CORE QUESTIONS and REPORT TEMPLATE
for
FY 2006 NSF COMMITTEE OF VISITOR (COV) REVIEWS

Guidance to NSF Staff: This document includes the FY 2006 set of Core Questions and the COV Report Template for use by NSF staff when preparing and conducting COVs during FY 2006. Specific guidance for NSF staff describing the COV review process is described in Subchapter 300-Committee of Visitors Reviews (NSF Manual 1, Section VIII) that can be obtained at <www.inside.nsf.gov/od/oia/cov>.

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. Committee of Visitor (COV) reviews provide NSF with external expert judgments in two areas: (1) assessments of the quality and integrity of program operations and program-level technical and managerial matters pertaining to proposal decisions; and (2) comments on how the results generated by awardees have contributed to the attainment of NSF's mission and strategic outcome goals.

Many of the Core Questions are derived from NSF performance goals and apply to the portfolio of activities represented in the program(s) under review. The program(s) under review may include several subactivities 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 subactivities of the program, with the latter requiring more time but providing more detailed information.

The Division or Directorate may choose to add questions relevant to the activities under review. NSF staff should work with the COV members in advance of the meeting to provide them with the report template, organized background materials, and to identify questions/goals that apply to the program(s) under review.

Guidance to the COV: The COV report should provide a balanced assessment of NSF's performance in two primary areas: (A) the integrity and efficiency of the ***processes*** related to proposal review; and (B) the quality of the ***results*** of NSF's investments that appear over time. The COV also explores the relationships between award decisions and program/NSF-wide goals in order to determine the likelihood that the portfolio will lead to the desired results in the future. Discussions leading to answers for Part A 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.* Discussions leading to answers for Part B of the Core Questions will involve study of non-confidential material such as results of NSF-funded projects. The reports generated by COVs are used in assessing agency progress in order to meet government-wide performance reporting requirements, and are made available to the public. Since material from COV reports is used in NSF performance reports, the COV report may be subject to an audit.

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/covs.jsp.

FY 2006 REPORT TEMPLATE FOR
NSF COMMITTEES OF VISITORS (COVs)

The table below should be completed by program staff.

Date of COV: January 19-20, 2006
Program/Cluster/Section: Science, Technology, Engineering and Mathematics Talent Expansion Program (STEP)
Division: Division of Undergraduate Education (DUE)
Directorate: Directorate for Education and Human Resources (EHR)
Number of actions reviewed: Awards: 20 Declinations: 32 Other:
Total number of actions within Program/Cluster/Division during period under review: Awards: 77 Declinations: 641 Other:
Manner in which reviewed actions were selected: **Selected awards were those with award numbers ending in a "6" or a "2" (the COV Chair chose these numbers). Selected declines were all those with proposal numbers ending in "6".**

PART A. 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, and withdrawals) that were *completed within the past three fiscal years*. Provide comments for *each* program being reviewed and for those questions that are relevant to the program under review. Quantitative information may be required for some questions. Constructive comments noting areas in need of improvement are encouraged. Expanded guidance for COV Members. Part A1 contains seven questions about the quality and effectiveness of how the review of proposals was conducted. They address efficiency, alignment of reviews with the program goals, and quality of information provided.

A.1 Questions about the quality and effectiveness of the program's use of merit review procedures.

Provide comments in the space below the question**.** Discuss areas of concern in the space provided.

QUALITY AND EFFECTIVENESS OF MERIT REVIEW PROCEDURES	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE[1]
1. Is the review mechanism appropriate? (panels, ad hoc reviews, site visits) Comments: The review process used in this program included review panels, ad hoc	Yes

[1] If "Not Applicable" please explain why in the "Comments" section.

reviews and reviews by NSF program staff. It is our opinion that review panels were appropriate for the Type 1 proposals since these panels provided good input to NSF program staff from professionals in the field, ensured a diversity of expert opinion on the various aspects of the proposals, and permitted the program staff to benefit from the on-the-ground experiences of practitioners. When used, ad hoc panels for the Type 2 proposals were appropriate since the number of submitted proposals was small thus permitting the expense of convening a panel to be avoided. The subsequent in-house review, which included a consideration of institutional diversity (size, type, geographic location), was also deemed appropriate given the mission and goals of the program.	
2. Is the review process efficient and effective? Comments: The review process was efficient and effective. In 2002, 177 proposals were reviewed, and sixteen were awarded. In 2003, 198 proposals were received of which 192 were Type 1. Awards were issued for sixteen Type 1 proposals and one Type 2 proposal. In 2004, the STEP program received 173 proposals of which 163 were Type 1 classification. After ten of the Type 1 proposals were withdrawn, the review process yielded nineteen Type 1 and three Type 2 awards. In general, panels for Type 1 proposals had 4 to 7 reviewers. The ad hoc panels, which reviewed Type 2 proposals, consisted of 3 to 4 reviewers. Each panel considered approximately 11 to 13 proposals, grouped according to the type of applicant institution (two-year, four-year, doctoral, and master's). In 2005, the STEP program received 206 proposals, including 181 Type 1 and sixteen Type 2 proposals, all of which were reviewed by panels. The program awarded twenty-two Type 1 proposals and two Type 2 proposals. In FY2005 at least 81% of the proposals were processed within six months of the deadline or target date for the solicitation year. For FY2002, 2003, and 2004, 99%, 96%, and 94%, respectively, of the proposals were processed within six months of the deadline.	Yes

