments for EEC proposals and panels were comprehensive and substantive. The nuances for Engineering Research Centers, Workforce Development/Broadening Participation, and Engineering Education proposals are summarized below: ERCs: Overall, the reviews were quite substantive and did a good job of explaining why the reviewers scored the proposals as they did. Some of the written reviews are comprehensive and well thought-out, but many, particular for the planning grants, were lacking details. The number of reviews for planning grants were inconsistent for those that were successful when compared to those that were not successful. WFD/BPE: In general, the reviewers' comments were substantive and consistent with their ranking of the intellectual merit and broader impacts of the proposal. A couple of exceptions are noted below. In general, comments were constructive in the sense that they would allow the team of investigators to improve the proposal and resubmit a more competitive proposal. Outliers: Occasionally the review lacked specific feedback to the team of investigators that would help them understand the rationale for the award decision or to improve the proposal for resubmission. At times, a reviewer would provide a "good" or "very good" rating without mentioning substantive weaknesses. In one case, an individual reviewer addressed a potential weakness in the proposal but provided attribution to only their own belief/opinion, without any additional evidence and without reference to specific information in the proposal. The same weakness was addressed more appropriately in the panel summary. ENGR Education: The substance of the reviews is a function of a reviewer's prior panel experience as well as their knowledge of the content area. This is to be expected and often explains the variation in reviews. Our recommendation from question 2, that of providing additional clarity regarding merit criteria for the solicitation, would provide improvements to neophyte reviewers as well as those who may have served on panels from different directorates or disciplinary areas.	YES
4. Do the panel summaries provide the rationale for the panel consensus (or reasons consensus was not reached)? Comments:	YES

In general, panel summaries did provide a rationale for panel consensus and noted when there was disagreement on the panel. The panel summaries tended to reflect discussion by panelists indicating strengths and shortcomings of a proposal. Again, there are some variations across the various EEC programs detailed below. ERC: The panel summaries do provide additional information and rationale from the group discussions and support the panel consensus. Some of the center and planning grants have what seems to be conflicting rationales. For example, in one case an ERC proposal did not receive an Excellent from any of the reviewers and did not follow the guidelines for the ERC solicitations but still was funded by the program. In this case the PIs had the opportunity to revise those portions of the proposal that were not compliant. In another case, a center was not funded under apparently similar conditions. The PO diary notes should better explain the decision and PO should explain the rationale to the proposal teams. WFD/BPE: The summaries did not explicitly provide rationale for consensus, but it can be inferred from the individual ratings and the panel recommendation that consensus was achieved through the synthesis of all of the reviews of the intellectual merit and broader impact. The Committee did not see reviews that were in conflict with each other, but reviewers had different observations based on their perspectives and expertise. The panel summaries and final ratings did seem consistent with the reviewer comments. Additional Notes: The wording of this question is odd. There is no place on the summary template to provide a “rationale” for consensus. More appropriate wording might be if the panel summary seems consistent with the review comments and final rating by the panel. All of the summaries did appear to reflect the panels’ consensus. The insights regarding the rationale for consensus were generally more apparent in the Review Analyses than in the Panel summaries. The information provided to the COV noted that panel summary templates were developed to improve consistency and better inform PIs of the panel discussion. Our subcommittee did not find a section in the Panel Summaries that discussed the rationale for consensus. Were the templates implemented? ENGR Education: By nature of the panel process, individuals state their reviews and the group comes together with discussion to create a unified summary. Where consensus is not met on all areas, panels included comments qualified by phrases such as “one reviewer did note…” or “some panelist thought…”. It is expected that consensus may not be possible on all details, so this phrasing is reasonable and gives more clarity on the level of agreement. Understanding team dynamics is important for the PO to have when putting panels together as well as facilitating the panels. Inclusivity of panelist opinions are important; and how the panel interacts as a group is a function of the outcome of the panel’s recommendation. PO’s should also include in their notes which panelist wrote the summary, as well as how many ratings were changed	

during the proposal review along with their rationale. These recommendations may require POs to have professional development in this area. Training should also be required of the reviewers or at best a set of ground rules for interaction.	
5. Does the documentation in the jacket provide the rationale for the award/decline decision? Comments: In general, the documentation did provide a rationale for the award/decline decision. The extent of the review analysis is highly variable. Again, there are some variations across the various EEC programs detailed below. ERC: The ERC review process is extensive, and each step is well documented and supported by the PO review analysis. For the most part, the rationale for the award/decline decision was well supported by the review documentation. The POs did a good job in providing the rationale for decisions and for the proposals reviewed, they aligned with the panel evaluations. In some cases, the jackets contain proposals that illustrate bad and good rationale for funding, including those such as rejected for too many references, or all PIs from same institution. One planning grant was rejected because the PI was from a CS department that was housed inside of the college of engineering. This situation could have been further clarified in the documentation. WFD/BPE: The documentation does provide a rationale that is consistent with the award/decline decision. Funded proposals consistently had higher reviewer ratings than those that were not funded. Notes: There was considerable variation in the format of the Program Officer review analyses. For example, some specifically had a section on COI and others did not. Some had bullets at the top giving relevant details such as review ratings and panel rating. Others had long statements where this information was embedded, and others did not include it. Some of the Review Analyses had typos such as the wrong institution being referred to. In some cases, copying and pasting led to awkward errors. The COV appreciates that program officers are responsible for a large number of jackets, so some typos are to be expected. However, perhaps it would be possible to “auto-populate” a Review Analysis form using information from the proposal, the Panel Summary and the Review Record (Form 7) with standard elements such as title, institution, review ratings, COIs, etc. This could both help minimize the burden on POs and lessen the chance of making some of the errors observed. It is worth emphasizing, however, that the Review Analyses were generally thorough in discussing strengths and weaknesses, as well as providing a solid rationale for the funding recommendation. ENGR Education:	YES

In general, decisions were consistent with the types of reviews at the various levels (i.e., individual, panel summary, PO summary notes). However, explicit rationales were not always provided in the PO summary notes. One program officer was particularly good at specifically acknowledging the rationale for award or decline decision in their summary notes, but this was not consistent across the various POs. One program, supplemental funding for the Graduate Preparedness Agency Priority Goal, did not have sufficient detail in the documentation to provide the rationale for the decision. There were no individual or panel reviews, just a program officer's analysis. In these cases, the response was the same for all the jackets and the rationale for the decision was not detailed. It is recommended that declined proposals receive more delineated specifics regarding the decision. POs should take additional time to share their additional knowledge to these proposals to enrich feedback. A second recommendation might be to create a PO summary template taking the best approaches from the various POs.	

