CORE QUESTIONS and REPORT TEMPLATE
for
FY 2018 NSF COMMITTEE OF VISITOR (COV) REVIEWS

Guidance to NSF Staff: This document includes the FY 2018 set of Core Questions and the COV Report Template for use by NSF staff when preparing and conducting COVs during FY 2018. 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/COV%20Policy%20and%20Procedures%20070915.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 2018 REPORT TEMPLATE FOR NSF COMMITTEES OF VISITORS (COVs)

The table below should be completed by program staff.

Date of COV:
October 11–12, 2018

Program/Cluster/Section:

Division-Wide COV for the Division of Graduate Education (for FY14–FY17):
- CyberCorps®: Scholarship for Service (SFS)
- Secure and Trustworthy Cyberspace–Education Track (SaTC-EDU)
- NSF Graduate Research Fellowship Program (GRFP)
- NSF Research Traineeship (NRT) and Innovations in Graduate Education (IGE)
- Promoting Research and Innovation in Methodologies for Evaluation (PRIME)

Division:
Division of Graduate Education (DGE)

Directorate:
Directorate for Education and Human Resources (EHR)

Number of actions reviewed: 318
- **Awards:** 33 proposals + 70 GRFP applications
- **Declinations:** 185 proposals + 15 GRFP applications
- **Other:** 15 GRFP Honorable Mentions

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

2,093 proposals + 69,286 GRFP applications

- **Awards:** 309 proposals + 10,000 GRFP applications
  - NRT, including IGE: 89 proposals
  - SFS: 119 proposals
  - SaTC-EDU: 69 proposals
  - PRIME: 32 proposals
- **Declinations:** 1,784 proposals + 48,681 GRFP applications
  - NRT, including IGE: 1,193 proposals
  - SFS: 262 proposals
  - SaTC-EDU: 252 proposals
  - PRIME: 77 proposals
- **Other:** 10,605 GRFP Honorable Mentions

Manner in which reviewed actions were selected:

A subset of proposals was selected using numbers chosen by the COV Chair. The Chair specified the numbers ‘1’ and ‘5’. All proposals ending with the number ‘1’ were pulled for the sample. To ensure an adequate number of proposals for the SFS, SaTC-EDU, and PRIME programs, proposals ending with the number ‘5’ were *also* pulled.

COV Membership

	Name	Affiliation
COV Chair or Co-Chairs:	Rory Cooper	University of Pittsburgh and U.S. Department of Veterans Affairs
COV Members:	Michael Ashby Sez Atamturktur Emmanuel Collins Sara Hernández Ernest McDuffie Diane Miller Deana Pennington Terri Pigott Keivan Stassun	University of Oklahoma, Norman Campus Pennsylvania State University University of Louisville Cornell University The Global McDuffie Group Northrop Grumman Corporation University of Texas at El Paso Loyola University of Chicago Vanderbilt University

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**: • In general, the panel agrees that the NSF's review methods are appropriate. However, there was difficulty in determining whether the review method for any given review was face-to-face panel, virtual panel, ad hoc review, or site visit. The method of participation should be added to the jacket to aid future COVs. • The panel agreed this question would be more informative if the focus was on effectiveness of methods rather than appropriateness. Assessing effectiveness would depend on examining data of predefined metrics and comparing outcomes of these different methods. For example, in what ways, and to what degree, are panel discussions impacted by virtual interaction? Does this impact the way the panel engages (we suspect it does), and does that impact the kinds of proposals that are selected? The NSF may have considered these issues, but there was no evidence in the eJacket. The COV recognizes virtual panels are useful for a number of reasons and recommends an assessment of the impact of different kinds of online collaborative tools on outcome/impact. **Data Source: EIS/Type of Review Module**	YES
2. Are both merit review criteria addressed a) In individual reviews?	YES

