FINAL REPORT 10/29/2009
For
FY 2009 NSF REESE COMMITTEE OF VISITOR (COV) REVIEWS

Guidance to NSF Staff: This document is the FY 2009 REESE NSF Committee of Visitors Final Report of the REESE Program. The COV followed the specific guidance for the COV review process as described in Subchapter 300-Committee of Visitors Reviews (NSF Manual 1, Section VIII) at <www.inside.nsf.gov/od/oia/cov>.

FY 2009 REPORT FOR REESE
NSF COMMITTEES OF VISITORS (COVs)

The NSF Program staff completed the table below.

Date of COV: October 22-23, 2009	
Program/Cluster/Section: REESE	
Division: DRL	
Directorate: EHR	
Number of actions reviewed: 51 **Awards: 23** **Declinations: 28** **Other: 0**	
Total number of actions within Program/Cluster/Division during period under review: 915 **Awards: 166** **Declinations: 691** **Other: 58**	
Manner in which reviewed actions were selected: **A REESE program staff member rolled a 20-sided die, which yielded an 8. All awards ending in 8 were selected, and every fourth declination ending in 8 was selected (or their respective collaborative, if appropriate). Additionally, the REESE Program Officers identified a small number of awards (8) that they wished the COV members to consider. Proposals were removed if COV members were PIs or co-PIs. The COV members were informed as to which proposals were selected randomly, and which proposals were identified by REESE Program Officers.**	

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.

A.1 Questions about the quality and effectiveness of the program's use of merit review process. Provide comments in the space below the question**.** Discuss areas of concern in the space provided.

QUALITY AND EFFECTIVENESS OF MERIT REVIEW PROCESS	YES, NO, DATA NOT AVAILABLE, or NOT APPLICABLE[1]
1. Are the review methods (for example, panel, ad hoc, site visits) appropriate? **Comments:** The COV found that review methods for the sampled jackets are appropriate but noted that the number of reviewers varied.	YES
2. Are both merit review criteria addressed a) **In individual reviews?** Usually, but not in every case b) **In panel summaries?** Yes, the panel summaries always address both criteria. c) **In Program Officer review analyses?** Yes, the COV found that the analyses were clearly present in all proposals they reviewed. **Comments:** We reviewed a subset of proposals that had two or more reviews which did not specifically address either Intellectual Merit or Broader Impact or both. Although not all proposals had panel summaries, those that did addressed both review criteria. It appears that some proposals did not have panel summaries because panel reviews were low or the reviews were ad hoc.	YES

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

3. Do the individual reviewers provide substantive comments to explain their assessment of the proposals? **Comments:** Most of the individual reviews showed that REESE reviewers provided substantive comments per the reviewer instructions. We found no noticeable relationship between the rating of the review and whether reviewers made substantive comments. Each proposal we examined had at least some reviews with substantive comments. The COV recommends providing guidelines to new panelists along with sharing examples of exemplary reviews.	YES
4. Do the panel summaries provide the rationale for the panel consensus (or reasons consensus was not reached)? **Comments:** The panel summaries include the rationale when consensus has been reached. Although we acknowledge the taxing program manager work loads, we feel that panel summaries could do a better job of more consistently explaining panel decisions in light of divergent reviews.	YES
5. Does the documentation in the jacket provide the rationale for the award/decline decision? (Note: Documentation in jacket usually includes context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), program officer review analysis, and staff diary notes.) **Comments:** The documentation in the jacket includes the program officer analyses which were very insightful and typically the most detailed and thoughtful elements in the review materials. Clearly the program officers were not simply summarizing the panel reviews and discussion but were discussing the proposals in light of the reviews. In some cases the program officers had consulted additional experts and this was noted. The POs seemed to exert great effort to give each proposal a thorough reading and analysis.	YES