3. Do the individual reviews (either mail or panel) provide sufficient information for the principal investigator(s) to understand the basis for the reviewer's recommendation? Comments: Based on reading more than twenty individual reviews, it is our opinion that the preponderance of the reviews do provide sufficient information to allow the principal investigator to understand the basis for the reviewer's recommendation.	Yes
4. Do the panel summaries provide sufficient information for the principal investigator(s) to understand the basis for the panel recommendation? Comments: Based on reading more than twenty panel reviews, it is our opinion that the preponderance of the reviews do provide sufficient information for the principal investigator to understand the basis for the panel's recommendation.	Yes
5. Is the documentation for recommendations complete, and does the program officer provide sufficient information and justification for her/his recommendation? Comments: Based on reading more than twenty program officer recommendations, it is our opinion that the preponderance of the recommendations do provide sufficient information for the principal investigator to understand the basis for the panel's recommendation.	Yes
6. Is the time to decision appropriate? Comments: The time to decision for the majority of the proposals was well within the recommended six-month timeframe from receipt of proposal to posting final decision.	Yes
7. Additional comments on the quality and effectiveness of the program's use of merit review procedures: The committee has made several observations based on our consideration of the quality and effectiveness of the program's use of the merit review process: A review of the jackets leads us to conclude that the quality of reviews has improved over time. Reviewers appear to be more knowledgeable and understanding of the program goals and objectives and individual reviewers have provided more detailed and useful comments. Additionally, the language in the program solicitations has improved in clarity and specificity over the life of the program. The program staff is to be commended for conducting well planned and thorough internal reviews, and, in some instances, taking a position not fully recommended by the review panels.	

A.2 Questions concerning the implementation of the NSF Merit Review Criteria (intellectual merit and broader impacts) by reviewers and program officers.

Provide comments in the space below the question. Discuss issues or concerns in the space provided.

Part A.2 contains four questions all of which are designed to guide the COV in evaluating the attention given in the individual, panel, and NSF reviews to the two merit review criteria, intellectual merit and broader impacts. **Intellectual merit** means: How important is the proposed activity to advancing knowledge and understanding within its own field or across different fields? How well qualified is the proposer (individual or team) to conduct the project? (If appropriate, the reviewer will comment on the quality of prior work.) To what extent does the proposed activity suggest and explore creative and original concepts? How well conceived and organized is the proposed activity? Is there sufficient access to resources? **Broader impact** means: How well does the activity advance discovery and understanding while promoting teaching, training, and learning? How well does the proposed activity broaden the participation of underrepresented groups (e.g., gender, ethnicity, disability, geographic, etc.)? To what extent will it enhance the infrastructure for research and education, such as facilities, instrumentation, networks, and partnerships? Will the results be disseminated broadly to enhance scientific and technological understanding? What may be the benefits of the proposed activity to society? The first three questions are the same except each references a different level of review, individual, panel, and NSF program officer. Each question should be answered by reading examples of the type of review specified and determining if the review as written used the two criteria as the basis for the decisions made about the project being reviewed. The fourth question is open-ended and should include comments on any reservations or exemplary practices observed in relation to the use of review criteria in reviews at the three levels.

IMPLEMENTATION OF NSF MERIT REVIEW CRITERIA	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE[2]
1. Have the individual reviews (either mail or panel) addressed both merit review criteria? Comments: The committee reviewed a total of twenty-seven reviews written by individual reviewers. We examined the reviews to determine if they were based on the two merit criteria. Our conclusion is that most reviewers made some attempt to fit their comments under the merit and broader-impacts categories suggested by the NSF. However, it was difficult for the reviewers to do this in any consistent fashion.	Yes

[2] In "Not Applicable" please explain why in the "Comments" section.

2. Have the panel summaries addressed both merit review criteria? Comments: We examined six reviews written by five panels. This includes three reviews from August 2003 and August 2005, respectively. We examined the reviews to determine if they were based on the two merit criteria. Our small sample suggests that compliance in the panel summaries has improved over time. Two of the three panel reviews from 2003 made no explicit reference to the two criteria. However, all three reviews from 2005 organized their comments along the lines of the two criteria.	Yes
3. Have the *review analyses* (Form 7s) addressed both merit review criteria? Comments: We examined six review analyses (Form 7s) to determine if they were based on the two merit criteria. In each instance, the program officer made explicit reference to the two merit criteria.	Yes

4. Additional comments with respect to implementation of NSF's merit review criteria:

Based on our consideration of the use of the two merit criteria in the three different levels of reviews, we believe the following issues should be mentioned:

These two criteria seem to have been introduced with one type of grant proposal in mind and are being used here to evaluate a very different type of proposal in the STEP program.