b) In panel summaries? c) In Program Officer review analyses? **Comments:** For individual reviews: • Both merit review criteria are addressed in the individual reviews; although there is much variability in the comprehensiveness of the reviews. In some cases only a single sentence is provided for each. For panel summaries: • Both merit criteria were comprehensively addressed in the panel summaries. For Program Officer review analyses: • Yes, the review analyses all address both criteria, in most cases comprehensively, and in some cases going beyond the panel reviews and identifying additional strengths and weaknesses. However, it does not appear that this information is sent back to the applicant along with the panel summary and reviews. • The COV recommends that the Division of Graduate Education (DGE) develop consistent procedures across their programs for conveying at least some of this important information back to investigators. **Data Source: Jackets**	
3. Do the individual reviewers giving written reviews provide substantive comments to explain their assessment of the proposals? **Comments:** • A pattern emerged from the samples reviewed. Reviewers who were provided templates to follow mostly provided more comprehensive, helpful individual reviews, in comparison with those who were not given templates. These templates specified that reviewers should address both strengths and weaknesses in both intellectual merit and broader impacts and sometimes prompted reviewers to address specific objectives in the solicitation. The template usually followed with a list of other criteria to be considered, again specifying strengths and weaknesses. The COV recommends that the use of templates for individual reviews become a best practice across DGE as well as throughout the National Science Foundation. • The COV observed that even with templates, broader impacts tended to be less comprehensively addressed by reviewers than intellectual merit. • The COV suggests DGE consider providing checkboxes of common types of broader impacts, along with specific criteria to rate associated with each type of broader impact. For example, if the proposed project plans to conduct outreach through citizen science, how well have the	YES

investigators specified their methods of a) recruitment; b) training; c) validation; and d) feedback? **Data Source: Jackets**	
4. Do the panel summaries provide the rationale for the panel consensus (or reasons consensus was not reached)? **Comments:** • For the most part, panel summaries are quite comprehensive. If consensus was not reached, the summaries effectively captured the alternative perspectives. • The COV commends the Program Officers for facilitating the rigorous documentation of the panel discussion. **Data Source: Jackets**	YES
5. Does the documentation in the jacket provide the rationale for the award/decline decision? [Note: Documentation in the jacket usually includes a context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), program officer review analysis, and staff diary notes.] **Comments**: • Yes, the documents in the jacket provide the rationale for the award/decline decision. • Three proposals in separate programs that were reviewed had major discrepancies among the individual reviews, the panel summary, and the final award decision. It appeared that the Program Officer decided to recommend award when the reviews, especially the panel summaries, rated the proposals low competitive (or lower). Two of these discrepancies were addressed in the diary note and the review analysis, and the decision may have been entirely appropriate. The third did not address the discrepancy. Given the number of proposals that are rated highly by panels, the COV recommends that DGE exercise caution when situations such as these arise and ensure that the rationale is well documented and comprehensive, and the approach is consistently applied across DGE. **Data Source: Jackets**	YES
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	YES

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 most cases the PI was provided sufficient rationale for the award/decline decision. • Without good individual reviews, it is unclear what feedback the Program Officer can provide to the applicant. **Data Source: Jackets**	
7. Additional comments on the quality and effectiveness of the program's use of merit review process: • It might be useful for DGE to apply emerging text analysis tools to conduct new types of data analysis to, for example, match proposals with reviewers; assess the adequacy of the reviews; or examine the consistency between the text of the reviews and overall proposal ratings.	YES

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**: • Yes, however there could be better documentation of data sources. • The COV suggests collecting and reporting more detailed and consistent characteristics of the panel members – demographic characteristics, position, rank (for faculty), areas of expertise, etc. • There is insufficient data to determine if the panelists are a representative peer group of those whose proposals are being reviewed. Additionally, the method(s) of review used for all panels is unclear from the data. It would be helpful to include more complete and robust information about the panelists in the jackets. **Data Source: Jackets**	YES
2. Did the program recognize and resolve conflicts of interest when appropriate? **Comments:** • Yes, the conflicts were dealt with and the documents that explain the COIs were detailed and comprehensive. **Data Source: Jackets**	YES
3. Additional comments on reviewer selection: • It is important for the Program Officers to continue efforts to ensure greater diversity of reviewers at the micro (e.g., individual reviewers on a proposal) and macro (programmatic) scale. • The demographics collected on the reviewers for the Graduate Research Fellowship Program (GRFP) should serve as a template for what data are collected on reviewers and how those data are organized. Additionally, the practices the GRFP employed to collect more complete data on GRFP reviewers should be employed in other DGE programs.	YES