6. Does the documentation to PI provide the rationale for the award/decline decision?	YES

(Note: Documentation to PI usually includes context statement, individual reviews, panel summary (if applicable), site visit reports (if applicable), and, if not otherwise provided in the panel summary, an explanation from the program officer (written or telephoned with diary note in jacket) of the basis for a declination.) **Comments:** Our analysis indicated that PIs received adequate explanation of funding decisions through reviews, panel summaries, and program officer communications. The communications were constructive and likely to improve the quality of the work whether or not the proposals were funded. Perhaps NSF would benefit from gathering information about resubmitted proposals.	
7. Is the time to decision appropriate? Note: Time to Decision --NSF Annual Performance Goal: **For 75 percent of proposals, inform applicants about funding decisions within six months of proposal receipt or deadline or target date, whichever is later**. The date of Division Director concurrence is used in determining the time to decision. Once the Division Director concurs, applicants may be informed that their proposals have been declined or recommended for funding. The NSF-wide goal of 75 percent recognizes that the time to decision is appropriately greater than six months for some programs or some individual proposals. **Comments:** In 2007, 29% of proposals took more than six months from submission to decision. In other years (2005, 2006, 2008) dwell time has been at 25% or less, with a range from 3% to 25%. The average over four years was 19% dwell time over six months. The COV recommends that NSF take measures to further reduce the dwell time.	NO

8. Additional comments on the quality and effectiveness of the program's use of merit review process:

Some variability in the quality of reviews is probably inevitable, so a key question is whether the reviews of a proposal, taken together, are thorough and fair and provide a sound basis for decision. When the differences were pronounced, the program officers seemed to do a good job of pulling together the judgments of reviewers, highlighting the most salient comments, and reconciling divergent reviews.

We are concerned that reviewers are inconsistent in attending to methodological rigor along

with theoretical grounding and impact on practice. Regarding methodological rigor, the question is whether the methods proposed are adequate to address the research questions posed in the proposal. We recommend that reviewers be advised specifically to attend to the quality of research methods as one component of intellectual merit and in balance with the quality of the questions asked and the theoretical frameworks used.

We reflected on the process in which the program officers appear to have considerable discretion in compiling the judgment of reviewers and making funding decisions, particularly for proposals judged "competitive" (as opposed to highly competitive or not competitive). Among the cases we reviewed, we did not find any in which we questioned the program officer judgment.

The implication is that program staff considers increasing efficiencies in proposal review.

A.2 Questions concerning the selection of reviewers. Provide comments in the space below the question. Discuss areas of concern in the space provided.

SELECTION OF REVIEWERS	**YES , NO, DATA NOT AVAILABLE, or NOT APPLICABLE[2]**
1. Did the program make use of reviewers having appropriate expertise and/or qualifications? **Comments:** In general, reviewers appear to have sufficient and relevant expertise to evaluate the proposals and represent various constituencies. However, it is difficult to determine whether the panels include reviewers with specific expertise (e.g., about teaching and learning issues with students from non-dominant communities, or specific knowledge of cultural influences on teaching and learning).	YES
2. Did the program use reviewers balanced with respect to characteristics such as geography, type of institution, and underrepresented groups? Note: Demographic data is self reported, with only about 25% of reviewers	YES

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

reporting this information. **Comments:** We noted gender balance across reviewers. The number of reviewers from many states is at the 0-5% level with the exception of California, Massachusetts, and Pennsylvania that reflect percentages in the low teens. In terms of types of institutions, the majority of reviewers come appropriately from PhD institutions and Research Intensive PhD institutions (64%). The number of reported minority reviewers (13%) is quite low in comparison to non-minority reviewers, but this may be related to the lower percent of minority faculty especially in PhD institutions.	
3. Did the program recognize and resolve conflicts of interest when appropriate? **Comments:** We found that the majority of the proposals did not have any COI with proposals assigned to their panel.	YES
4. Additional comments on reviewer selection: The departmental affiliation of the reviewers is known but not their expertise. The particular expertise that reviewers have is important in the review process. We suggest that perhaps a database similar to those that associations like AERA and NARST have would be useful and would allow individuals to choose appropriate theoretical and methodological expertise.	

A.3 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.