In a traditional grant for a scientific discipline, a primary goal might be to "advance knowledge." A secondary goal would be to promote "teaching, training and learning."

For a Type 1 STEP grant, the primary goal is to promote "teaching, training, and learning," an activity that in the current taxonomy falls squarely under "broader impacts." As a result, reviewers and panels seemed to be unsure about how to apply the broad "intellectual merit" criterion to this kind of proposal.

Several reviewers seemed to translate "intellectual merit" as pertaining to "primary goals" of the project and "broader impact" as secondary goals because this is pattern for a traditional grant. Many typically used the intellectual merit section as the place to summarize the proposal and evaluate the likelihood of success. Others reviewers simply moved one level down in the hierarchy by addressing the specific questions that the foundation offers to illustrate intellectual merit: "advancing knowledge," "qualified … to conduct," "creative and original," "well conceived and organized." They addressed the main activity under the proposal in the second part in response to the "teaching, training, and learning," question under broader impact.

More guidance for reviewers and panels about how to apply these criteria and where in a review to put different parts of the discussion might be useful.

A.3 Questions concerning the selection of reviewers. Provide comments in the space below the question. Discuss areas of concern in the space provided.
Part A.3 contains five questions related to the selection of reviewers. The first three questions address the adequacy of the numbers of reviewers, their expertise, and their diversity. The fourth question asks about recognizing and resolving conflicts of interest and the last question is open-ended asking for any issues that need to be identified about the selection of reviewers.

SELECTION OF REVIEWERS	YES , NO, DATA NOT AVAILABLE, or NOT APPLICABLE[3]
1. Did the program make use of an adequate number of reviewers? Comments: The program made use of 318 reviewers (not all of whom were unique) for 745 proposals. It is our opinion that this number of reviewers was adequate.	Yes
2. Did the program make use of reviewers having appropriate expertise and/or qualifications? Comments: Based on our examination of the qualifications of the 318 reviewers and the relationship of their qualifications to the goals of the program, it is our opinion that the preponderance of the reviewers did have appropriate expertise.	Yes
3. Did the program make appropriate use of reviewers to reflect balance among characteristics such as geography, type of institution, and underrepresented groups?[4] Comments: We examined the characteristics of 318 reviewers, not all of whom were unique. Reviewers were from a variety of geographic regions of the U.S. including 43 different states and the Virgin Islands and Puerto Rico. By self-report, reviewers were from a variety of types of institutions including two-year colleges (51), doctoral institutions (128), masters institutions (81), baccalaureate institutions (40), industry (5) and other (5). Forty-seven reported that they were minority, 2 reported that they had disabilities, and 122 reported that they were female. The program would benefit from including individuals from two-year colleges panel in order to provide a better balance since many of the proposals involve	Yes

[3] If "Not Applicable" please explain why in the "Comments" section.
[4] Please note that less than 35 percent of reviewers reported their demographics last fiscal year, so the data may be limited.

partnerships between two-year and four-year colleges as well as between high schools and colleges. The programs that involve collaboration with community colleges should have two community college reviewers on it. Some of the panels that were examined did not have as much geographic diversity as might be desirable.	
4. Did the program recognize and resolve conflicts of interest when appropriate? Comments: We reviewed 20 proposals, and 1 of conflicts of interest, and determined that the program did resolve conflicts of interest appropriately. Because of the small number of conflicts, no generalizations are appropriate. Of the proposals that were reviewed there was only one in which there was a statement of potential conflict of interest. There was an acknowledgement that the reviewer resided in the area of the college in question, Disclosure of the proximity and possible conflict was an appropriate resolution.	Yes
5. Additional comments on reviewer selection: Based on our consideration of the selection of the reviewers, we believe the following issues should be mentioned: Reviewers should be encouraged to provide more thorough reviews that highlight the strengths and provide concrete suggestions when possible to overcome weaknesses.	

A.4 Questions concerning the resulting portfolio of awards under review. Provide comments in the space below the question. Discuss areas of concern in the space provided. Part A.4 contains thirteen questions relating to the quality and balance of projects that received funding. The first question asks about the overall quality of the projects supported. The next nine questions relate to the balance of the projects in terms of size and duration, risk, multidisciplined, innovation, types of groups, experience of investigator, geographic distribution, institution types, integration and disciplines. Question 11 asks about the participation of under-represented groups, and Question 12 asks about the relevancy of the projects in terms of national priorities, agency mission, and relevant fields. The final question is open-ended, asking about any concerns relevant to the quality of projects or the balance of the portfolio.