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:** • GRFP and NRT program plans include logic models that could be emulated within the management plans for other DGE programs. • The rationale for distribution of funding for GRFP across EHR and OIA is not clearly connected to programmatic objectives or priorities. • In general, dwell time appears appropriate. Clearly the multiple years of operating on continuing resolutions has taken its toll on dwell time in some programs. There is some evidence for increased efficiency in processing more proposals with a bit less dwell time. • Proposal pressure on NRT has taken a clear toll on the success rate, which has dropped [lower than the COV believes appropriate]. In light of the non-trivial effort involved in preparing and submitting these proposals, it could be valuable to implement a pre-proposal process for NRT. • It may be advantageous for NSF and the larger community to perform a cost-benefit analysis as proposal pressures increase and success rates decline. An example from the AST division can be found here: https://arxiv.org/abs/1510.01647. A two-stage process (a pre-proposal selective process) may be helpful as well. • The COV is concerned that the low ratio of program staff to numbers of proposals and panels is not sustainable and that ultimately quality of process and/or outcomes may suffer. This includes processes and outcomes pertaining to review, but also to monitoring and supporting active awards. • The COV would have welcomed information regarding the ways in which program quality measures and outcomes are collected, synthesized, and dash-boarded for Program Officers to effectively and efficiently track and assess programs.
2. Responsiveness of the program to emerging research and education opportunities. **Comments:** • Program solicitations, in general, include very good overview/rationale with literature citations to explain the emergence/timeliness of the opportunity. • NRT has identified priority areas that relate to NSF Big Ideas. However, the COV noted an apparent strong funding skew toward priority areas.

- GRFP has developed some programmatic innovations, such as GRIP and GROW, that expand opportunities for student training and development in ways that are timely and relevant.

- There is some evidence that the previous language of interdisciplinarity may be colliding with the new language of NSF Big Ideas. The Big Ideas are in many cases, but are not all necessarily, interdisciplinary. Some "merely disciplinary" projects that are eminently in line with the Big Ideas may get overlooked as a result of the sense that interdisciplinarity is a goal in and of itself.

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

Comments:

- It is very difficult, if not impossible, to assess the program planning and prioritization process across DGE in light of the change of COV purview from individual program level to pan-division level. The COV recommends that, if it does not already exist, there be created a closed-loop division-level plan for mapping overall division objectives and priorities through its programs and to outcomes, and that this be provided to the next COV well in advance of the panel meeting. This context is critical for assessing how well the portfolio meets division and NSF-wide goals.

- NRT funding from directorates/divisions is very uneven but this appears to be appropriate given the stated priority areas and the strong emphasis on cybersecurity and data-enabled science and engineering in years under the COV's review. The COV thinks it is appropriate that the program is not limited to these topics and is open to innovative ideas that emerge from the STEM community. The COV recommends continuing this procedure.

- The COV applauds the decision to move IGE to its own program, separate from NRT, enabling new approaches to graduate education to be pilot tested.

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

Comments:

- At the programmatic management level there was insufficient evidence to evaluate continuous improvement and performance measurement.

- There were comprehensive responses to previous COV reports across the various programs. However, there were only updates to COV responses for GRFP and SFS; more consistency in closing the loop on COV recommendations, responses, and conclusions would help to ensure that COV recommendations are ultimately acted upon as appropriate.