RESULTING PORTFOLIO OF AWARDS	APPROPRIATE, NOT APPROPRIATE[3], OR DATA NOT AVAILABLE

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

1. Overall quality of the research and/or education projects supported by the program. **Comments:** The newness of the REESE program makes this a difficult question to answer. We had no final reports and only a limited number of annual reports to review. Based on the content of the review and annual report materials, the quality of the research appears promising. For example, the annual reports from projects funded in 2006 demonstrate clear progress towards meeting project goals. The quality of the lines of work supported by the REESE program appears high. The typical award proposal received reviewers' scores of 4.02, or "very good"; those declined received a mean score of 2.70.	APPROPRIATE
2. Does the program portfolio promote the integration of research and education? **Comments:** The main strength and focus of the REESE program is the integration of research and education. At a deeper level, the REESE portfolio has an appropriate balance of projects with immediate application to education and projects that are more remote from practical application. Integration of research and education is also evident in the mentoring of doctoral and postdoctoral students that appears in many proposals. The COV commends the NSF for its integrated program portfolio.	APPROPRIATE
3. Are awards appropriate in size and duration for the scope of the projects? **Comments:** In the materials we reviewed, award size and duration appear appropriate to the scope of work of the projects.	APPROPRIATE
4. Does the program portfolio have an appropriate balance of: • Innovative/potentially transformative projects? **Comments:** The COV reviewed a sample of funded projects indicating that a small fraction were highly innovative, risky, and yet potentially transformative	APPROPRIATE

projects. These projects tend to be basic research with practical application for long-term, not short-term payoff, which we consider an appropriate balance. We did not find that proposals lacked innovation. On the contrary, the reviews show an appreciation for the scientific advances offered by all the funded proposals.	
5. Does the program portfolio have an appropriate balance of: • Inter- and Multi- disciplinary projects? **Comments**: Based on the portfolio analysis conducted by the REESE research network, almost half of REESE projects combine STEM and other disciplines, and in 2008, the fraction of interdisciplinary work was even greater (49/75). Investigator self-reports are the basis of these data. Furthermore, it is not clear in the portfolio analysis document how decision makers categorized a project as, for example, "physical and natural sciences" vs. "STEM and other discipline combinations." Conversations with the POs and the text of the program solicitations identified the importance of combining theories and methods from multiple disciplines, yet this was not the definition used by the portfolio analysis. The COV suggests improving the portfolio analysis system by standardizing REESE and DR-K12. The sample of projects the COV examined yielded some evidence of multidisciplinary teams working together.	APPROPRIATE
6. Does the program portfolio have an appropriate balance considering, for example, award size, single and multiple investigator awards, or other characteristics as appropriate for the program? **Comments:** Based on proposals reviewed, it appears that a majority of the proposals are multiple investigator awards, although there is no summary of such data available. The proposals awarded show a normal distribution by award size, with an average of $227,171 per year and a duration of nearly 3 years.	APPROPRIATE
7. Does the program portfolio have an appropriate balance of: • Awards to new investigators? NOTE: A new investigator is an investigator who has not been a PI on a previously funded NSF grant.	DATA NOT AVAILABLE

Comments: How one decides what constitutes an appropriate balance is not clear. The data show that funding for PIs with prior NSF support occurs at a higher rate than with new PIs for several years. We note that in the last year of funding a significant imbalance of new and prior funding of PIs occurred at 9% and 29% respectively. Based on the stats provided, the odds for a proposal to be funded are 1:5 but the odds for a new investigator are 1:9. Although prior funded PIs may have more experience and thus may produce higher quality proposals, we suggest that REESE develop a means to achieve a better balance and work out a plan to improve the new PI capacity. We do not know how to determine appropriateness or inappropriateness at this point.	
8. Does the program portfolio have an appropriate balance of: • Geographical distribution of Principal Investigators? **Comments:** Several states have an acceptance rate of about 30%. Those that have 0% tend to have a smaller number of proposals. Outliers: Texas with 3% and a decent number of proposals (N=38); but Connecticut with 69% (N=13), DC and Arizona with about 50%; and a small number of states cluster around 40% (IL, MA, MI, OR, WI) which may be justified by clusters of faculty at institutions in these states who do fundable research.	APPROPRIATE
9. Does the program portfolio have an appropriate balance of: • Institutional types? **Comments:** The majority of the funded projects come from Research Intensive PhD institutions, about 50%, which fits with the 53% of proposals submitted by such institutions. The other major type of institution, PhD institutions, contributed 22% of the proposals. The percents of awards per institution matches the percents submitted except in FY2009 when awards from Research Intensive institutions jumped to 76% whereas only 5% came from PhD institutions.	APPROPRIATE
10. Does the program portfolio have an appropriate balance: • Across disciplines and sub-disciplines of the activity?	APPROPRIATE