RESULTING PORTFOLIO OF AWARDS	APPROPRIATE, NOT APPROPRIATE[5], OR DATA NOT AVAILABLE
1. Overall quality of the research and/or education projects supported by the program. Comments: Based on our review of twenty project grant jackets, we believe that the overall quality of the research and/or education projects supported by the program is appropriate.	Appropriate
2. Are awards appropriate in size and duration for the scope of the projects? Comments: In total, we reviewed twenty projects to determine the size and duration of awards. Project funding amounts ranged from approximately $250,000 to $2,000,000. The time spans for projects ranged from 3 to 5 years, with most projects lasting 5 years. It is our opinion that the projects appear appropriate both in terms of award amounts and duration.	Appropriate
3. Does the program portfolio have an appropriate balance of: • Innovative/high-risk projects?[6] Comments: Most of the twenty projects we reviewed had some element of innovative/high risk when defined in terms of current *best practices*. This includes practices which may be common in some sectors of higher education but considered innovative/high risk to the particular grantee institution. It is our opinion that this type of innovative/high risk is appropriate for this program. That is, the awardee institution intends to introduce *best practices* not previously used at the institution.	Appropriate

[5] If "Not Appropriate" please explain why in the "Comments" section.

4. Does the program portfolio have an appropriate balance of: • Multidisciplinary projects? Comments: We determined that many of the twenty projects that were reviewed were multidisciplinary. That is, the primary investigators and co-primary investigators were from different disciplines and the topics covered were often multidisciplinary. Additionally, people from many different groups and different institutions were involved. It is our opinion that the number of multi-disciplinary projects is appropriate for this program because the effort to increase STEM participation often requires an institutional perspective and not just a single department/unit commitment.	Appropriate
5. Does the program portfolio have an appropriate balance of: • Funding for centers, groups and awards to individuals? Comments: We reviewed 20 projects and determined that most involved funding for groups (multiple disciplines and/or institutions) and few, if any, were funding for individuals. It is our opinion that this type of participation is not only appropriate for this program, but may even be necessary for both immediate and sustained success.	Appropriate
6. Does the program portfolio have an appropriate balance of: • Awards to new investigators? Comments: We reviewed 20 projects. Most (if not all) of those awarded had lead primary investigators with previous funding experience. Given the nature of this program, this outcome is not unexpected.	Appropriate
7. Does the program portfolio have an appropriate balance of: • Geographical distribution of Principal Investigators? Comments: We reviewed 20 projects from fifteen principal investigators from fifteen different states in three to four different regions of the country. This appears to be an appropriately balanced group of Principal Investigators.	Appropriate

[6] For examples and concepts of high risk and innovation, please see Appendix III, p. 66 of the Report of the Advisory Committee for GPRA Performance Assessment, available at <www.nsf.gov/about/performance/acpga/reports.jsp>.

8. Does the program portfolio have an appropriate balance of: • Institutional types? Comments: Of the twenty projects we reviewed, nine were from research 1 institutions, eight were from 4-year colleges or universities and three were from 2-year colleges. Given the number of 4-year and 2-year institutions in the country compared to the number of research 1 institutions, it is our opinion that the program might want to seek methodologies that will increase awardee representation from among the former group of institutions.	Appropriate
9. Does the program portfolio have an appropriate balance of: • Projects that integrate research and education? Comments: Most of the twenty projects reviewed had some element of integrated research and education activity. This discovery is not surprising given that undergraduates' involvement in research is now a widely accepted retention and persistence to graduation factor in STEM fields.	Appropriate
10. Does the program portfolio have an appropriate balance: • Across disciplines and subdisciplines of the activity and of emerging opportunities? Comments: Of the twenty projects reviewed, many had some elements that cut across disciplines and subdisciplines of the activity and of emerging opportunities. In general, given the goals of this program, it was not unexpected to discover these elements.	Appropriate
11. Does the program portfolio have appropriate participation of underrepresented groups? Comments: Some of the reviewed projects had an exclusive focus on underrepresented groups and many included, among other elements, an intent to include underrepresented groups.	Appropriate
12. Is the program relevant to national priorities, agency mission, relevant fields and other customer needs? Include citations of relevant external reports. Comments: All of the projects addressed factors that support the goals of STEP and, as such, are relevant to national priorities, relevant fields, and other customer needs and the agency mission.	Appropriate

13. Additional comments on the quality of the projects or the balance of the portfolio:

Based on our consideration of the quality of the projects and the balance of the portfolio, we believe the following issues should be mentioned:

Many of the proposals reviewed shared common themes including summer bridge programs; institutional partnerships (particularly 4-year colleges/universities with community colleges); mentoring; focus on underrepresented groups; curricular reform; high school outreach; undergraduate research; constructive intervention; multi-faceted approach.