- Some examples found in a spot review of random proposals in GRFP indicate that there may still be an issue with some applicants being knocked out on the basis of "one fatal flaw" instead of holistic review. This is especially problematic in cases where the "one fatal flaw" pertains to a perceived lack of prior achievement. The GRFP as constructed is not intended to be an achievement award but rather an award based on promise and quality of the proposal.

- The GRFP program statistics indicate that while gains have clearly been made in the proportion of awards to underrepresented groups, individuals with disabilities and veterans continue to be significantly underrepresented. Studies suggest that this population of students may be more likely to take a “GPA hit” due to challenges in the transition to college, even if subsequent and overall performance is strong.

- There was no prior COV report for NRT because of a programmatic change from IGERT to NRT.

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

RESULTING PORTFOLIO OF AWARDS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE
1. Does the program portfolio have an appropriate balance of awards across disciplines and sub-disciplines of the activity? **Comments:** • In general, graduate education and workforce development programs should be based on projected national workforce development needs and areas of growth. For example, within GRFP the disciplines of chemistry (8%), CISE (2%), engineering (26%), geosciences (6%), life sciences (29%), materials research (2%), mathematics (3%), physics and astronomy (5%), physiology (8%), and social sciences (9%) are all well-represented, albeit not evenly distributed. The percentage of funded applications is approximately the same in each discipline. Within each discipline, a plethora of sub-disciplines is represented. The success rate within each sub-discipline appears approximately constant. Since the panels are organized around each discipline (and sometimes sub-disciplines), a statistically equitable distribution of awards can be expected. It would appear that the percentage of the overall budget that is awarded by each panel is proportional to the number of applications. It appears that the process of selecting the successful GRFP applicants is driven by the applicant pool, not the projected nationwide workforce development needs. • Based on 2013-2017 data by institution type, over 20% of GRFP awardees and funding go to three institutions. This may conflict with NSF's vision for national capacity building in graduate education and workforce development. • NRT – Projects are evenly distributed across demographics, geographic areas, etc. The majority of the funding is allocated to research intensive institutions, as expected. • IGE – Projects are evenly distributed across demographics, geographic areas, etc. Twenty-one out of 50 states are represented by 31 awards since 2014. • PRIME – Insufficient data to evaluate. **Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data EIS/Committee of Visitors Module. From the Report View drop-down, select the Funding Rate module to see counts of proposals and awards for programs. The Proposal Count by Type Report View will also provide a summary of proposals by program.**	APPROPRIATE

2. Are awards appropriate in size and duration for the scope of the projects? **Comments:** • Overall, given the scope of each project, the awards are of appropriate size and duration. • SFS – The scholarships are standardized. The typical budget is for five years (8-13 scholarships). • GRFP – The program provides 3-year fellowships over a 5-year period; the fellowships include an annual $34,000 ($32,000 in 2013) stipend and an annual $12,000 (unchanged since 2013) cost-of-education allowance to the host institution. While responsive to inflation of cost-of-living, the cost-of-education allowance may not be keeping pace with rising tuition costs. • NRT – The large scope of the program may warrant larger investments. • PRIME – A limited number of projects were funded in 2015 for 2 or 3 years with an average annual budget of ~$250,000. **Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data EIS/Committee of Visitors Module. From the Report View drop-down, select Average Award Size and Duration.**	APPROPRIATE
3. Does the program portfolio include awards for projects that are innovative or potentially transformative? **Comments:** • Many, but not all the awards, are innovative and some are potentially transformative. • SFS and SaTC-EDU – In general, the projects reviewed are innovative and transformative. • GRFP – The review criteria are the same as all NSF grants (intellectual merit and broader impacts), and reviewers are specifically asked whether the proposed activities are potentially transformative. Although the applicants describe their proposed research plans in some detail, most of the reviews for successful applications tend to emphasize the past accomplishments of the applicants and their future potential. Subsequent awardee annual activities reports evidence that innovative and potentially transformative research is conducted. • GRFP students have little accountability to complete the work proposed; this may be something worth investigating.	APPROPRIATE