Comments: Data from the ARC Portfolio analysis were difficult to use to address this question due to limitations in how the survey categorizes disciplines (the identification of the “STEM and other disciplines” category did not distinguish between disciplines such as “engineering” as a subject of study, e.g., engineering education, versus as a source of theory and methodology for the research). However, the portfolio does exhibit an appropriate balance between mathematics and science (15% versus 12%), and a comparable number of proposals combining more than one STEM discipline (16%). We had one question about balance that was difficult to evaluate given the limitations of the data in the portfolio analysis. That challenge was the focus in the portfolio on educational outcomes as defined within traditional disciplinary boundaries (e.g., math versus sciences, biology versus chemistry) versus the presence in the portfolio of projects examining teaching and learning in more innovative multi-disciplinary contexts. It is consistent with REESE’s goals for innovation to encourage multi-disciplinary problems such as climate change literacy, which draws on earth and atmospheric sciences, biology, and chemistry ideas. We recommend giving more explicit attention to soliciting proposals and tracking areas of focus along such multi-disciplinary lines.	
11. Does the program portfolio have appropriate participation of underrepresented groups? **Comments:** The tables provided in the proposals and awards statistics document provide relevant evidence. The yearly percentages of minority participation in incoming proposals ranged from 14-17% and percentages of minority participation in awarded proposals ranged from 10-17%. Similarly the percentages of women involvement in incoming and awarded proposals track each other well. Although the percentages of minority involvement in awarded proposals track the corresponding percentages in submitted proposals, the committee recommends that the program proactively find ways to increase the number and quality of submitted proposals with minority involvement to raise these percentages. We note that in the yearly statistics on minority involvement, the percentage of minority participation in incoming proposals held steady at about 16% whereas the percentage of minority participation in awarded proposals declined steadily from 17% to 10% during the four award cycles starting in 2006. The committee recommends that program personnel address this decline. The number of proposals submitted to NSF reflects low minority involvement (16%). We suggest that you look at the percent of minority faculty in the field.	APPROPRIATE

However, we note that the rejection rates for proposals with minority and non-minority involvement are nearly the same (27% and 33% respectively). The COV also recommends that the program consider minority involvement beyond the PI and co-PI status, and examine whether the trends are different when consideration is given to senior personnel and graduate students.	
12. Is the program relevant to national priorities, agency mission, relevant fields and other constituent needs? Include citations of relevant external reports. **Comments:** We evaluated the sample jackets available, the portfolio analysis, and program solicitations. The proposed research aligns well with the program's identified goals, national priorities, and the NSF mission. Criteria for relevance to the proposed research potential to "strengthen scientific and engineering research potential" were brought to bear in evaluating proposed research under the broader impacts criterion. The highlighted research demonstrates clear accomplishments relevant to this aspect of the mission. The mathematics results document improvement in statistical reasoning skills and advanced mathematical reasoning skills (rate, proportionality, linear function) through curriculum materials and technology developed in REESE research. These results also align with the constituent needs, such as the NCTM standards, which call for greater attention to the application of the mathematics to data applications, and inclusion of statistical reasoning in the middle and high school curriculum. Similarly, the science results help identify the challenges in advanced concepts of evolutionary thinking and demonstrate success in learning about biological diversity using REESE-funded materials, both areas of focus identified in national science standards documents from NRC and AAAS. The analyses of teacher shortages have clear implications for NSF's mission of improving "science and education programs at all levels" and represents a strand of REESE work designed to support NSF's mission of "providing a source of information for policy formulation." The identification of focus "emerging research" areas helps the REESE program respond to areas identified by its constituents. The areas flagged as emerging research respond to areas identified by the research community (e.g., neural bases of learning) and by the community of practitioners and policy makers (e.g., cyberlearning). We examined broad areas of focus identified in the portfolio analysis document and noted that the programs are aligned with attempts to investigate and improve STEM learning, but the portfolio analysis does not enable us to evaluate the significance of these projects' contributions.	APPROPRIATE

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

In order to maintain continuous assessment of the portfolio balance, we suggest that the program consider compiling summary data that capture whether underrepresented groups are studied and how; whether graduate students and postdocs who are involved are members of underrepresented groups; and whether teachers who are project participants are members of underrepresented groups.