6. Does the documentation to the PI provide the rationale for the award/decline decision? [Note: Documentation to PI usually includes context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), and, if not otherwise provided in the panel summary, an explanation from the program officer (written in the PO Comments field or emailed with a copy in the jacket, or telephoned with a diary note in the jacket) of the basis for a declination.] Comments: In general, the documentation provided to the PI provided the rationale for the award/decline decision. The extent of the PO comments is highly variable and the use of boilerplate language, though understandable, does reduce the amount of information given to some PIs when they are declined. Again, there are some variations across the various EEC programs detailed below. ERC: Overall, the documentation to the PI explained the rationale for the award/decline decision well. In one example, the decision of which ERC was awarded was made after the blue-ribbon panel met and was based on other priorities and considerations. WFD/BPE: For most proposals, the documentation to the PI was complete and clearly indicated the rationale for the ultimate award or decline decision. In most cases this is in the form of a statement that refers the investigators to the panel summary and the individual reviewer comments referring to the strengths and weaknesses of proposal merit. In many cases, Program Director's correspondence to the PI sought response to weaknesses identified by reviewers. In such cases PI response was included in the diary notes and appropriately utilized in funding decisions. The COV felt that additional insight from the PO to the PI would often be helpful for proposals that are declined, especially when highly rated. Outliers: For the occasional case in which the individual reviewers did not provide substantive weaknesses to justify low ratings, the investigators may not understand the rationale for the decision. Collaborative proposals should be sure to have sufficient information in each jacket. For example, one jacket (proposal X) was collaborative with another (proposal Y). This proposal was returned without review. No details are available in Project X's jacket. There is just one sentence that says the proposal did not respond to the solicitation for the REU. Very specific problems with the proposal are detailed in PO comments for Proposal Y. This could only be found by going to the review analysis that was not the pdf. The pdf version did not include a link to the collaborative proposal. This made it difficult to track down the rationale for the return.	YES

ENGR Education: This is a difficult question to address. Panelists and panel summaries cannot provide specific documentation to the PI regarding award/decline – in fact they do not provide award or decline decisions. The PO's review analysis stays in-house at the NSF is not a public document, yet it is often rich in the rationale. In some cases, what is sent to the PI is boilerplate language that does not provide the specific rationale for award/decline from the PO. However, the reviews and panel summaries are rich in detail. An outlier was the Graduate Preparedness Agency Priority Goal where the PIs of declined supplements did not get detailed comments. Instead, the response was not specific and was a generic template. It was a missed opportunity to explain what was lacking in the package and perhaps in the PIs understanding of what the intent of the program was and why it was an important issue. As a second recommendation, complete documentation should be provided between the PO and the PI. A few jackets were incomplete. As example, one jacket must have been sent back to a PI but there was no documentation to indicate this was done.	
7. Additional comments on the quality and effectiveness of the program's use of merit review process: Comments; In general, the COV was impressed with the quality and effectiveness of the division's use of the merit review process. Program specific comments are below. ERC: The use of templates/rubrics means that the reviewer is constantly reminded of the review criteria and this helps to "even out" the reviewing by multiple reviewers and panels. However, it is more difficult to make the case for the satisfaction of the merit review criteria on the planning grant applications. The potential for addressing the merit review criteria should be addressed for planning grants. WFD/BPE: In all cases, the proposals were deemed to have been reviewed quite well using the merit review criteria. The panelists, individual reviewers, and program officers effectively assessed the quality of proposals and used the resulting assessments in making award/decline decisions. EEC dwell times are within the 6-month goal for NSF which is good. The declines have a higher average number of reviews than the awarded proposals. This might suggest that more care is being taken to give a fair assessment of the proposals that are not funded. In panel reviews, each proposal typically received 3-4 reviews. One ad hoc review had 5 reviews consistent with this trend of more reviews on declined proposals.	YES

For some Workforce Development programs, Reviewers cited significant weaknesses that a proposal did not build on prior results. Recommendation: Include language in the REU and RET program solicitation to indicate that the proposal should address building on prior results if applicable. It should also describe the extent to which innovation/novelty is important in distinguishing from prior award, and the extent to which sustainability is important and whether there is a limit to how many times a program can be renewed. For proposals that are not competitive, consider a swift decline notification. This would allow PIs to pursue other opportunities. ENGR Education: The merit review process is complex. The terms intellectual merit and broader impacts are broad, allowing for many different interpretations. Increased structure for individual reviews would greatly improve a reviewer's ability to fully address the merit review criteria. This template should be shared broadly to help inform proposers, so the rubric used to assess a proposal is not vague. Guidance to panelists on how best to read proposals and address the merit criteria might be helpful. These five elements should be considered in both criteria, but they do not necessarily 'fit' for both the IM and BI; hence, it adds to the complexity of the review process.	

II. Questions concerning the selection of reviewers. Please answer the following questions about the selection of reviewers and provide comments or concerns in the space below the question**.**

SELECTION OF REVIEWERS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE
1. Did the program make use of reviewers having appropriate expertise and/or qualifications? Comments: In general, the COV found that the division made use of reviewers having appropriate expertise. The division should continue its efforts to diversify the reviewer pool and to gather additional data about reviewers. Program specific comments are below. ERC: Overall, the reviewers have appropriate expertise and qualifications. The use of industry reviewers is appropriate for ERC programs. Some may not be as familiar with the review process and requirements for intellectual merit and broader impacts. Not clear how industry reviewers are recruited, selected, and trained for the ERC review process. The COV members recognize the difficulty in finding available members of the scientific and engineer community to serve on panels. The efforts to bring in reviewers with diversity on all axes is a huge amount of work and laudable. However, the lack of diversity among the reviewers both in ethnic diversity and gender diversity was troubling. In addition, panel composition in which all the scientific participants were male, and the broadening participation participants were female (or the lone female participant) should be avoided. This can be especially surprising in disciplines such as chemical engineering where there are many women scientists and engineers and also black women and women of color. The stature of the reviewer seems well-matched to the scale of the award—ERCs have reviewers with significant stature. In some cases, additional subject matter experts would have helped in the review process. WFD/BPE: To the extent that the panelist expertise was evident from the information provided in the eJackets, the program used reviewers with appropriate qualifications. From the information provided in the eJacket, it was sometimes possible to ascertain department affiliation, but not expertise. The Committee requested CVs of the reviewers, and a review of a subset of those CVs showed that the reviewers were highly qualified to review within the program areas of Broadening Participation and Workforce Development. It was evident that an effort was made to achieve an overall panel composition that was appropriately tailored to the goals of the program. It was also noted that the panels comprised a cross section of engineers, scientists (including social scientists) and leaders from the research and education community, and there was a	YES