• PRIME – Though there were only a few awards provided, all were innovative in the field. **Data Source: Jackets**	
4. Does the program portfolio include inter- and multi-disciplinary projects? **Comments:** • A large percentage of the awards involve inter- and multi-disciplinary projects. • SFS and SaTC-EDU – By definition, cybersecurity is inter-/multi-disciplinary. • GRFP – The program solicitation specifically encourages inter- and multi-disciplinary projects, and ~40% of the funded projects fall into this category. The percent of inter-/multi-disciplinary proposals that are funded is proportional to the percent of the applicants that are inter-/multi-disciplinary. • NRT – Based upon the titles and/or abstracts, most of the projects appear to be inter-/multi-disciplinary. • IGE – Based upon the titles and/or abstracts, most of the projects appear to be inter-/multi-disciplinary. • PRIME – Approximately half of the limited number of awards may be considered inter-disciplinary. **Data Source: eJacket -COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data. If co-funding is a desired proxy for measuring inter- and multi-disciplinary projects, the Co-Funding from Contributing Orgs and Co-Funding Contributed to Recipient Orgs reports can be obtained using the EIS/Committee of Visitors Module. They are available as selections on the Report View drop-down.**	APPROPRIATE
5. Does the program portfolio have an appropriate geographical distribution of Principal Investigators? **Comments:** • The awards in each program are generally geographically dispersed, with the exception of GRFP (see the third bullet below). • SFS and SaTC-EDU – During 2014-2017, 188 awards were made in 39 out of 52 states and territories (including DC and Puerto Rico). However, only six states (AL, CA, NY, NC, PA, and TX) comprised 41% of the awards. • GRFP – Although awards were made in all 50 states, DC, and Puerto Rico, three institutions represented more than 20% of the awards,	APPROPRIATE

and less than 15 institutions represented over 50% of the awards, during the years for which data was provided (e.g., 2013-2018). • NRT – During 2014-2017, 54 awards were made in 30/50 states. • IGE – During 2015-2017, 35 awards were made in 21/50 states. • PRIME – While there are a limited number of awards to consider, a large percentage have been funded in California. **Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data**. **EIS/Committee of Visitors Module. Select Proposals by State from the Report View drop-down.**	
6. Does the program portfolio have an appropriate balance of awards to different types of institutions? **Comments:** • Given the focus of each program, the awards reflect an appropriate balance of different types of institutions. • SFS and SaTC-EDU – 25% of the awards were made to EPSCoR states, and 15% to MSIs. The majority of the awards (67%) were made to PhD institutions, with Master's institutions making up the majority of the rest (20%). The majority of the awards were made to public institutions (78%). • GRFP – Not applicable as the awards are made to individuals. • NRT – 20% of the awards were made to EPSCoR states, and 13% to MSIs. The majority of the awards were made to R1 (80%) public institutions (87%). No Tribal College was represented. • IGE – No Tribal College was represented. • PRIME – Insufficient data. **Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data. EIS/Committee of Visitors Module. Select Proposals by Institution Type from the Report View drop-down. Also, the Obligations by Institution Type will provide information on the funding to institutions by type.**	APPROPRIATE
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	APPROPRIATE

defined as someone within seven years of receiving his or her last degree at the time of the award. **Comments:** • Overall, an appropriate balance of awards have been made to new and early-career investigators. • SFS – 21% of the awards were made to new PIs. • GRFP – Not applicable as the awards are made to individual students. • NRT – Nearly all of the grants were awarded to established investigators, which is appropriate. • IGE – Nearly all of the grants were awarded to established investigators, which is appropriate. • PRIME – There are insufficient data to evaluate. **Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data. EIS/Committee of Visitors Module. Select Funding Rate from the Report View drop-down. After this report is run, use the Category Filter button to select New PI for the PI Status filter or New Involvement (PIs & coPIs) = Yes.**	
8. Does the program portfolio include projects that integrate research and education? **Comments:** • When appropriate, some of the programs in the portfolio include projects that integrate research and education. • SFS – Yes • GRFP – Yes. By design, this program integrates research and graduate education. • NRT – Yes • IGE – Yes • PRIME – Yes **Data Source: Jackets**	APPROPRIATE
9. Does the program portfolio have appropriate participation of underrepresented groups[3]?	APPROPRIATE