A.5 Management of the program under review. Please comment on:
Part A.5 contains four questions relating to the overall management of the program under review. The first question asks about the overall management of the program. The second question asks about the responsiveness of the program to emerging research and education opportunities. Question 3 asks about the planning and prioritization process (both internal and external) which guided the development of the program. Question 4 is open-ended, asking about any concerns relevant to the management of the portfolio.

1. Management of the program.
Comments:
We reviewed approximately 30 of project annual reports. We also considered the program solicitations and announcements, discussions with program officers, and program/division annual reports. Finally, we considered if the program included any of the following: pre- and post-award site visits and reports; outreach workshops; PI meetings and reports; fiscal management workshops; and mid-point reviews. Based on these items, we determined management of the program is effective and efficient. As is the case with some other NSF DUE programs that we are familiar with, the program officers who have worked with STEP institutions do an enormous amount of detailed work in support of their grantees, particularly before grants are awarded.

The STEP Management Plan gives a good top-line overview and serves as a basis for continuous improvement. For instance, NSF could provide a more detailed outline of the STEP program solely for the purposes of internally assessing if program management is done effectively. The design of the program could also better facilitate the collection of leading and lagging success indicators. For example:

The STEP Program is a strategic and tactical implementation of the "Technology Talent Act of 2001" and is designed to *increase in the number of students (US citizens or permanent residents) pursuing and receiving associated and bachelors degree in established or emerging fields within STEM.* DUE's management of the program includes the following responsibilities and success metrics:

Program Design

- Establishing SMART (specific, measurable, actionable, realistic and time-bound) goals and objectives, and annual success metrics/milestones for the program (metrics: budget, qualitative, quantitative, incremental increase in degrees, etc.)
- Developing the Program Solicitation, procedures and processes (metrics: clear requirements, effective process, etc.)
- Communicating and Marketing STEP opportunities to institutions and faculty (i.e. target audience: solicitation, e-mail, e-news, conferences, etc.), (metrics: # of proposal, demographics, etc.)

Proposal Processing and Review

- Coordinating and managing the solicitation process (metrics: deadlines, cycle time, etc.)
- Coordinating and managing the peer-review process (metrics: participation, process outputs, etc.)
- Managing the NSF internal selection and award process, including pre-site visits (metrics: cycle time, awards, rejects, etc.)

Tracking and Reporting

- Monitoring the progress of funded projects, site visits, workshops, PI meetings, reports, (metrics: reports, milestones, data, mid-point reviews, etc.)
- Managing the grant or projects closure process.
- Assessment of the program, e.g. COV, Post-site visits, etc.
- Report of outcomes vs. goals and objectives

The program description in the solicitations were revised yearly ('02 thru '04). The refinements seem to improve clarity of scope and requirements and also helped to identify possible leading and lagging indicators of success. These mentioned indicators should be measured (not just mentioned). The '04 and '05 metric requirement of "…specific numerical targets for these increases" should also apply to the overall program metrics. That is, NSF should also have overall numerical target increases in respect to total spending, even if only used internally.

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

There were six Type II proposals awarded. We reviewed 1 of the 6 projects in our packets, where NSF followed-up with panel reviewers and determined that the proposal was much stronger than rated by the panel. The proposal was funded at the amount and durations proposed. The program solicitations and announcements, program officers' expertise, and program/division annual reports provide evidence that the program's management has been responsive to emerging research and education opportunities.

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

Based on our review of program solicitations and announcements, interviews with key program

personnel, distribution of funds, and (cite any other evidence used), we determined the program planning and prioritization process is appropriate for guiding the development of the portfolio. (Cite examples.)

4. Additional comments on program management:

The STEP program as with other NSF programs requires an enormous amount of resources to design, re-design, plan, execute and track. We feel the NSF staff is effective in managing this program. Continuous improvement is best accomplish when goals are specific and measurable, and aimed at delivery of the desired outcomes. We feel that continued investment in the development of NSF staff to improve strategy planning, implementation, management, measurement skills and tools would allow for more innovation in how to manage and measure the success of the program. Examples of such management tools include Six-sigma certification (DMAIIC), Kepner-Tregoe Analysis (Potential Problem Analysis), etc.

PART B. RESULTS OF NSF INVESTMENTS

NSF investments produce results that appear over time. The answers to the first three (People, Ideas and Tools) questions in this section are to be based on the COV's study of award results, which are direct and indirect accomplishments of projects supported by the program. These projects may be currently active or closed out during the previous three fiscal years. The COV review may also include consideration of significant impacts and advances that have developed since the previous COV review and are demonstrably linked to NSF investments, regardless of when the investments were made. Incremental progress made on results reported in prior fiscal years may also be considered.