balance across sectors especially for the INCLUDES program. Most reviewers were from R1 institutions, but minority serving institutions and masters-level universities were also represented. The reviewer selection process sounds reasonable and multi-faceted to achieve a diverse pool of reviewers with various perspectives. Comments: Quite often, the information provided about the panelist in the eJacket was insufficient to ascertain their disciplinary affiliation. In some cases, the institutional affiliation was also missing. Outliers: One proposal did not seem to have social science expertise needed. Recommendations: Provide expertise as a separate category in the reviewer information. Expand the disciplinary classifications to include more meaningful data. The codes provided in the eJacket materials are much more general than the codes supported in PRIM. It would be helpful for the Committee to have this more detailed information in order to evaluate expertise. Additionally, PRIM should allow more than one choice (maybe up to three) for area of expertise. Some areas seem to be missing (such as social science), and some seem too broad (such as education). ENGR Education: In general, we were pleased with the backgrounds and expertise of the panelists that were brought together. Panels were created given various backgrounds. As example RED proposals had both disciplinary persons as well as engineering education specialists. Note that the information provided should have indicated where the expertise of the panelists could be found. Several members of this sub-committee Googled the panelists as we did not know where else to find this information.	
2. Did the program recognize and resolve conflicts of interest when appropriate? Comments: In general, the COV found that the division appropriately identified and resolved conflicts of interest. Program specific comments are below. ERC: NSF does an outstanding job with COI, especially in the ERC program where so many universities and companies are involved. Good idea to require ERC pre-proposals to not list companies to eliminate potential COI with industry reviewers and increase the reviewer pool. WFD/BPE: Different Program Officers seem to have different styles when it comes to documenting conflicts. Some explicitly address conflicts in the Context	YES

Statement or Review Analysis, indicating that panelists with a conflict left the room during the discussion and that the conflict was identified in advance so that the reviewer could not access the review in the system. In other cases, there was no explicit evidence of COI as mentioned by the Program Directors. The subcommittee noticed a lack of consistency in the identification of COI in the Review Analyses. For example, one jacket had a panelist with a COI (same institution) but this is not called out in review analysis. Another jacket lists two COI in the panel overview, but review analysis only says that there was one. ENGR Education: By in large COI was handled correctly; however, in a few cases the PO's notes failed to mention that the COI individuals left the room during discussion and deliberations. This might be a minor point, but it provides that 'safety' check for the jacket that the process was conducted properly.	
3. Additional comments on reviewer selection: ERC: COI can reduce the available pool of potential reviewers so that the resulting panels may not be as effective. The committee recommends that reviewers from industry should be sought out. A better geographical distribution of reviewers is recommended, especially to represent topics that are important for certain regions of the country. WFD/BPE: Overall, the reviewer selection process appears reasonable and multi-faceted to achieve a diverse pool of reviewers with various perspectives. The subcommittee observed good geographic diversity, type of institution diversity, and sector diversity among the reviewers. It is particularly noteworthy that NSF is reaching out to organizations such as NSBE to identify reviewers. Reviewer demographics are very important – but this is better addressed in the data in aggregate form than in the information in the eJackets. The demographic information of the reviewers was often incomplete in the Review Record, presumably because the reviewers opted not to report. The information provided indicates that the reviewer gender is roughly evenly split between women and men. This is an excellent and noteworthy balance between men and women; however other genders are not represented. An analysis of the underlying geographic data provided further insights: While all regions are represented, the West is least represented (in terms of the number of reviewers per capita), which makes sense given the distance that reviewers would have to travel to participate in an in-person panel. Given the recent shift to remote/virtual meetings due to the pandemic, this might be an opportunity to increase the participation of reviewers from the West via virtual meeting technology. Also, larger states tended to have less representation on review panels, in general, than smaller states. Only one of the 7 largest states had more than the national average representation (reviewers per capita) on panels. Outliers:	YES

For one proposal, 4 out of 5 reviewers were from large public R1 institutions; two of those reviewers were from one institution. The fifth reviewer's affiliation is not listed. With 5 reviewers it should be possible to achieve more balance regarding institutional perspectives. Recommendations The eJacket should provide specific data on the type of institution for each reviewer, including minority serving institutions and type of institution such as 2-year, 4-year, PhD granting, etc. Ideally this would be auto populated (from PRIM) when a reviewer indicates their institution, to ensure consistency in the way institutions are categorized. Establish a mechanism to achieve more inclusive gender representation among panelists. (This may require providing additional gender options, such as gender nonbinary, in the demographic questions) Use virtual meeting technology to engage more reviewers from the West in order to achieve an even better geographic distribution. ENGR Education: Based on the information provided, it was found that reviewer selections are appropriate, including both experienced PI's and new reviewers. Gender, minority status, geographic locations, and the types of institutions where they come from are also considered during reviewer selection. However, it was found that certain individual panels may have lacked diversity in one or more forms (e.g., gender, under-served populations, etc.).	

III. Questions concerning the management of the program under review. Please comment on the following:

MANAGEMENT OF THE PROGRAM UNDER REVIEW
1. Management of the program. Comments: ERC: The EEC division is efficiently managed with a dedicated team of professionals. It is very clear that the team members are service-oriented and care deeply about the ERC program. The ERC program is truly the crown jewel of the ENG directorate, and it is in very good hands. Many workshops have been conducted by POs to organize and disseminate to the community. WFD/BPE: The structure of the Division and the responsibilities of the Program directors are well defined. There are frequent opportunities for communication and collaboration through biweekly meetings. The addition of discipline-specific co-Program Directors who have cross appointments with other divisions is a strength. POs, administrative staff, Division director and deputy directors collectively address and manage a host of program management issues. It is difficult to ascertain whether there is a fair distribution of workload across the various programs. Recommendation: NSF should look at workload across programs / Program Directors using a number of metrics (number of awards, number of proposals, budget allocation). ENGR Education: There seems to be a clear hierarchy at NSF regarding the decision for awards. Decisions are made through a series of levels, which provides the necessary oversight to ensure no proposal is funded simply because one person is in strong support of the effort. With that said, no documentation is included in the jacket regarding insights the DD or DDD have toward a proposal or cluster of proposals. From the jackets, there is no indication as to how the higher levels at NSF review proposals and make decisions regarding funding, such as the group meetings with other POs to review the clusters and allocations of funding and awards to fund. There is no documentation trail for these meetings in the jackets. Further, from this template, there is not much information regarding the follow through on the proposals and what happens after they become grants. Sections I and II focus predominantly on the pre-award process. The data presented to the COV is not self-contained such that the team needs to review additional resources (e.g., presentations). This team is fortunate to have a former PO on this sub-committee to add context to the management of programs. In fact, there are many outreach and professional development opportunities available to program directors. However, it is widely known that POs are very busy and so this sub-committee wonders how much additional training/professional development would better aid in the review process? Further, if there is an NSF wide checklist, this would be very helpful to POs as there are many nuances that need to be addressed in reviewing and managing proposals. In all, this

question is a large and broad question and the resources available to properly address it are not offered to the COV.

2. Responsiveness of the program to emerging research and education opportunities.

Comments:

ERC:
EEC management and POs are actively scanning the research and education landscape for emerging opportunities. There were only a few workshops on technical topics and thematic priorities. This can be expanded in the future, especially to highlight convergent topics.