[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

Comments:

- There is evidence that there is appropriate participation of underrepresented groups.
- Veterans and people with disabilities remain highly underrepresented across all programs.
- There are no data on first generation college students or the socio-economic status of the students.
- SFS and SaTC-EDU – 15% of the awards were made to MSIs. Both minorities and women were involved in a significant percentage of the awards.
- GRFP – The 2012 GRFP COV report suggested that women and minorities were underrepresented in the program. That COV was concerned that measures of participation by women and underrepresented minorities (URMs) may have been distorted by overrepresentation of URMs within a sub-discipline while there was severe underrepresentation at the discipline level. It was suggested that this is an artifact of the affinity of URMs for specific sub-disciplines. NSF's response noted that it would continue to explore ways of appropriately balancing the portfolios of awards, including the use of electronic and social media. Between 2000-2010, the success rate of URMs was comparable to that of other demographic groups. Since 2010, the success rate of applications of URMs (as an aggregate group) grew significantly while the success rate of other demographic groups remained essentially unchanged. Since 2010, the success of applications from Asians was comparable to that of non-URM applicants, whereas the success of applications from American Indians or Alaska Natives, Blacks or African Americans, and Native Hawaiians or Other Pacific Islanders increased.
- Data on first-generation college status and socio-economic status (such as Pell-grant eligibility) of GRFP awardees were not available to the COV.
- The success rate of underrepresented minorities has increased, but the pool of underrepresented minorities remains the same. The program was responsive to the 2012 COV but was not successful in growing the pool of applicants.
- NRT – From the data for FY2014–2017, reasonable percentages of the PIs were women and underrepresented minorities.
- IGE – From the data for FY2015–2017, a preponderance of PIs were women and the vast majority were white.
- PRIME – Insufficient data.

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.

Data Source: eJacket-COV Module Documents: 4 – Awards Portfolio – Proposal & Award Data. **EIS/Committee of Visitors Module. Select Funding Rate from the Report View drop-down. After this report is run, use the Category Filter button to select Women Involvement = Yes or Minority Involvement = Yes to apply the appropriate filters.**	
10. Is the program relevant to national priorities, agency mission, relevant fields and other constituent needs? Include citations of relevant external reports. **Comments:** • As each program in the portfolio targets unique aspects of needs, we offer specific comments for each program below. • SFS – The 2011 report by the NSTC "Trustworthy Cyberspace: Strategic Plan for the Federal Cybersecurity Research and Development Program" recommended a research agenda to make cyberspace secure. Furthermore, the White House is engaged in the multi-agency National Initiative for Cybersecurity Education (NICE) program. The SFS is supportive of these initiatives in recognizing that cybersecurity education and workforce development are critical elements of cybersecurity research and development. • GRFP – The distribution of GRFP funds is skewed to large institutions with the infrastructure to support student applicants in the preparation of the application material. This appears to be in conflict with the mission of NSF of advancing "the capability of the *Nation* to meet current and future challenges." • IGE and NRT – STEM graduate education is poised to undergo substantive transformation, driven by a plethora of national studies, an accelerating pace of STEM discovery, national STEM workforce trends, globalization of STEM fields, and improved understanding of pedagogy (NAS, Board on Higher Education and Workforce. *Revitalizing Graduate STEM Education for the 21st Century*, 2018). The IGE and NRT programs recognized that the path forward to advancing STEM graduate education will involve new models for training graduate students. • PRIME – Federal agencies are being called upon to show that funding priorities are evidence-based and to provide plans for how both project and program evaluations demonstrate or validate impact and are used to support budget priorities (US Office of Management and Budget. "Next Steps in the Evidence and Innovation Agenda," Memorandum to the heads of Departments and Agencies, Document M-13-17, Washington, DC, 2013). **Data Source: Jackets, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National**	APPROPRIATE

Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission, eJacket-COV Module Documents: 4 – Awards Portfolio – National Priorities & NSF Mission	
11. Additional comments on the quality of the projects or the balance of the portfolio: • DGE should consider if the GRFP students are receiving mentoring and professional development.	