A.4 Management of the program under review. Please comment on:

1. Management of the program.

Comments:

We considered the solicitation, reviewer selection, and post award steps in order to evaluate management of the REESE program. We reviewed the evolution of the REESE solicitation during fiscal years 06-10, a subset of jackets, stats on reviewers, proposals, awards, and dwell times in order to come up with our comments. The changes in the solicitation are evidence of an engaged and active program management process. Most of the award jackets included detailed information on award negotiations revealing that program officers discussed the substantive issues raised by panels and communicated these findings to the proposers who, following requirements, addressed these issues and questions satisfactorily in order to be funded. This policy is evidence of good program management.

Regarding reviewer statistics, we note that while there is a good gender balance, the percentage of minority reviewers was somewhat lower in 2008 and 2009 than in the prior years of 2006 and 2007. However, the large percentage of “unknowns” in this data makes accurate judgments about representation difficult. We recommend that reviewers and panelists be reminded and encouraged to provide this information to facilitate accurate data collection.

Institutional type and state-wise distributions of reviewers appear to be adequate. Award statistics reveal a clear separation between awarded and declined proposals in terms of mean ratings and a trend of higher success rates of experienced PIs compared to new PIs. Award distribution by state does not show any apparent bias. The REESE program met its dwell time goals in all years except for the 07 solicitation, when 29% of proposals had dwell times longer than 6 months (higher than the 25% target). Given that there are only a few external factors that could prolong the dwell times (such as, perhaps, delay in getting IRB permissions), we recommend that the program bring dwell times to below 25% consistently.

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

Comments:

We reviewed the portfolio analysis and the evolution of the REESE solicitation during fiscal years 06-10 in order to answer this question. The solicitation stayed relatively constant between 06 and 07 fiscal years, then in 08 included an explicit category, “frontier research” for tapping into emerging research and education opportunities, renamed “emerging research” in FY09 and continuing through FY10. This action is evidence of the program seeking out emerging opportunities identified by the research community. These emerging topics emphasized inter-disciplinary approaches and research teams. Our analyses of the subset of awarded REESE projects revealed that the program funded at least some proposals in all four of these emerging areas and that certain emerging areas received more proposals than others.

We note that NFS restricted submissions under the “emerging topics” to four subcategories – neural basis of learning, cognitive processes underlying teaching and learning, measurements, modeling and methods, and cyberlearning and teaching. The first of these morphed from neural basis of math learning to neural basis of STEM learning, whereas the others remained the same during 08-10. If a proposer identifies a new emerging opportunity that does not fit in these four subcategories, it is not clear whether he or she will be able to submit that proposal to REESE as an emerging topic. We could not tell what process, if any, identified these emerging topics and how the research community conveyed feedback. Though the committee feels that these four subcategories are well articulated and address critical emerging research and education opportunities, we recommend establishing a process for identifying future emerging topics if it is not already in place.

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

Comments:

We did not have access to documents detailing the program planning and prioritization process; instead we reviewed the REESE solicitations as products of the program planning and prioritization process. The REESE proposal solicitations from 2006 to 2009 reveal refinement from year to year to make the program priorities explicit in the solicitations. For example, we noted the expansion and clarification in later years of the categories, “synthesis” and “empirical research” that originally appeared in the 2006 solicitation. Program solicitations in these later years have specifically identified target areas of focus, including cognitive processes underlying STEM, neural basis of STEM, and cyberlearning and teaching; and have added another dimension to categorize the proposed research as “pathways,” “knowledge diffusion,” “empirical research,” and “large empirical research.” The program portfolio reveals evidence in 2008 and 2009 awards for proposals in these identified focus areas.

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

Comments:

REESE was formed in 2006, so there were no previous COVs.

5. Additional comments on program management:

No additional comments noted.

PART B. RESULTS OF NSF INVESTMENTS

The NSF mission is to promote the progress of science; advance national health, prosperity, and welfare; and secure the national defense (NSF Act of 1950).

The COV commented on (1) noteworthy achievements based on NSF awards in the portfolio under discussion; (2) ways in which funded projects have collectively affected progress toward NSF's mission and the strategic outcome goals of Discovery, Learning, and Research Infrastructure: and (3) expectations for future performance based on the current set of awards.

NSF investments produce results that appear over time. The COV review gave consideration to significant impacts and advances that have developed during the last four years but had no previous COV to review based on the newness of the REESE program.

