National Science Foundation Committee of Visitors Report
Interagency Education Research Initiative

May 10, 2002

Committee of Visitors

Chair, Dr. Anne C. Petersen, W K Kellogg Foundation

Dr. Douglas Carnine, Director and Professor, National Center to Improve the Tools of Educators, University

of Oregon

Dr. Guinevere Eden, Laboratory of Brain Function and Behavior, Georgetown University Medical Center

Dr. James Fey, Department of Curriculum and Instruction, College of Education, University of Maryland

Dr. Joseph K. Torgesen, Department of Psychology, Florida State University

Committee Meeting, April 25-26, 2002

Report Submitted May 10, 2002

NSF COV IERI Executive Summary

The Interagency Education Research Initiative (IERI) addresses an extremely important gap in US research funding: education research. Given the importance of education in the United States, and the high investment in education at the federal, state, local, and family levels, sound scientific information about effective educational practices for strong student outcomes is urgently needed. The disappointing state of current educational outcomes for young people intensifies the urgency even further.

Interagency efforts are extremely difficult to launch and maintain. The National Science Foundation (NSF) Committee of Visitors (COV) sees this interagency effort as a very promising approach to achieving a national program of education research able to achieve the necessary results. A sustained national effort has served us well in many other areas to bring important knowledge to bear on important societal issues such as national defense, space science, medicine and health care more generally, among other crucial areas. The lack of such a program in education research has meant more than the lack of a systematic knowledge base; it also means that promising students and top scientists are unlikely to focus their careers in this area. We applaud the three agencies – NSF, Department of Education (Office of Educational Research and Improvement), and National Institute for Child Health and Human Development -- for working together so effectively and urge them and all those who make policy for and fund these agencies to understand the importance of sustained commitment with stable funding for education research so that there will be a pipeline of outstanding investigators producing a comprehensive body of knowledge to address the nation's most pressing educational challenges.

The COV's overall perception of IERI management is that the interagency teams have done a remarkable job of translating the initial general goal into programmatic objectives with a generally effective implementation plan. The agency teams have been appropriately agile in refining funding guidelines and approaches with each successive year of experience with this initiative. The refinements have been intelligent and effective in yielding improved proposals.

The COV finds overall that the review process has integrity and is efficient. We make suggestions for further improvements, especially to achieve the initiative goals. Our analysis of the portfolio finds many strong proposals and we suggest ways to increase the strength of those areas with some limitations. It is too early to assess the quality of results of NSF's investments in this initiative as none of the major awards are yet completed. We identified some promising awards and have described these "nuggets."

We conclude with some suggestions on possible steps that might speed progress toward the vision and goals of the initiative. Again, we believe this is an extremely important step for the three partner agencies to be taking and we believe that their action should be strongly supported. At the same time, because we feel a degree of urgency about achieving the original goals within a realistic time frame, we urge adoption of appropriate strategies to (1) recruit a larger number of strong scientists to this education research, and (2) enhance the skills of current researchers in areas particularly needed for this initiative. Broader efforts beyond the scope of this initiative are also needed in developing the pipeline of researchers for education research and in developing a more extensive and innovative knowledge base in specific areas.

| Date of COV: April 25-26, 2002 |
| Program/Cluster: IERI |
| Division: REC |
| Directorate: EHR |
| Number of actions reviewed by COV: |