OTHER TOPICS

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

   - GRFP – The program needs modernization and to be made more strategic. Given evolving labor markets, GRFP awards might be distributed in areas of current and future national need, for example, as advised by the Bureau of Labor Statistics.
   - There needs to be more follow-up to determine effectiveness/return on investment. A mentorship plan would be beneficial in reporting outcomes during and after funding. The recipients of GRFP grants should be required to submit Independent Development Plans (IDPs) with their first annual report, signed by their mentor(s).
   - The unique challenges/traditions of individual disciplines and institutions are not recognized. A one-size-fits-all approach to addressing challenges in graduate education is unlikely to succeed. DGE is unique in that it is responsive to all NSF directorates. Several of the DGE programs require buy-in from other divisions. When a program is required to address the needs of many who have diverse needs, natural tension exists.
   - For GRFP, there needs to be an investment in post-award evaluation that builds upon the work NORC has already done in 2014.
   - DGE should consider implementing follow-up surveys of fellows/trainees/scholars in GRFP/NRT/SFS and institutions to cover five years past program participation. The goal would be to determine the impact of the programs on the fellows/trainees/scholars and the institutions that trained them.
   - Per data provided for 2013-2017, over 20% of GRFP funds are awarded to three institutions.
   - Overall, all the panel reviews could potentially benefit from a bibliometric analysis to ensure higher quality reviews and to improve consistency between ratings and the text of the reviews.
   - DGE's programs tend to have 'tribal knowledge' and could benefit from sharing best practices and templates across NSF.
   - Discipline-specific training needs to be considered in designing the DGE programs that span the entire directorate. In NSF's Strategic Plan (https://www.nsf.gov/about/performance/strategic_plan.jsp), please see the discussions of "A Nation that is the global leader in research and innovation" (NSF's vision) and "Advance the capability of the Nation to meet current and future challenges" (Strategic Goal 2).

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.

- DGE should reconsider the decision to limit the number of times an applicant can apply for the GRFP as a master's student. The change made for the 2017 competition, limiting to *one* the number of times a graduate student may apply, is likely disadvantaging some of the students it aims to advance and benefit. The master's degree is often a stepping stone for the PhD especially for those who take disadvantaged paths to the PhD (see, e.g., Lange 2006). The NSF GRFP is excluding students who are participating in bridge programs such as those funded by the NSF LSAMP BD program. Such programs are intended to provide additional avenues of access to individuals – including/especially women, URMs, and people with disabilities – who otherwise have been traditionally underutilized by top PhD programs (at least in part because most top PhD programs still rely heavily on GRE scores for admission; see, e.g., Posselt 2016). Specifically, the current GRFP eligibility requirement denies eligibility for individuals who have completed a master's degree unless they have experienced a long interruption before starting a PhD program. However, by design, bridging programs such as LSAMP BD and other NSF-supported bridge programs serve students who are intentionally using a master's degree as a continuous stepping stone to the PhD. Eligibility similar to other PhD students – i.e., in the first year or start of the second year of the PhD program – would be appropriate for such students. DGE might consider a tiered approach to the GRFP, allowing students to apply in the senior year of undergraduate studies, in the master's program, and in the first year of doctoral studies.
- In general, NSF should better leverage its existing programs at the undergraduate and introductory graduate levels; for example, INCLUDES, LSAMP, AGEP, TCUP, HBCU-UP, etc. Additionally, partner with other agencies and foundation programs with complementary goals, such as McNair, Switzer, Nielsen, AAAS – Entry Point!, etc.