WFD/BPE:
There are a variety of mechanisms in place to ensure that the Division is well-positioned to be agile and responsive to emerging research and education opportunities. NSF management and Program Directors engage in portfolio review to proactively explore changes to solicitations that are needed. The Program Directors maintain currency in the field through journals, conferences, workshops, and other resources. An array of workshops is organized to identify emerging trends through leaders of research community (e.g., AP in Engineering, Energy, Mechatronics and Robotics, Diversity, Makerspace, among others). NSF-sponsored workshops in a broad range of emerging areas are an asset in terms of community building that results in identification of future directions.

In general, BPE and WFD seem responsive to emerging research and education opportunities. Evidence of this is found in workshops that have been supported by BPE as described in the information provided. The subcommittee commends the focus on emerging leaders, junior women faculty, making academic change happen, and diversity in EEC faculty. These topics of emerging research also come up in reviews, panel summaries, and review analyses. For WFD, topics of most REU and RET were called out as important emerging areas.

Recommendations: Consider looking at outcomes and best practices of research-focused education programs and develop funding opportunities for translating this research into practice. More cross-pollination across programs is desirable.

ENGR Education:
EEC has established a good mechanism to identify emerging research and education opportunities through its broad communication network. Among the proposals reviewed, funded RED, CAREER and EAGER proposals especially address these opportunities. Program officers have opportunities to attend conferences where they can identify new areas. Program officers also look to journal papers where authors indicate the “next big thing” and follow up with these areas. Of the workshops presented in the information provided, many of these are broadening participation and other specific areas, but it seems that over the past few years there is a lack of what is also known to be important areas such as data science, AI, cyber security used in education at the course/curriculum level (note, these might be found in other areas of NSF). Also, if this is happening in other areas, has NSF been bringing together two different areas to see where intersections for new research can be formed. What is missing is a symposium or survey to the community on what is the next thing to be researched? For the COV, the participants should have been provided with the NSF’s strategic plan regarding new programs, re-evaluating current programs, or investigating emerging areas; and the process for how changes are made or instituted.

3. Program planning and prioritization process (internal and external) that guided the development of the portfolio.

Comments:

ERC:
Internal reviews are focusing on the right questions and provide opportunities to improve program planning and prioritization. Seeking feedback from panel reviewers provides valuable information from people intimately involved in the associated programs.

WFD/BPE:
The portfolio review at the division level is an appropriate and necessary part of the prioritization process. The planning and prioritization are aligned to ensure that proposals with fit in cluster themes are recognized and supported. The feedback from panelists is also very important as it provides a different perspective on the solicitations, process and programmatic operation.

Comment:
The questions or topics to be discussed in review meetings with Program Director and Division management are good. It is also good that the EEC DD or DDD meet with panel reviewers for debrief sessions. Without any details, it is hard to comment on the effectiveness of these activities.

ENGR Education:
This is a difficult question to address as very little information about this was provided. It seems that the program portfolio review is relatively new to NSF, and as a result there may not be an overly formal process in place yet. It was mentioned earlier that there is no information in the jackets regarding discussions after the panel and the PO's summary and before announcement of the award/decline. As example, the DD/DDD often meets with the panelists just prior to the team adjourning to obtain formative feedback about the panel and how the programs may be improved. It was mentioned that on more than one occasion, changes were made.

It would have been helpful to know how a jacket fits into the overall cluster and division prioritization. There is boilerplate information regarding the jacket in terms of the overall program. The questions posed were provided, but there is no data, or summary of the data, about the responses to these questions or how they were used. It was appreciated that a gap analysis had been conducted and that they are using the results for planning and prioritization.

At the end of the fiscal year, a new budget is proposed and awarded from Congress; hence, there should be analysis and prioritization that has been made before the various programs receive their budgets. It is recommended that data is collected, and a summary is provided for each of the 6 questions.

4. Responsiveness of program to previous COV comments and recommendations.

Comments:

ERC:
The ERC program has responded very well to the COV comments and recommendations from 2016. Documentation provided to the COV in the form of reviewer CVs gives more information on reviewers' qualifications. Virtual participation in panels has enabled more geographical diversity. More industry reviewers are being used in ERC panels. The committee was pleased to see the division leveraging data analytics to take a more systematic approach to their portfolio and would

like to learn more about the efforts of their AAAS Fellows in this regard. The development of Gen-4 ERCs is a significant step forward and the planning grants are a great way to broaden participation of targeted groups that may not have been able to participate in ERCs. The program facilitated this by providing the COV with the previous report, the response to the previous report, and documentation of other actions. It made the current effort seem important and worthwhile.

WFD/BPE:
Overall NSF has been very diligent and thorough about responding to COV comments and recommendations. Variability in addressing merit review was well addressed through the Panel Summary template.

The issue of reviewer expertise (vs. qualification) still remains. It would be too onerous for the COV to review the CV of every reviewer in their portfolio. It would be very helpful if the PRIM system provided this information.

ENGR Education:
Overall, the NSF provided detail responses and has made changes to aspects of their programming and processes. COVID has provided a better opportunity to have a true geographical distribution of the reviewers as distance no longer is a factor (only reliable internet). In general, the NSF has made a positive effort to address the 2016 COV concerns.

IV. Questions about Portfolio. Please answer the following about the portfolio of awards made by the program under review.

RESULTING PORTFOLIO OF AWARDS	APPROPRIATE, NOT APPROPRIATE, OR DATA NOT AVAILABLE
1. Does the program portfolio have an appropriate balance of awards across disciplines and sub-disciplines of the activity? Comments: In general, the COV found that the division had an appropriately balance of awards across disciplines and sub-disciplines. Program specific comments are below. ERC: Overall, the ERC portfolio is balanced driven by important thematic areas. WFD/BPE: Within EEC, we can see awards by thematic clusters rather than disciplines. Each program represents a cluster. There is a distribution of awards across the four clusters, although Center funding represents a significant portion of the EEC budget. ENGR Education: We do not have data balanced across the disciplines of engineering. Hence, this question and the data provided is somewhat skewed as engineering education does not conform to one engineering discipline and is its own discipline. The number of awards by clusters seems to be appropriately balanced. However, in term of the dollar amount of funding, a majority of the funds go to the Centers and Networks cluster, but how much of this comes from this directorate and how much comes from other directorates? The education piece for the Centers is very small compared to the disciplinary spending towards the ERCs. As a result, there is somewhat a mismatch of the name of the Division! It would have been good to see what the data was associated with around the disciplines of engineering. It is important to know if the funds are going predominantly to bioengineering versus civil engineering. This is more relevant to the Centers given that the lion share of the funding is going to Centers.	YES
2. Are awards appropriate in size and duration for the scope of the projects? Comments:	YES

In general, the COV found that size and duration of the awards was appropriate for the scope of the projects. Program specific comments are below.

ERC:
The implemented increase in ERC award size was sorely needed. Ten-year duration for ERCs is appropriate for the scope and impact of the proposed work. If anything, the ERCs are underfunded given the degree of commitment and involvement required from each PI to keep the program integrated and on track. Many of the PI's have large groups, so the ERC funding does not command enough of their bandwidth. It's imperative to assess degree of commitment at the start and then throughout. Planning grants recipients should participate in a future ERC preproposal submission.