PART A.   INTEGRITY AND EFFICIENCY OF THE PROGRAM'S PROCESSES AND MANAGEMENT

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

QUALITY AND EFFECTIVENESS OF MERIT REVIEW PROCEDURES	YES, NO, or DATA NOT AVAILABLE
Is the review mechanism appropriate? (panels, ad hoc reviews, site visits) Comments: The overall operation of the review process seems appropriate to the program goals and structure. Almost all reviews are by panels that are constructed to provide the diverse expertise needed to judge the complex proposals that are submitted to the program. We encourage the IERI program officers to continue this practice, with special efforts to include reviewers from underrepresented groups. The review panels need to include reviewers with expertise in research design and data analysis, and a positive review from that reviewer should be an important criterion for funding. It would also be important for the methodology reviewer to provide constructive suggestions for that proposal's methodology.	Yes
Is the review process efficient and effective? Comments: Our overall impression is that the review process is efficient and effective. Panel reviews mean that the program officers get advice in a timely fashion. To assure that only solid proposals are funded, the program officers have, in at least one year, granted less than their budgeted funds. There are some concerns about the consistency of written review prose and numerical scores. Given the large and complex proposals for this initiative, it is important that reviews be thorough and provide enough feedback for promising investigators to be able to strengthen their revised applications. One concern about the review process is that the requirement for methodological rigor has not been applied as consistently as is appropriate for initiatives like this one. In particular, only a limited number of current funded projects are designed to test hypotheses or scale up educational interventions that are supported by evidence from prior research.	Yes
Is the time to decision appropriate? Comments: In general, the program officers seem to be meeting the GPRA time to decision goals. However, we noticed that in the most recent year there seems to be some drift toward longer dwell times. We suspect that this is a sign of heavy demands on program officers and that those who receive awards and those whose proposals are declined actually receive informal decision information in a more timely pattern than the official data show.	Yes

Is the documentation for recommendations complete? Comments: In the many proposal "jackets" that we reviewed, we were impressed that documentation for recommendations is careful and complete.	**Yes**
Are reviews consistent with priorities and criteria stated in the program's solicitations, announcements, and guidelines? Comments: In our analyses of individual proposal reviews and the summary data provided by an NSF contractor, we got the overall impression that reviewers did apply the priorities and criteria stated in the program solicitations. We noticed, for instance, in quite a number of the "jackets" for declined proposals comments such as, "this proposal does not seem appropriate for the IERI program." In another, reviewers commented that "the scalability requirement is not addressed in this proposal." We did notice with some awards that the reviewers did not speak to the initiative criteria as vigorously, particularly the criteria related to methodology. Among the specific review criteria detailing judgment of intellectual merit and broader impact, we noted careful attention to what we believe to be the most important items. There is somewhat less attention to broader impact than to intellectual merit, but we are not concerned by that difference. We did note that the question about attention to underrepresented groups is not as consistently answered as we would hope for.	**Yes**

Discuss issues identified by the COV concerning the quality and effectiveness of the program's use of merit review procedures: None beyond those identified above.

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

IMPLEMENTATION OF NSF MERIT REVIEW CRITERIA	% REVIEWS
What percentage of *reviews* addresses the <u>intellectual merit</u> criterion? In answering this question, we relied on a content analysis of 58 reviews of projects that were awarded funding during the first three years of IERI competition. The content analysis was conducted by an independent contractor, PRAMM consulting. A random examination of 5 folders suggested reasonable accuracy in the content analysis. The analysis indicated whether or not a particular aspect of intellectual merit was commented on. The percentage of applications on which the individual merit criteria were commented on by at least one of the reviewers for an award were: 1.1-84%, 1.2-97%, 1.3-33%, 1.4-83%, 1.5-48%. This pattern is consistent with expectations, with the most frequently addressed criteria being qualifications of reviewers, and the least frequently being creativity.	1.1 – 84% 1.2 – 97% 1.3 – 33% 1.4 – 83% 1.5 – 48%
What percentage of *reviews* addresses the <u>broader impacts</u> criterion? The percentage of applications on which the broader impacts criteria was commented on by at least one of the individual reviewers for an award was: 2.1-67%, 2.2-45%, 2.3-16%, 2.4-40%, 2.5-1%. Benefits to society were almost never mentioned and infrastructure enhancements were seldom mentioned. The smaller frequency with which these criteria were referred to in reviewer's comments probably reflects the salience of the intellectual merit vs. impact criteria in the views of reviewers. This seems appropriate for reviews of research.	2.1 – 67% 2.2 – 45% 2.3 – 16% 2.4 – 40% 2.5 – 1%
What percentage of *review analyses* (Form 7's) comment on aspects of the <u>intellectual merit</u> criterion? The percentage of review analyses that commented specifically on individual aspects of intellectual merit was: 1.1-45%, 1.2-52%, 1.3-9%, 1.4-65%, 1.5-14%. It is not clear why these criteria would be commented on less frequently than by reviewers, except that the analysis may have simply commented on the one or two most salient features of the application, while reviewers felt responsible to comment more broadly on the application. In addition, the reviewer data combines data from three different reviewers.	1.1 – 45% 1.2 – 52% 1.3 – 9% 1.4 – 65% 1.5 – 14%
What percentage of *review analyses* (Form 7's) comment on aspects of the <u>broader impacts</u> criterion? The percentage of review analyses that commented specifically on individual aspects of intellectual merit was: 2.1-15%, 2.2-9%, 2.3-5%, 2.4-5%, 2.5-2%. The broader impact criteria seem even less salient to those writing the review analyses than they were for individual reviewers.	2.1 – 15% 2.2 – 9% 2.3 – 5% 2.4 – 5% 2.5 – 2%