The following questions are developed using the NSF outcome goals in the NSF Strategic Plan. The COV should look carefully at and comment on (1) noteworthy achievements of the year based on NSF awards; (2) the ways in which funded projects have collectively affected progress toward NSF's mission and strategic outcomes; and (3) expectations for future performance based on the current set of awards. NSF asks the COV to provide comments on the degree to which past investments in research and education have contributed to NSF's progress towards its annual strategic outcome goals and to its mission:

- To promote the progress of science.
- To advance national health, prosperity, and welfare.
- To secure the national defense.
- And for other purposes.

Excellence in managing NSF underpins all of the agency's activities. For the response to the Outcome Goal for Organizational Excellence, the COV should comment, where appropriate, on NSF providing an agile, innovative organization. Critical indicators in this area include (1) operation of a credible, efficient merit review system; (2) utilizing and sustaining broad access to new and emerging technologies for business application; (3) developing a diverse, capable, motivated staff that operates with efficiency and integrity; and (4) developing and using performance assessment tools and measures to provide an environment of continuous improvement in NSF's intellectual investments as well as its management effectiveness.

B. Please provide comments on the activity as it relates to NSF's Strategic Outcome Goals. Provide examples of outcomes (nuggets) as appropriate. Examples should reference the NSF award number, the Principal Investigator(s) names, and their institutions.

B.1 <u>OUTCOME GOAL for PEOPLE</u>: Developing "a diverse, competitive and globally engaged workforce of scientists, engineers, technologists and well-prepared citizens."

Comments:

The principal objective of the STEP program is to explore ways in which the infusion of NSF funds can be used to increase the number of STEM majors. As is noted in (of all places) the enabling legislation and the program solicitations (NSF program solicitation 05-519 page 5) existing research and practical experience contains many suggestions on how STEM educational programs could be changed or supplemented in order to make them more attractive. STEP institutions will implement some existing, or some new, ideas in order to increase the number of STEM graduates at those institutions. The 19 institutions who were awarded STEP grants in FY2004 project that 1640

more STEM students will graduate per year at the end of the five-year award period. (This number comes from statements in tab 8, page 23, of our briefing book.) The additional investment per year per student is about $625. The approximate $625 figure does not, of course, reflect the total cost of educating that STEM student, once she or he has been enticed to major in a STEM discipline. Given the present scope of the STEP program, these investments, while very worthwhile, are still insufficient to be really noticeable in national statistical terms.

We note that there are indirect impacts of STEP which, though difficult to measure, are still worth commenting on. Students who participate in STEP activities develop a more positive image of science and technology, even if they don't major in a STEM discipline. These students can affect the technology sector of the U.S. economy by becoming part of the sales force which will increase the market for technologically advanced products, some of the venture capitalists who will invest in companies that make these products, some of the managers who will run and found such companies, and most importantly some of the consumers who will buy and use these products.

STEP has been a successful beta test. It shows that strategies for increasing STEM student enrollment will succeed at a variety of institutions. Earlier projects were mostly undertaken by the small group of pioneers who developed these strategies or had other reasons for being personally invested in them. STEP investigators are not all pioneers. They have implemented these projects successfully at a reasonable cost

B.2 OUTCOME GOAL for IDEAS: Enabling "discovery across the frontier of science and engineering, connected to learning, innovation, and service to society."

Comments:

There has been a great deal of research on why students drop out of STEM majors, and there is a growing body of literature about best practices. The NSF program officers know about this research and have done an excellent job in disseminating it to potential proposers. Another approach is to ask why students persist in STEM, to ask what does work. A currently funded STEP project is investigating that question at the University of Oklahoma. We are aware of, but less familiar with, a number of the other research and scholarly projects which have been conducted under the overall STEP umbrella (referred to within the STEP program as "type 2" proposals) and find them both interesting and worthy of support. We recommend continued operation of and possible expansion of the research and scholarly component of STEP.

B.3 OUTCOME GOAL for TOOLS: Providing "broadly accessible, state-of-the-art S&E facilities, tools and other infrastructure that enable discovery, learning and innovation."

Comments:

Tools, as defined in this box, are not one of the objectives of the NSF STEP program.

B.4 OUTCOME GOAL for ORGANIZATIONAL EXCELLENCE: Providing "an agile, innovative organization that fulfills its mission through leadership in state-of-the-art business practices."[7]

Comments:

We have nothing particular to add in this area, other than a general appreciation for the way that NSF manages its business.

[7] For examples and further detail on the Organizational Excellence Goal, please refer to pp. 19-21 of NSF's Strategic Plan, FY 2003-2008, at <http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf04201>.

PART C. OTHER TOPICS

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

C.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.

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

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

The COV discussed a number of additional suggestions to the NSF which deal with issues C1, C2, and C3 above. Rather than try to artificially divide our comments among those three areas, we are presenting our thoughts as the answers to a series of questions that are specific to the STEP program.

1. Do the existing funding levels appear to allow for a reasonable mix of projects of the appropriate scope at a good mix of institutions?