WFD/BPE:
It is difficult to assess the appropriateness of the award size and duration, since this depends on the scope of work that is detailed in the proposal narrative. This is not reflected in the proposal summary or other information provided.

From the data provided, planning grants appear to have decreased the average duration of awards. Grants with longer durations may have a higher likelihood of accomplishments that are more long-term. This would be worth exploring.

The subcommittee noted that the BPE award size and both BPE and WFD award durations are decreasing. This is a concern because these awards are supposed to be broadening participation and developing the skills of the next generation of workforce professionals. The size and duration of the awards should allow time for that to occur. Since NSF awards not only provide support for activity and/or research to be conducted but also support education of graduate students and play a critical role in building the careers of PIs, it is important to have some stability, if not actual growth, in award sizes with passage of time.

Recommendation: Assess the institutional diversity and PI diversity of the planning grants (EEC in general). Are they well-advertised?

Look at the entire budget structure and allowable costs for WFD grants – ensure that they are up to date for today's economy and present attractive opportunities to the participants as well as adequate support for staff. Funding should align with program goals and provide adequate financial support/incentive to attract high quality teams with diverse representation.

ENGR Education:
The size and duration of awards highly depends on particular programs (CAREER, RED, REF, RIEF, etc.). While the information provided gives statistical data on the average size and duration for all awards made within EEC, the size of individual awards varies from one program to another.

CAREER awards, which typically involves only one PI, seem to be underfunded for the amount of effort required and the length of time. In contrast, ERC awards received the largest amount of funding with many

individuals supplementing the PI's. RED awards seem to be the most ideal based on length and activity. We recommend the NSF to organize more CAREER proposal development workshops for potential PI's to develop more competitive CAREER proposals and increase the funding amount of CAREER awards as well. Further by removing an ERC, many more RFEs or RIEFs may be funded; thus, enhancing the number of engineering education research projects and would ultimately enhance the portfolio.	
3. Does the program portfolio include awards for projects that are innovative or potentially transformative? Comments: In general, the COV found that the program portfolio included innovative and/or potentially transformative projects. There continues to be discussion over what precisely is meant by "potentially transformative". Program specific comments are below. ERC: The ERC program is all about generating innovative ideas that have excellent transformative potential. ERCs have the resources to fully explore these ideas, demonstrate new devices and systems in testbeds, develop a specially trained workforce, and spinoff new technologies to established and startup companies for maximum impact. ERCs cover many emerging research areas WFD/BPE: NSF's long-time focus on potentially transformative awards is crucial to impactful investments. For WFD, the information provided calls out two programs that sound great: Engineering for US ALL (led by U of MD) and Mega RET/REU site led by Morgan state The IP disclosures, start-up companies, E4USA, Mega RET-REU, etc., illustrate that the innovation and transformation of EEC is on-going. The titles of some of the jackets support innovative and potentially transformative work particularly in the BPE projects. Recommendation: A project does not necessarily have to be novel or innovative in order to be potentially transformative. Develop opportunities for sustainability, implementation, adaptation. Consider co-funding partnership opportunities with external organizations and industry. ENGR Education: It is difficult to determine whether a proposal will be innovative or potentially transformative so close to having been funded or completed. Panel jacket reviews found only a few instances where the terms of transformative, innovative, or some variant of these were used, across the jackets. The data	YES

provided included information on lag indicators of innovation (e.g., invention disclosures, spinoffs, patents, etc.) that were predominate outcomes of the ERCs. Unfortunately, these types of metrics are not likely appropriate lag indicators for the innovativeness or transformative-ness of engineering education research. Two exemplar examples were provided for Workforce Development, but metrics were not provided in their description. One mark of transformative-ness is that the research changes engineering education. In doing so, one metric might be the long-term propagation and/or uptake of the innovations or results of the research. This is often tough data to collect because it requires longitudinal data collection; and such data does require close attention be paid to aspects of dissemination. A recommendation might be to determine what lag indicators might be best suited for engineering education research and when they should be measured. Here is where the NSF can capitalize on Broader Impacts to serve as an avenue for stronger metrics of transformative-ness.	
4. Does the program portfolio include inter- and multi-disciplinary projects? Comments: In general, the COV found that the division included inter- and multi-disciplinary projects. Program specific comments are below. ERC: ERCs, especially Gen-4, are transdisciplinary and convergent. They are the best examples of inter and multi-disciplinary collaboration that NSF has to offer. The cross-disciplinary category in the rubric addresses exactly this. Many other NSF programs are contributing to the ERC program. The thematic areas around which ERCs are organized are all connect various disciplines. They connect multiple disciplines and technology sectors. WFD/BPE: Co-funding of proposals with other NSF directorates is an important tool for leverage EEC funds and to support interdisciplinary work. It was good to see a large number of proposals in EEC that were co-funded, and that EEC is co-funding proposals in other directorates. The COV is surprised that there is not more of this type of funding in partnership with EHR and CISE. The projects in WFD (REU and RET) and BPE tend to involve inter-and multi-disciplinary teams and goals. For example, BPE often includes engineers and other disciplines including social sciences. REU/RET should include some folks with expertise outside of engineering or science to help with the professional development although this was not clear from the information in the eJackets. It should be noted that the information provided did not contain specific data for the WFD or BPE programs.	YES

ENGR Education: Given that two-thirds of the funding goes to ERCs, information provided clearly demonstrates that multi-disciplinary work is certainly happening in the directorate. For the portion of funding that goes to Engineering Education and Workforce Development, there was evidence that collaboration with other Directorates (e.g., CSE, ENG, GEO, MPS, EHR, etc.) occurs to fund EEC awards. Lastly, a strong argument can be made that engineering education research, by its nature is both inter- and multi-disciplinary as it integrates multiple engineering disciplines along with education and learning sciences.	
5. Does the program portfolio have an appropriate geographical distribution of Principal Investigators? Comments: In general, the COV found that the division had an appropriate geographical distribution of Principal Investigators given the pool of proposals that it received. Program specific comments are below. ERC: ERCs are well distributed across the U.S. and have foreign participating organizations around the world. ERC lead institutions are concentrated in the South and South East. WFD/BPE: Overall the distribution is geographically diverse. It should be noted that the data provided represent all of EEC, not a breakdown by program. The data shows submissions from all over the USA with generally more submissions from states and territories with higher populations which is reasonable. Overall, the geographic distribution of PIs looks reasonable. There were four states with no awards, but also a low number of submissions. It would be helpful to know the percentage of proposals funded indicated in addition to the numbers. Recommendation: It would be helpful to include information about the number of submissions and percentage of proposals funded by state. ENGR Education: The program portfolio includes competitive proposal submissions from all 50 states, plus Puerto Rico and Guam. The relative number of proposals generally aligns with the 2018 population for each given state, i.e., states with larger populations and a greater number of institutions have a larger number of potential PIs. Some states or territories have not received an award. The larger number of submissions and awards is to be expected based on the existence of lager research institutions (R1s) and potentially engineering education programs. NSF awarding top scholars that may all be at one institution should not diminish any abnormalities in geographical diversity. For example, Indiana and Virginia have a higher number than expected based on population. This is likely due to Purdue University's School of Engineering Education and Virginia Tech's Department of Engineering Education. Top	YES