Discuss any concerns the COV has identified with respect to NSF's merit review system.

This is not really a concern, but it was interesting to note how infrequently the <u>broader impacts</u> criteria were addressed in the review analyses. This suggests either that these criteria are not highly salient for funding decisions, or that they are taken for granted by those writing the analyses.

A.3   Questions concerning the selection of reviewers.

SELECTION OF REVIEWERS	YES , NO Or DATA NOT AVAILABLE

Did the program make use of an adequate number of reviewers for a balanced review? Comments: An analysis of reviewers across the three IERI competitions indicates that the average number of reviewers who wrote reviews for applications was very close to three. This is judged to be adequate for a balanced review.	**Yes**
Did the program make use of reviewers having appropriate expertise and/or qualifications? Comments: The review panels were composed of experts from appropriate domains, and their expertise was sufficiently varied to allow multiple perspectives to contribute to the evaluation of applications. The disciplines of reviewers included: economics, computer science and engineering, neuro/biological/psychology, anthropology, linguistics, sociology, mathematics, life sciences/biology, physics, chemistry, social sciences, atmospheric sciences, cognitive psychology, educational psychology, statistics/methodology. As noted earlier, the review panels should each include one member with expertise in research design and data analysis, and a positive review from that panel member should be given important weight in funding decisions. It would also be important for the research methodology reviewer to provide constructive suggestions for that proposal's methodology.	**Yes**
Did the program make appropriate use of reviewers to reflect balance among characteristics such as geography, type of institution, and underrepresented groups? Comments: There was generally a low percentage of minority scientists on the review panels. However, this is not noted as a criticism, but reflects an ongoing challenge to draw scientists of African/American and Hispanic identify into the scientific review process at the highest levels. It was apparent from conversations with NSF staff that they have made a concerted effort to increase minority representation on the review panels, and will continue to make this effort.	**Yes**
Did the program recognize and resolve conflicts of interest when appropriate? Comments: The review process is carefully attuned to seeking out and resolving conflict of interest situations in the review process. However, we did note one declined application (9979795) in which one of the reviewers who wrote a review was from the same institution as the PI. The PI is in the college of arts and sciences, while the reviewer was in the college of education.	**Yes**
Did the program provide adequate documentation to justify actions taken? Comments: This appears to be consistently done in a careful way. In the several cases in which individual reviews seemed to be discrepant from the funding outcome (2 high reviews, one low review), the review analysis clearly identified the reasons for the declination, and these reasons consistently reflected the fact that the proposal did not meet one of the primary objectives outlined in the RFP.	**Yes**

Discuss any concerns identified that are relevant to selection of reviewers in the space below. No concerns beyond those identified above.

A.4 Questions concerning the resulting portfolio of awards under review.