Given the current funding level of the STEP program, we believe there is a reasonable mix of projects of the appropriate scope at a good mix of institutions (see question A.4, subquestions 6-10). If funding to the STEP program were to expand significantly, in addition to funding more PIs who write to the current program goals, STEP might want to consider soliciting projects that require more than two million dollars for full implementation. Such projects could, for instance, be ambitious, institution-wide projects located at some of the country's very largest universities, multi-institutional collaborations among universities in an entire state or large urban region, or projects involving extensive data collection.

2. Might projects be sufficiently significant under lower funding levels?

We believe that there may be such projects. NSF's experience (see question under part B, PEOPLE) indicates that a regular STEP program, implementing existing ideas, requires an investment of $500-$1000 per student per year. Suppose a person working at a smaller institution wants to do a STEP program that would result in an increase of 20 students per year. If this project came to the NSF with the suggested small-institution budget of $0.5 million for five years, it would be a hard sell. It would represent an investment of $5K per student per year, about 8 times the nominal average. While we do not suggest the establishment of a strict cost/student ratio, such a proposal would be much stronger if the budget were pruned to a fraction of the nominal $0.5 million.

We are concened (see question 4 below) that if this cost per student number were included in program solicitations that it could constrain proposal types and be taken seriously. However, program solicitations, informal conversations between program officers and prospective PI's, and other communication mechanisms should make it clear that proposals of smaller scope are welcome and appropriate. Sometimes it is the small institutions, or small-scope proposals at larger institutions, which generate big ideas.

3. When the currently funded projects reach the end of their 5-year grant periods, what types of opportunities should exist for them to seek an additional grant?

In general, we would expect institutions that have had grants for five years to sustain those grants using local resources. We can visualize a number of cases where there would be exceptions.

If an institution has a new idea for a project to accomplish the objectives of STEP they might submit a proposal to implement the new project. It should be substantially different from the first project and there should be assurances that the first project, if effective, will be sustained.

An institution might be involved with another institution's proposal to implement their original project. In most cases the principal investigator would come from the institution that is starting to implement the project. If an institution is working with a number of institutions to implement their original project, it is possible that they would submit the proposal since they would be facilitating implementation for multiple institutions. This would be a model that is used by the Gates Foundation with the implementation of Early and Middle College programs.

An institution might be eligible for a smaller grant for the ongoing collection of data that would be indicators of the success of their initial project. Although the number of STEM majors is the important indicator, as discussed in part 4 below, it would be useful to define and collect indicators along the way to graduation that might inform whether or not a project was on track to increasing the number of STEM graduates.

4. Has the NSF developed an appropriate set of quantitative metrics which can be used to manage the program, to respond to congressional inquiries, and to manage projects at individual institutions?

Incremental Investment per additional STEM major:

The very last pages of our briefing book contain an estimate from the NSF of the incremental institutional investment per additional STEM major promised in the grants funded so far. We have explored, both with the Foundation officers and among ourselves, the appropriate uses of this number. As indicated in part B, we believe that establishing this kind of rough estimate of the required investment per incremental STEM major is a very valuable output from the STEP program. To the committee, it suggests that feasible levels of funding (albeit ones that are higher than the current pilot levels) could, for example, bring the US back into a position of worldwide leadership in terms of the fraction of 24 year olds who receive a first degree in science. For example, if we round up the required investment to $1000 per additional student per year and assume that it takes 4 years of study (perhaps over more than 4 calendar years) for a major to graduate, $1 billion per year would generate 250,000 additional STEM majors per year.

Going forward, the committee hopes that the foundation will continue to track this number and update it as information comes in about what institutions actually do, not just what they promise. Presumably, this will be an important part of the mandate for ORC.

The STEP program was able to solicit accurate information about costs in part because it gave proposers wide latitude in their proposals and then selected proposals for funding at least partly on the basis of their cost effectiveness. We agree with the NSF staff that publicizing the existing average incremental investment per student or forcing all institutions to conform to a particular investment per student could undermine the foundation's ability to continue to solicit accurate information. In particular, we are not persuaded that any information about the average or range of

values for the investment per incremental student should be highlighted in request for proposals or other information provided to proposers.

However, we do believe that it might be helpful for proposers to be told that cost effectiveness in this sense is one of the criteria used to evaluate proposals and that the program officers could experiment with telling reviewers the range of values for accepted proposals so that they have a benchmark for evaluating the proposal assigned to each of them.

Intermediate indicators

Because it takes many years for a program to move to a new steady state level of majors graduating per year, it would be useful for institutions and the foundation to have some intermediate indicators about the effectiveness of the interventions being undertaken. One such indicator might be a measure of the retention rate in or attrition rate out of a STEM degree track. Another might be a measure of intended major for students at various stages in their careers.

We suggest that the program officers work with ORC and the various PIs to work toward a set of intermediate indicators. If there were some consistency of these indicators across programs, it would clearly help the foundation and scholars interested in these interventions to evaluate how they work. One forum where these might be discussed might be the annual meeting of the PIs.