We commented on the impact of NSF supported contributions to the field and made recommendations to use investments to stimulate emerging new areas, and addressed potential for transformative impact in research or education.

The COV used the program and award portfolio information that NSF provided, as well as members' own knowledge of the field, and other appropriate information to develop the comments for this section.

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

B.1 OUTCOME GOAL for Discovery: "*Foster research that will advance the frontier of knowledge, emphasizing areas of greatest opportunity and potential benefit and establishing the nation as a global leader in fundamental and transformational science and engineering.*" This category includes NSF's disciplinary and interdisciplinary research in science and engineering, education research, and centers.

Comments:

Of the sample proposals provided to the COV, we found several projects that meet the outcome goal for discovery. Two examples of this work are:

#0909588 PI: Robert Jacobs, University of Rochester. This newly funded project investigates the differences between expert and novice geoscientists in the field and in a virtual semi-immersive display environment. The composition of the research team includes scientists and educators with expertise in perceptual learning, geology and geophysics, recording and analyzing eye movements and high resolution, and large-field-of-view image capture of natural environments. This research program is potentially transformative because it is the first effort to understand and improve visual training in the geosciences using techniques commonly employed in the cognitive sciences.

#0910188 PI: Albert T. Corbett, Carnegie Mellon. Innovation in this new project draws on research in cognitive science, artificial intelligence, and genetics to develop an intelligent tutoring system in genetics. Project plans indicate the development of a system that can tell the difference between learners who actually understand the concepts versus learners who have figured out how to respond in a way that gives the appearance of genuine understanding, but who do not actually understand the concepts. This example represents a well-established technology in mathematics, but has yet to be adapted to other disciplines.

B.2 OUTCOME GOAL for Learning: "*Cultivate a world-class, broadly inclusive science and engineering workforce, and expand the scientific literacy of all citizens." This category includes K-12, undergraduate, graduate, and postdoctoral education and training; public understanding of science; and lifelong learning.*

Comments:

 # 0633952, Transforming Elementary Science Learning through LEGO Engineering Design is a project that studies new engineering-based science curriculum in 15 3rd and 4th grade classrooms to determine how children learn science from engineering-design-based instruction and also identifies best practices. This work gives elementary school students an opportunity to experience engineering that focuses on urban classrooms. The curriculum and its development are viable learning tools for teachers.

 # 0910191, Empirical Research: Breaking through the Reputational Ceiling: Professional Networks as a Determinant of Advancement, Mobility, and Career Outcomes for Women and Minorities in STEM, proposed by Julia Melkers and colleagues, addresses the characteristics and role of networks in career advancement, outcomes, and mentoring for women and underrepresented minority academic scientists in non–Research I institutions. We give particular attention to the role of mentorship and aspects of specific mentor resource exchange in affecting network access and participation. This study highly aligns with expanding outcomes for learning for under-represented groups in STEM. Specifically, the proposed study documents the content and dynamics of professional networks in the less research-intensive academic science setting and relates them statistically to both tangible and intrinsic career outcomes.

In examining the titles of the whole database of funded projects, we note a balance among proposals that focus directly on learning in mostly formal settings (i.e., classrooms); proposals exploring approaches / tools for learning but in design experiments; and proposals studying outcomes of classroom learning. CAREER proposals seem more field-based, studying the "activity" instead of only its outcomes. In looking for the outcome on learning of underrepresented groups, it seems that a handful of projects explicitly address learning of such groups; others argue that all students, including such groups, will benefit because of the nature of the activity (e.g., gaming); yet many other projects tend not to address such groups.

From the proposals reviewed, we note a gap between what the proposals aim to accomplish and specifics about how the work will cultivate a world-class broadly inclusive STEM workforce and scientific literacy. In the cases we reviewed, we noted superficial attention to the application or implication for using or influencing practice; in even fewer cases, these applications or implications were embedded and elaborated in the larger discussion of the work itself.

In regard to methods and research design we conclude that causal studies are the most frequently funded projects. We note that there is an under-representation of funded descriptive studies by comparison. Such studies would be of value in that they can respond to questions not addressed in other funded projects, as well as provide insightful information and documentation about what the treatment actually was, what the processes entailed, and relevant contextual information.