scholars in engineering education at each institution warrant a larger number of proposals be awarded to these PIs and their institutions.	
6. Does the program portfolio have an appropriate balance of awards to different types of institutions? Comments: In general, the COV found that the division had an appropriate balance of awards to different types of institutions given the pool of proposals that it received. Program specific comments are below. ERC: ERCs provide excellent opportunities for minority serving institutions and HBCUs. The mixture of participating institution types in the portfolio could be improved with more minority serving institutions leading ERCs. WFD/BPE: The portfolio shows awards to different types of institutions for all of EEC. The data are not available for WFD and BPE, specifically. The top 100 Research intensive universities have by far the most awards. 2-year institutions have awards only in 2017. Recommendations: Include Minority Serving Institutions in the data. Show collaborating institutions in the data. Encourage partnerships with 2-year colleges. Encourage collaborations with 4-year institutions and with 2-year institutions. Allow PIs to propose innovative models of collaboration with faculty and students from such institutions, beyond the timeframe of summer. The subcommittee underscores the importance of involving 2-year institutions in WFD and BPE awards, especially since there tends to be greater diversity at two-year institutions. The subcommittee recommends an effort to solicit more proposals involving 2-year institutions, either as host or collaborative institution. ENGR Education: The total number of funded proposals (and subsequent award obligations) is swayed heavily toward PhD granting institutions, particularly 'Research Intensive PhD Institutions (Top 100).' This is not surprising for a number of reasons. Faculty attracted to these institutions are likely interested in doing research. This requires obtaining funding to support doctoral students and there is a greater research expectation to receive promotion and/or tenure. This will always impact the program portfolio. A far smaller number of proposals are submitted by 2-year, 4-year, Masters, and other organizations (non-profits, local government, professional societies, U.S. Government agencies, small businesses, large corporation businesses, and other private and public organizations). These proposals	YES

have similar success rates, which suggests that the program portfolio is providing opportunities for all types of institutions. The reality is that NSF cannot force diverse institutions to submit proposals and NSF can only fund what is submitted. A recommendation is to increase outreach that encourages more diverse institutions to collaborate with PhD granting institutions to increase their representation within the program portfolio.	
7. Does the program portfolio have an appropriate balance of awards to new and early-career investigators? NOTE: A new investigator is an individual who has not served as the PI or Co-PI on any award from NSF (with the exception of doctoral dissertation awards, graduate or post-doctoral fellowships, research planning grants, or conferences, symposia and workshop grants.) An early-career investigator is defined as someone within seven years of receiving his or her last degree at the time of the award. Comments: In general, the COV found that the division had an appropriate balance of awards to new and early-career investigators. Program specific comments are below. ERC: While new investigators are typically not PIs of ERCs, they play a significant role in many ERCs. ERCs can provide an easy path for new faculty to develop collaborations and become part of a research team, learning important life skills in the process. New PI's can be funded via seed grants, mentored, and included in the leadership succession planning. WFD/BPE: The information given to the COV did not provide program-specific data on WFD and BPE programs, so the subcommittee examined the data for EEC overall. The EEC portfolio attracts a large number of proposal and makes actual awards to new PIs and PIs with prior submissions. Moreover, the program also funds CAREER PIs. It would be helpful to have this data broken down by program. New investigators would not be likely to be competitive for an ERC but could be for other programs. Does the data distinguish single PI submissions? Are the successful new investigators more likely to be in collaborative proposals with more experienced co-PIs? If so, encourage more collaborations with new investigators. Recommendation: It was not possible to determine whether grants were awarded to new investigators without going into the PI history in each	YES

eJacket. It would be helpful to provide this information in summary form for future COVs. ENGR Education: New and early investigators receive funds primarily through the RIEF and CAREER programs. The funding rate of CAREER awards appears consistent from year to year. In terms of the other types of awards, RED must be submitted by a senior administrator (likely seasoned PI), RIEF should be for people wanting to get into engineering education research – albeit associate or full professor (and some assistant professors). The RFE and EAGER should be across the board. The overall success rate of new PIs ranges is lower than PIs with prior proposal submissions, which is to be expected as people learn how to write successful proposals. The data provided about the percentage of new PI proposers and new PI awards was encouraging given that the pool of experienced PIs is likely much larger than the pool of new PIs.	
8. Does the program portfolio include projects that integrate research and education? Comments: In general, the COV found that the division's portfolio included projects that integrate research and education. Program specific comments are below. ERC: ERCs integrate research and education in a strong way with workforce development playing a major role in the center's efforts. WFD/BPE: By the nature of the programs in the EEC portfolio, projects integrate research and education. All CAREER proposals must include this and RET and REU integrate research experiences and education. Some BPE and INCLUDES include research experiences. For example, an INCLUDES project investigating the experiences of Black women in graduate engineering programs aims to give back to supporting current and future Black women in graduate engineering programs and educating the community about how their experiences. ENGR Education: The EEC program has appropriate participation of underrepresented groups. However, the participation of underrepresented groups varies from one program to another. The Broadening Participation program, by its nature, has done excellent work in terms of involving participation of underrepresented groups including underrepresented PI's. Other programs, such as CAREER, RED, RIEF, REF and ENGAR, involve the participation of underrepresented groups to a certain extent. Females were particularly well represented in proposals.	YES

Statistical data showed that proposals with minority involvement have a higher success rate than those proposals without minority involvement. We recommend the NSF collect data to determine how successful and/or what impact the studies getting funded with minority PI's are having relative to proposals that do not have minority PI's.	
9. Does the program portfolio have appropriate participation of underrepresented groups[3]? Comments: In general, the COV found that the division had an appropriate participation of underrepresented groups given the pool of proposals that it received. The division should continue its efforts to increase the number of proposals it receives from PIs from underrepresented groups and institutions that serve substantial numbers of people from underrepresented groups. Program specific comments are below. ERC: The ERC portfolio has an appropriate participation of underrepresented groups. In fact, the statistics for women in ERCs are the highest they have ever been in every category. The women faculty numbers, however, are still too low. ERCs do a great job of providing a diverse and inclusive environment and that is a major reason why the statistics are much better than national averages. The results for underrepresented racial minorities, however, seem to be going in the wrong direction, a very disappointing result considering the efforts that are being made to increase diversity. Many of the underrepresented groups cited in the proposals that this subcommittee read were targets of outreach programs and not necessarily participants in the planning of the research or serving as PI's or Co-PIs. It is noteworthy that several of the proposals did not include a plan for workforce development or outreach to underrepresented groups. Underrepresented minority PI's among the proposals reviewed were lacking. The participation of underrepresented leaders is essential at the outset and continuing through the life of the center. WFD/BPE: The information provided was very helpful, particularly that broken down by program. It would be helpful to have data that combined URM and women. For EEC overall, far more proposals are submitted without minority involvement then with, but the success rate is higher for those with minority involvement. The trend seems to be that those with and without minority involvement are approaching the same success rate.	YES

[3] NSF does not have the legal authority to require principal investigators or reviewers to provide demographic data. Since provision of such data is voluntary, the demographic data available are incomplete. This may make it difficult to answer this question for small programs. However, experience suggests that even with the limited data available, COVs are able to provide a meaningful response to this question for most programs.