RESULTING PORTFOLIO OF AWARDS	**APPROPRIATE, NOT APPROPRIATE, OR DATA NOT AVAILABLE**

Overall quality of the research and/or education projects supported by the program. **Comments:** The funded projects in this portfolio are of variable quality in relation to the goals of the IERI initiative. Many of the projects have important goals, strong prior research support for scaling up studies, and strong research designs. A number of current funded projects have notable limitations in research design and/or potential for significant payoff from scaling up studies. We expect that as the IERI program goals become more sharply focused and better known by the literacy, science, and mathematics education research communities, the quality of resulting proposals will improve. However, there are a number of ways that the IERI program officers can be proactive in making that progress a reality. Some of our ideas are discussed in section B5 of this report.	**Appropriate**
Are awards appropriate in size and duration for the scope of the projects? **Comments:** The awards were either large grants or planning grants, with the latter smaller and shorter than the former. In general, the size and duration was appropriate for the scope of the project.	**Appropriate**
Does the program portfolio have an appropriate balance of · **High Risk Proposals** **Comments:** The current funding plan involving planning grants and Phase I proposals is an appropriate means to help increase the number of proposals that will meet the major goals of the IERI program. We also hope other approaches such as some of the current outreach activities engaged in by NSF will be followed in order to increase the number of proposals that meet all the program goals. For example, NSF has had six workshops or public presentations on IERI in the past three months alone. These sessions, however, were similar to the current portfolio in that there appeared to be only one presentation integrating mathematics and pedagogy and nothing on science. It will be essential to target researchers from the currently under-represented areas.	**Appropriate**
· **Multidisciplinary Proposals** **Comments**: Most of the proposals funded under this initiative are multidisciplinary.	**Appropriate**
· **Innovative Proposals** **Comments**: This initiative emphasizes scaling up of existing models. While the proportion of applications that directly address questions about scaling up has so far been limited, it is anticipated given the beginning nature of the field, that future proposals addressing these questions will involve innovative ideas and designs.	**Appropriate**
Of those awards reviewed by the committee, what percentage of projects addresses the integration of research and education? **Comments**: The nature of this initiative in inherently integrative, applying research findings to improve educational practice.	Percentage 100% of funded proposals

Discuss any concerns identified that are relevant to the quality of the projects or the balance of the portfolio in the space below.

It is critical to point out that our review of the portfolio does not represent a review of the fields of reading, math and science education in general. Moreover, within our COV group, we differed in terms of our interpretation of scaling. For example, is the intent of IERI to scale up interventions or to scale up research, such as conducting large-scale

survey studies? Another question had to do with the robustness of the research that must under gird an intervention before scaling up. For example, could quasi-experimental studies with converging findings justify scaling up research on an intervention or are randomized control studies a minimal requirement?

The analysis that follows is based on best guess answers to the questions that arose among the COV members:

    a.   IERI intends to scale up interventions not research studies in general. This is not to say that large-scale studies involving correlational analyses, surveys, etc are not important but they are not the focus of IERI.

    b.   IERI requires some level of at least quasi-experimental findings in support of an intervention before going to scale in IERI. Again, this position is not intended to denigrate research on interventions using other methodologies; rather it is an assumption related to IERI specifically. We also value qualitative research and believe that it could have a complementary but not primary role in IERI studies.

If the portfolio had been reviewed with different assumptions, this summary would be different. It would be useful for IERI program officers to review the program solicitation and make sure they are consistent with the current goals of IERI.

Out of 13 funded proposals reviewed in greater detail, we found wide variability in terms of evidence base, scalability, methodology, and broader impact. Specifically only 5 out of 13 applications demonstrated sufficient evidence base as defined by experimental and quasi-experimental research that allows for causal interpretations. Even fewer applications had a sufficiently rigorous methodological approach that would ensure meaningful interpretation of the data. In spite of the limitations of the methodology employed, the applicants intended to go to scale (though some of the applicants proposed developing the necessary evidence during the IERI funding period to have a viable model to take to scale). About three quarters of the proposals included interventions/instruction, which seem central to scaling and the criteria of IERI. As a result of the limited evidence base and limits in methodology, it is difficult to assess the potential broader impact of most of these proposals. On the other hand, the proposals with a strong evidence base and strong methodology were exemplary the type of complex studies that can provide social benefit. Overall, the stronger proposals focused on reading and reading related skills. This field has had the benefit of a stable funding base, which has lead to well-established methodological standards and outcome measures and a long tradition of involvement by social science researchers.

We also note that most of the projects in the analyzed portfolio focused on the early school years or even pre-school. There were a handful of projects at the middle grades level but none at the high school level. This represents another gap to be addressed in the IERI initiative.