B.3 <u>OUTCOME GOAL for Research Infrastructure: "</u>*Build the nation's research capability through critical investments in advanced instrumentation, facilities, cyberinfrastructure and experimental tools. "*This category includes facilities, research instrumentation, and cyberinfrastructure.

Comments:

The components of the nation's research infrastructure to which REESE contributes are not instrumentation, facilities or the hardware, or even middleware of cyberinfrastructure. Rather,

REESE projects have the potential to build the educationally relevant application layer of the cyberinfrastructure, through funding projects that develop and utilize the application layer conducive to facilitating communication, collaboration and learning across geographic boundaries. Although this potential is present, our review of the available list of 174 funded REESE projects revealed only about 7 projects (4%) that could possibly be building on cyberinfrastructure. These findings match self- reported PI data from Table 3.7 of the REESE Portfolio Analysis, which states that 5% of all funded projects expect to generate software products. Thus, evidence exists that REESE is contributing to cyberinfrastructure. We suggest that REESE program officers may want to consider whether this proportion is adequate, and whether it tracks the proportion of submitted cyberinfrastructure related proposals.

Self-reported data show that 61% of funded projects intend to generate products such as curricula, TPD materials, new instruments for measuring outcomes, or new cyberlearning techniques. Even if only a small fraction of these actually generate products that pertain to experimental tools and cyberinfastructure, the outcome would result in REESE contributing to the nation's research infrastructure. A corollary of this assumption is that tracking the actual products generated would be useful for a future COV to determine the extent of actual contribution, as opposed to the extent of expected contribution on which we are commenting.

PART C. OTHER TOPICS

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

Proposers and reviewers are not directed to focus strongly enough on methodology; hence there is some inconsistency in attention to methodology, even in funded proposals. The COV recommends providing more direction to investigators and reviewers while recognizing the importance of maintaining an appropriate balance between the theoretical contribution, the contribution to educational practice, and the methodological rigor of the research.

The COV notes a low rate of minority PI participation and recommends exploring options to increase it.

We further suggest that program solicitation give greater attention to the possibility for broader impacts on policy. The current statement of broader impacts includes no mention of policy.

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.

We found the quality of the Program Officer review to be high with respect to synthesis, analysis and communication.

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

Based on past experience, some COV members questioned the value of PI meetings while others found some value in these meetings (e.g., networking, broader knowledge of NSF initiatives, etc). All agreed, however, that these meetings would benefit from better design to increase their value, especially in these tight budget times when travel money is limited.

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

The COV was not asked to comment on the portfolio distribution between elementary, secondary, and undergraduate. Moreover, we would like to know what the common themes are, if any, across grade level and across content level.

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

The COV process is good but could benefit from a more directive and detailed overview of the process in advance of the start of the meeting. While we appreciated the Webinar, we did not have an expectation of how the meeting would flow and did not review some of the material, such as the template, that was critical to the process. Some members did not understand the roles, scope and nature of their work before the meeting and were surprised to find they would be responding to questions and writing responses throughout the two-day session. The agenda was not available early in the process and did not match the actual sequence of the work that took place during the session. A briefing that includes the Chair, writer, panelists and NSF would be helpful a week or two before the COV begins to share information about the format of the meeting, the types of recommendations that would be helpful to the REESE staff, the scope and nature of the work and the deadline for completing the report. In addition, not everyone understood the benefit and high priority given to the closeout briefing with NSF.

The meeting room for REESE participants could have been better equipped with office supplies that participants needed during the course of their work. It also might be a good idea when working with two COV groups to have refreshments available in both rooms to minimize disruption, even though we acknowledge that all participants were invited to partake of the refreshments in the conference room occupied by the DR K-12 group.

One member would have preferred a mid-grain size evaluation versus evaluation of the portfolio as a whole. The reviewer thought that by identifying, for example, 3 or 4 pressing problems for mathematics education, the COV could then look across the portfolio to evaluate the extent to which the portfolio of related projects are making progress on these pressing issues. This reviewer felt that NSF could make it a requirement to write preliminary reviews in advance of the meeting.

NOTE: The following questions (C.6 – C.11) were provided by the REESE and DR K-12 Program Officers and reviewed and approved by Dr. Joan Ferrini-Mundy, DRL Division Director, and Dr. Wanda Ward, EHR Acting Assistant Director.