For BPE, in most years there were more proposals with minority involvement and success rate is higher with minority involvement. Success rates and the number of proposals submitted vary quite a bit over the time frame examined. Is there an explanation for this? For WFD, the data show that in most years there were more proposals without minority involvement and the success rate is varied. In general, we do not think the number of proposals submitted, or the number proposals funded with minority involvement meets the "appropriate representation" criteria. This is an area that should be carefully considered and investigated. For BPE, success rates are higher with women involvement for all years. There is considerable variation in the number of proposals submitted and the number proposals funded with women involved. Is there an explanation for this? For WFD, the number of proposals with women involvement indicates appropriate representation. Because of the nature of the WFD programs, involving women and URM is really important. It is hard to convince more diverse students to enter engineering if they do not see role models. Involvement should be as leaders/investigators and as mentors (personnel). Perhaps more women and URM are involved in the mentoring or other aspects of the program beyond being PI or co-PI. In one panel summary, it was specifically called out as a "unique" strength a majority of the mentors for the REU would be from underrepresented groups. This should be the norm, not unique. We must be careful not to put too large of a service burden on women and URM but it is appropriate for them to be rewarded for their efforts which could be done via an REU. NSF should do more to encourage submissions from diverse project teams. Recommendation: Why are REU and RET's limited to one PI and one co-PI? Consider allowing multiple investigators. If women and minorities are involved as senior personnel, this may be perpetuating inequities around people getting credit for their work and hampering their ability to get future grants. It is important for women and URM to be included on the leadership team. Consideration should be given to elevating the prestige of participating in such programs – for example, not allowing funding for a graduate student or faculty summer salary sends a message about the importance of the program and may influence perception by tenure and promotion committees and limits the pool of PIs who can realistically apply for this program. ENGR Education: The EEC program has appropriate participation of underrepresented groups. However, the participation of underrepresented groups varies from one program to another. The Broadening Participation program, by its nature, has done excellent work in terms of involving participation of underrepresented groups including underrepresented PI's. Other programs, such as CAREER, RED, RIEF, REF and ENGAR, involve the participation of underrepresented	

groups to a certain extent. Females were particularly well represented in proposals. Statistical data showed that proposals with minority involvement have a higher success rate than those proposals without minority involvement. We recommend the NSF to collect data to determine how successful and/or what impact the studies getting funded with minority PI's are having relative to proposals that do not have minority PI's.	
10. Is the program relevant to national priorities, agency mission, relevant fields and other constituent needs? Include citations of relevant external reports. Comments: In general, the COV found that the division's programs were relevant to national priorities, agency mission, relevant fields and other constituent needs. Program specific comments are below. ERC: ERCs are working on some of the nation's greatest technical challenges and informed by the NAE grand challenges and the National Research Council, for example 2014 NRC Report on Convergence and the 2017 National Academies consensus study report on A New Vision for Center-Based Engineering Research. ERCs have a huge visibility and impact beyond academia. WFD/BPE: Broadening participation, workforce development, integrative education and research, interdisciplinary research, etc., are all national priorities identified in numerous NAE reports and other publications. National Academy of Engineering (2006). Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering. Washington, D.C., National Academies Press. Committee on Science, Engineering and Public Policy, Division on Policy and Global Affairs, National Research Council (2011). Expanding Underrepresented Minority Participation: America's Science and Technology Talent at the Crossroads. Washington, D.C., The National Academies Press. National Academy of Engineering (2008). Changing the Conversation: Messages for Improving Public Understanding of Engineering https://www.nap.edu/catalog/12187/changing-the-conversation-messages-for-improving-public-understanding-of-engineering National Academy of Engineering (2005). Educating the Engineer of 2020: Adapting Engineering Education to the New Century https://www.nap.edu/catalog/11338/educating-the-engineer-of-2020-adapting-engineering-education-to-the#:~:text=Educating%20the%20Engineer%20of%202020%20is%20ground	YES

ed%20by%20the%20observations,Engineering%20in%20the%20New%20Century . ENGR Education: Given national and agency priorities, it does appear that engineering education programs do serve national interests by providing solicitations aimed at conducting research that improves the understanding and producing of "better" engineers. Of the solicitations issued in the EEC, 5 of the 12 solicitations are specifically aimed at engineering education research. Although most of the funding goes towards ERCs, there are multiple opportunities to make an impact from engineering education research (e.g., research, interventions, etc.). Further, through evaluation of jackets, topics of sustainability, diversity (including neurodiversity, First Gen, Hispanic, women of color), K-12 education, ethics and society, socio-economic impact (program for underserved groups) were well represented in the engineering education proposals.	
11. Additional comments on the quality of the projects or the balance of the portfolio: An additional effort should be made at the aftermath of COVID 19 to cover topics impacted by our need to be prepared for such emergencies.	YES

OTHER TOPICS

1. Please comment on any program areas in need of improvement or gaps (if any) within program areas.

   ERC:
   Propose thematic priority areas for planning grants. Possibly using an RFI process.

   Compressing the timeline for the ERC proposal process would be beneficial for both successful and unsuccessful ERC teams.

   WFD/BPE:
   Recommendation: Include data on % of REU and RET that are renewals. How are these categorized – difference between renewal, supplement, etc.? What is the breakdown by awards given back to same groups?

   The WFD REU and RET programs should make more of an effort to obtain submissions from diverse teams that include women and people of color. The percentages of proposals including people of color as reported here are unacceptable. One suggestion is to have a webinar with current awardees to help future awardees as was done with the RED program.

   RETs are limited to engineering and CSE. The subcommittee feels that they would be very valuable to other disciplines and could be a broader program within NSF. Expanding to include science would also be consistent with the goals of the Next Generation Science Standards (NGSS) which are increasingly important for K-12 education in the USA.

We suggest a careful review of REU and RET programs including solicitation, budget allocation, number of investigators, scope, etc.

ENGR Education:
There was not enough on the area of cross over research between other directorates and engineering education. There needs to be a push to the envelope in emerging areas of how data science, cyber learning, artificial intelligence can be infused into the current solicitations (and future solicitations). Perhaps creating a workshop or set of workshops to bring together individuals from the different fields.