5. Is the scope of the evaluation contract, called for in the enabling legislation and recently awarded to the contractor (QRC, a division of Macro, Inc.) appropriate?

We think so, but we do have some suggestions.

Overall, the contract RFP is well defined and clear in describing the program and both the Type I and Type II projects. The SOW explains NSF's expectations of the contractor role clearly. Examples of the outlined deliverables would better help the contractor provide the appropriate information. The verification and validation plan for the web-based instrument and electronic database should be reviewed in-advance to assure sufficient design inputs are considered.

We note that on the NSF's suggested work statement, there were lines on the form (sample chart) which permitted institutions to list other majors than those which were specifically listed. The option of "other" does not seem to exist within the proposed web-based instrument submitted by ORC. The instrument should account for someone who graduated with a degree in materials science, or someone with a joint major in biotechnology and computer science, or someone with Allied Health degrees in fields like nuclear medicine technology.

A second suggestion from the Committee is that ORC collect data on leading and lagging indicators that correlate to the desired outcomes of STEP, i.e. "increases in STEM degrees earned". For example, the contract should monitor retention rates from the freshman to the sophomore year. Many STEP projects concentrate their efforts on the freshman year and so such data, which can be gathered relatively early in a project's existence, can provide prompt feedback on the effectiveness of a project. Some individual committee members might want to parse these data even more finely and gather information on the transition between the first semester and the second semester of the freshman year; we leave it to the NSF program officers to decide.

6. How can the NSF best utilize the peer review system, which we accept as an integral part of the NSF management strategy, to make decisions on an implementation program?

We noticed in our reviews that the peer review system has produced reviews for STEP that are sometimes less useful than is the case for other programs within NSF. The STEP program officers whom we talked to about this issue agreed with our assessment. The characteristics of non-useful individual reviews varied some, but included such examples as general, uncritical, unsupported statements of great praise, restatements of the abstracts accompanied by blandly enthusiastic endorsements, brief reviews that were two sentences long, and the like.

We discussed various reasons for the causes of the relatively greater prevalence of such reviews and, while we discussed various possible solutions, we do not see this problem as going away. Perhaps an increased emphasis during the orientation session with reviewers on the usefulness of review comments may be helpful. As discussed in Parts A2 and C4, it might also help to provide more information and guidelines to reviewers.

The NSF already seeks to include a variety of people on panels. They already seek to include panelists who have experience in actually operating programs (such as Deans and Provosts). We encourage program officers in the STEP program to expand even further their search for review panelists. Another source of panelists is the pool of people who lead the diversity efforts within various professional societies.

We did notice that program officers have exercised discretion in selecting STEP proposals for funding, by declining programs with strong reviews and supporting programs which have relatively modest reviews. We endorse and encourage this process.

We encourage the practice within DUE of having a small group of program officers review a relatively homogeneous group of proposals before final decisions have been made. We note that program officers have often been very appropriately selective in their choice of comments of individual reviewers in order to support a funding decision. In those cases, we found ourselves agreeing with the assessment of program officers compared to the panel consensus ranking.

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

SIGNATURE BLOCK:

_____

For the STEP COV
Harry Shipman
Chair

MEMORANDUM

DATE: December 31, 2006

TO: Bernice Anderson**,** Senior Program Director for Evaluation
Directorate for Education and Human Resources

FROM: Susan Hixson, Lead Program Director Program (EHR/DUE)
Science, Technology, Engineering, and
Mathematics Talent Expansion Program (STEP)

SUBJECT: Science, Technology, Engineering, and
Mathematics Talent Expansion Program (STEP)
COI and Diversity Memo

The Committee of Visitors report for STEP was approved at the EHR Advisory Committee meeting held at NSF on November 1, 2006. The COV consisted of six members selected for their expertise related to the goals of the program. They provided a balance with respect to the type of institutions supported through the program, gender, and representation from underrepresented groups. The following table shows the main features of the COV's diversity.

Category of COV Membership	**No. of COV Members in Category**
Member of EHR Advisory Committee……	……1…….
Institution Type:	
❑ University…(2 PhD, 1 Masters)	……3…….
❑ Four-year College……	……1…….
❑ Two-year College……	……1…….
❑ K-12 School or LEA……	……
❑ Industry……..	……1…….
❑ Federal Agency……	……
❑ Non-profit Organization ……	……
❑ Consultant ……..	……
Location	
❑ East……..	……3……
❑ Midwest/North ……	……1……
❑ West……	……1……
❑ South……	……1……
Gender	
❑ Female……	……2……
❑ Male……	……4……
Race/Ethnicity	
❑ White……	……4……
❑ Black……	……2…
❑ Hispanic……..	……
❑ Asian……	……
❑ Pacific Islander……..	……

The COV was briefed on Conflict of Interest issues and each COV member completed a COI form. Individual COV members neither read nor discussed any proposals or files with which they had conflicts.