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

- Most of the goals of the programs in DGE involve affecting systemic change on graduate education, yet there appears to be very little information available that assesses the effectiveness of each funded project beyond the period of funding (amplification, scalability, sustainability, etc.). Longitudinal studies are desperately needed in graduate education.
- One major opportunity to study impact is with regards to holistic review. NSF's elimination of the GRE requirement in GRFP was a tectonic shift at the time; it would be valuable to the scholarly literature and to the community to understand the longitudinal impacts of this move. DGE could potentially exert even more systemic change through continued leadership in setting expectations for holistic review. For example, could DGE set program eligibility requirements for institutions on the basis of holistic review practices?
- Reviewers need to be educated to recognize and reduce the impact of their personal biases to better identify potential and talent.
- NSF could take better advantage of natural pipelines used to identify talent and potential.
- NSF will likely not be able to accomplish these recommendations without an increase in staffing.
- The COV believes NSF might improve efficiency and effectiveness in monitoring awarded proposals through the implementation of data analytics techniques and tools. The portfolio reviewed has a significant amount of related data; use of advanced tools to analyze performance at both the individual award and portfolio levels would facilitate greater oversight by the Program Officer, Division leadership, Directorate leadership and NSF leadership.

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

- The current main outcome measure for DGE is students in graduate school, but the COV feels a more appropriate/impactful measure would be graduation and post-graduate employment.
- There needs to be greater outreach for the GRFP to grow the pool of students from four-year colleges and universities across the board.
- There needs to be exploration into why 50% of the GRFP students enroll in 10 institutions and whether this is beneficial to the program or if changes should be implemented.
- The GRFP strategic outreach plan should be implemented, evaluated, and regularly updated according to foundation priorities and available data. Adequate resources need to be allocated. This practice will aid in the mission of broadening participation in graduate programs of students from underrepresented and underserved populations.

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

- The COV review process has been problematic. DGE is nearly the last division at NSF to move from a program-specific to a division-wide COV review. However, it appears the DGE COV review was organized without guidance from the previous COV reviews, as evidenced by unrealistic expectations and a presentation of overwhelming information that can only be described as a “data dump.” Importantly, the DGE staff are not being criticized, because as expressed elsewhere in this report, it is clear the division is understaffed, and many of the key positions within the division are staffed with new or temporary personnel. The COV is a good example of where cross-divisional cooperation could improve the efficiency of such activities. What are the best practices associated with division-wide COV reviews?
- The template was serviceable, but too generic for addressing the unique issues associated with individual programs. The answer to all of the questions was “yes, but ….”
- The data that was presented to the COV was in desperate need of curation. In many cases (e.g., the individual jackets), the data was too raw to accept as representative of, for example, GRFP applications. In the end, the GRFP files were best organized of all of the programs, but only after requests that the data be reorganized.
- The members of the COV were originally told, when the members were recruited, that the work would be completed during the visit itself. Later, only a few days before the meeting was scheduled, the members were asked to do “homework” before the meeting. If the available information had been curated and well-organized, and if the questions and priorities for the COV had been articulated, it might have been possible to complete the requisite tasks in a single day. However, some members who reviewed material over several days – and only scratched the surface of the data provided – were told later their work was not the focus of the COV review. The “rules of engagement” need to be articulated with greater clarity from the get-go.
- The staff of the division know better than anyone what factors are impediments to achieving their goals. Self-evaluation and an even-handed presentation of the metrics that evidence said impediments would have positioned the COV to provide more useful feedback. This could be done in advance of the onsite COV.

The Committee of Visitors is part of a Federal advisory committee. The function of Federal advisory committees is advisory only. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the Advisory Committee, and do not necessarily reflect the views of the National Science Foundation.

SIGNATURE BLOCK:

_____

For the 2018 DGE Division-Wide COV
Rory A. Cooper, PhD
Chair