Another area might be to further RED to push the envelope on curricular changes in engineering, as well as other academic areas such as advising so that interventions may be employed.

A third area is to look at engineering education at the graduate level. Curriculum and models at the graduate level have not changed in eons. It might be great to extend RED to the graduate level. Aspects of credit hours, semester/term, and admissions. Many companies (e.g., Amazon, Google, etc.) are creating 'schools' with radically different platforms of education that provide credentials that put at risk the current framework of graduate engineering. The opportunity to be experimental is great; plus, it contributes to a national need. Additional areas that are growing are post baccalaureate certificates with stackable opportunities to create graduate degrees. These innovations are occurring and being implementing but there are aspects that need research regarding the efficacy and learning that is occurring. Cross cutting programs that integrate the research and emerging practices. This might also include efforts around classroom environments (e.g., augmented, virtual, in person) and how learning is impacted.

2. Please provide comments as appropriate on the program's performance in meeting program-specific goals and objectives that are not covered by the above questions.

ERC:
The challenge of working across geographical distances is baked into the ERC selection criteria, which require 3 distinct lead institutions; however, this leads to a huge challenge in integrating and coordinating the efforts. Interestingly, the pandemic taught us something very important about virtual meetings—they are much more worthwhile for everyone involved when everyone is virtual. We encourage the use of virtual ERC proposal review panels for this reason.

ENGR Education:
From the directorate's website, "the creation of the 21st century engineers" is a fundamental to this division and to engineering education research. The skills they need are often beyond what we traditionally need, and this division does fund these current needs (e.g., teamwork, design, leadership), but culturally we need to see this shift occur across the U.S. Further, has the NSF funded research on skills that may be needed in the future?

How researchers consider sustainability and long-term impact is not necessarily something that is measured or followed-up on. There needs to be true buy-in for how the research results are "taken up" and infused in curriculum and in pedagogical practices across faculty. It speaks to the importance of cultural shift, which is ingrained in the RED solicitation. ERC's are also good with this as seen in their sustainability plans. In doing so, for solicitations like RIEF and RFE broader impacts become almost as important as the research itself.

CAREER awards in engineering education research have much promise, but it is disturbing that the numbers being submitted are decreasing. As reviewers and submitters to this program, the education and outreach plan seems to be a struggle for engineering education researchers. The

research and education/outreach plans are heavily interwoven; and hence, there seems to be a disconnect. Perhaps think about how this may be 'revised' for this area.

3. Please identify agency-wide issues that should be addressed by NSF to help improve the program's performance.

ERC:
More funding is needed for NSF, ENG and EEC. ERCs are frequently reviewed and held accountable for research, workforce development, and diversity deliverables. They are uniquely positioned to address grand challenges that are growing in scale and complexity. Additional ERCs would have a significant positive impact on the ability of the United States to maintain its global leadership in engineering and technology.

The ERC award decision should be synchronized with the academic calendar to permit on-boarding of students and faculty.

WFD/BPE:
To achieve the larger goals regarding national priorities and broadening participation, we recommend that NSF be a leader in this area and insist on inclusive, bias-free language in all communications including proposals, annual reports, and solicitations. Language matters. For example, it was noted that "PI History" lists "PI Handicap" Is this the best term for indicating a disability? Also, NSF could encourage PIs to use bias-free language as well, for example to replace "freshmen" with "first-year".

Some resources that might be helpful are the new APA guidelines https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/#:~:text=The%20American%20Psychological%20Association%20emphasizes,demeaning%20attitudes%20in%20their%20writing

Other interesting information regarding the use of the term "underrepresented minority (URM)" is available in Byron White's 2016 piece on anti-deficit thinking and "rising stars" https://www.insidehighered.com/views/2016/04/19/importance-viewing-minority-low-income-and-first-generation-students-assets-essay and Tiffani Williams's 2020 piece on "URM" as harmful racist term https://cacm.acm.org/blogs/blog-cacm/245710-underrepresented-minority-considered-harmful-racist-language/fulltext.

ENGR Education:
Clearly there are budgetary constraints, but perhaps the current funding could be managed differently. For example, if the education piece in the ERC is changed or expanded to include engineering education research, there may be additional opportunities to further education. Note that currently the education piece includes a REU, RET, outreach component and a young scholar's piece. These areas speak to the workforce development and broader participation, but not engineering education research. This is a missed opportunity; but would require funds be shifted within the ERC to go towards engineering education research. Perhaps model ERCs in a similar light to elements of the RED; or, expand additional monies to the RED program.

NSF should encourage more cross-cutting programs where discipline-based fields intersect with education research. There is a need and push to engage more interdisciplinary research – yet silos still reside in the NSF. With that said, PIs are often confused or lack knowledge in submitting cross-cutting proposals as they fear they will not be evaluated properly in a division or will get "lost" across directorates. It requires the NSF knowing how to best encourage and evaluate these cross-cutting proposals.

4. Please provide comments on any other issues the COV feels are relevant.

ERC:
All of the lead PIs for ERCs that this subcommittee reviewed are male. This needs to be improved through data collection and analysis of contributing factors that lead to this result.

ENGR Education:
To reiterate, a majority of the total EEC budget goes to the Centers and Network cluster. This could potentially result in excellent engineering education proposals submitted to other EEC programs not being funded due to insufficient funds.

5. NSF would appreciate your comments on how to improve the COV review process, format and report template.

Very well organized and run COV. It was a pleasure to participate. In view of the experience with the pandemic, the NSF should consider adopting virtual meetings as the basis for all of their review processes except the site visits. This will help with geographic diversity and lower the burden for URM members. It is a great leveler, eliminating the schmoozing aspect of the review process, which can disadvantage URM applicants. Finally, it will save a tremendous amount of time and money.

Some information in the jackets is in the form of email which is only available as outlook. For those of us who do not have access to outlook, this complicates our COV process. NSF staff were very helpful in converting items when requested.

The COV would have liked to have more access to the data underlying the figures in the information provided. Consider different ways of presenting data to enhance readability and understanding. Consider the work of a leading expert in visualization to help better present the data for maximum clarity. The next COV would benefit from this.

Consider a double-blind review process as a first stage of the review process. This might serve the same purpose of the triage. This would require some creative thinking and innovation in reimagining the review process. Overall NSF's merit review process is still the gold standard, but we still feel there is room to improve.

For future COVs, having a rubric by which to assess jackets would have been helpful. Additionally, in reviewing the information provided, often the type of data being sought was not the type of data that was provided. The information provided was also inconsistent. One panelist indicated that it looked as though different people complied different sections. Information was not provided on the management of the program which made it difficult to answer the questions about program management.

SIGNATURE BLOCK:

For the Engineering Education and Engineering Centers Committee of Visitors

Darryll Pines, Ph.D.
Co-Chair

Lance C. Pérez, Ph.D.
Co-Chair