After the joint discussion of the REESE and DR K-12 COV members, the Chair elected to respond to this section via narrative. Here are the questions the COV members addressed and their collective response.

C.6. How do (or should) the REESE and DR K-12 programs complement each other?

C.7. How do (or should) the REESE and DR K-12 programs accommodate emerging concepts in research and/or development?

C.8. How do (or should) the REESE and DR K-12 programs attract or promote potentially transformative research?

C.9. How do (or should) the REESE and DR K-12 programs support cyber-infrastructure for learning, and what are the potential risks in this area?

C.10. How do (or should) the REESE and DR K-12 programs engage a broad spectrum of researchers and developers?

C.11. How well do the program solicitations and funding decisions reflect important issues in the field?

DR K-12/REESE Joint Considerations
(Questions C6-C11 Combined Narrative)

The portfolio analysis approach to understanding the characteristics of projects and people is essential and helps to frame thinking about how the DR K-12 and REESE programs complement each other. While the respective COV teams found flaws and ambiguities in the individual analysis of each program's portfolio, these findings should not be construed as a negative. Similarly, the idea of having two different organizations perform independent analyses of REESE (by ARC) and DR K-12 (by ABT) is good. This approach served to illustrate that self-reporting as categories can lead to a remarkably different statistical characterization than one based on coding information in proposals as well as to different definitions or interpretations of them. To ensure more effective joint consideration of DR K-12 and REESE in the future, the COV recommends that program management look for opportunities to standardize (where possible) and improve on the portfolio analysis process for both DR K-12 and REESE.

In terms of the complementariness between the programs, the subcommittees found that there are real ambiguities in the field about the distinctions between DR K-12 and REESE. The extremes are clear – basic versus applied – however the overlap between the programs seems substantial enough to justify giving serious thought to options that will improve understanding and/or mitigate confusion in the field. Accordingly, the subcommittees highly recommend that the NSF think about three potential options: sharpening the distinctions between the programs, consolidating the programs into one, or looking at some combination of the two.

This issue of "overlap" of course, relates to Pasteur's Quadrant – the notion that research is often an interaction between the basic and the applied (lab and field) rather than a linear progression. The matter invites one to think about a portfolio diagram that augments the cycle of innovation diagram that the NSF currently uses. The new joint portfolio diagram would look more like a Venn diagram (basic, applied, overlap) or quadrants, as in Pasteur's model.

Both DR K-12 and REESE use formal and informal methods (e.g., panel debriefings, annual PI meetings, workshops, etc.) to identify emerging concepts in research and/or development as well

as sow the seeds for potentially transformative projects. The subcommittees encourage the continued use of these approaches. We also recommend that program staff for both DR K-12 and REESE seek to further leverage future COVs as an additional source of ideas about emerging concepts. Given the in-depth review of the portfolio by the COV team members, big ideas, such as thematic research on cognition and minorities, are just as likely to come from a COV as from other external sources.

With respect to the DR K-12 and REESE programs' support for cyberlearning, the subcommittees determined that approximately 20-30% of the projects in each portfolio comprised cyberlearning-related projects. The subcommittees agreed that this was both an appropriate and reasonable balance in each program's portfolio. In the discussion that followed the presentations by the COVs to the NSF on October 23, it came to light that there was some disparity in the number of cyberlearning projects in the REESE portfolio. This ambiguity further supports the subcommittee's aforementioned recommendation to standardize and fine-tune the portfolio analysis process.

The DR K-12 and REESE subcommittees were uncertain about the meaning of the "potential risks" with regard to investments in cyberlearning as referenced in question C-9. Consequently, the group did not give much consideration to this aspect of the question. If this is an area that the respective program staffs would like future COVs to consider, the group suggests rewording the question to ensure clarity.

Finally, DR K-12 and REESE engage a broad spectrum of researchers and developers across a range of disciplines as evidenced by each subcommittee's responses to many of the questions in sections A1, A2, and A3 in the COV report. Specific information about where these researchers are based (e.g., Education Department of a university) was more detailed for DR K-12 than for REESE. Some subcommittee members took the position that question C.10 was ambiguous and therefore difficult to answer because there was no agreed upon definition of "researchers" and "developers" and the distinction between the two was unclear.

SIGNATURE BLOCK:

_____

Research and Evaluation on Education in Science and Engineering (REESE)
Robert Boruch
Chair