Improving the quality of the IERI program might entail the following:

1. Attracting researchers with a strong methodological background into educational research, particularly math and science education. This capacity building will require a number of actions that are summarized elsewhere in this report.
2. Stability in funding to encourage researchers to develop a program of research appropriate for IERI.
3. Deciding how to fund and support the development of evidence bases for interventions that would be suitable for scaling through IERI support.
4. NSF and program officers need to be clear about the relative importance of the benchmarks; for example, technology and interdisciplinary teams are not essential. In contrast, methodology is critical and a methodologist should be a member of every review panel. The methodologist's review should be seen as a vital part of the overall review and should provide sufficient feedback so that promising applicants can strengthen their proposal. Another issue is the need to clarify to applicants what constitutes an adequate evidence base to justify scaling up. Finally applicants must understand that scaling up is itself a target of investigation (in addition to testing the generalizability of the intervention to other populations and settings), requiring appropriate measures and methodologies that are continuing to emerge. A technical assistance center could be of help in planning for future research through which multiple grantees would measure common aspects of the implementation process.
5. Proposals to IERI need to include sufficient support from a methodologist to conduct large-scale studies.
6. Support prospective researchers in gaining access to subjects in schools by providing technical assistance, encouraging payment of stipends to schools, etc.

PART B. RESULTS :   OUTPUTS AND OUTCOMES OF NSF INVESTMENTS

Note: None of the Part B questions can be appropriately applied at this time to this initiative as there are no major projects (other than planning efforts) that have been completed. We have assessed a sample of the progress reports from among the first year awards and believe that most already have reported some significant results with the others having potential.

B.1.a COV Questions for PEOPLE Goal

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

PEOPLE GOAL INDICATORS	PROGRAM ACHIEVEMENT
Development of well-prepared scientists, engineers or educators whose participation in NSF activities provides them with the capability to explore frontiers and challenges of the future; **Comments**: Some of the awards are for teacher training and some of the awards focused on student outcomes involve teacher training.	**Data not available**
Improved science and mathematics performance for U.S. K-12 students involved in NSF activities; **Comments:** Almost all of the awards are directed toward improvement in science and mathematics achievement. Some of the awards focus on literacy, which we infer is generally important for science and mathematics.	**Data not available**
Professional development of the SMET instructional workforce involved in NSF activities; **Comment**s**:**	**Does not apply in any systematic way.**

B.2.a COV Questions for IDEAS Goal

<u>NSF OUTCOME GOAL for IDEAS</u>: **Enabling "discovery across the frontier of science and engineering, connected to learning, innovation, and service to society."**

IDEAS INDICATORS	PROGRAM ACHIEVEMENT
Discoveries that contribute to the fundamental knowledge base; **Comments**: While this initiative is not focused on the development of fundamental knowledge, we believe that most of the awards in the first three rounds will make some contribution. The new focus on research on scalability will likely yield significant contributions to fundamental knowledge of that topic.	**Data not available**

Leadership in fostering newly developing or emerging areas; **Comments**: Because this initiative draws attention to education research, many of the investigators will play a leadership role in fostering further developments in this area.	**Data not available.**
Connections between discoveries and their use in service to society; **Comments**: All of these awards are focused on the connection between discovery and service to society so there is high potential for significant results.	**Data not available.**
Connections between discovery and learning or innovation; **Comments**: Most of the awards in this initiative have a clear connection so there is high potential for significant results.	**Data not available.**
Partnerships that enable the flow of ideas among the academic, public or private sectors. **Comments**: Almost all of the awards made in this initiative have a connection between academics and schools, yielding high potential for significant results.	**Data not available.**

Comment on steps that the program should take to improve performance in areas of the IDEAS goal.   We believe that many of the ideas will be significant but we do not yet have results.

B.4 Please comment on any program areas in need of improvement. It is premature to make judgments here. We offer below suggestions for continued improvement of proposals and resulting findings and impact.

B.5 Provide comments as appropriate on the program's performance in meeting program-specific goals and objectives, which are not covered by the above questions.

While we believe that this initiative has made good progress thus far, we believe that the pace could be quickened by adopting some strategic approaches to recruit scientists into the field and by improving the skills of scientists already involved in education research. Other fields have faced similar challenges in rapidly progressing. For example, the early AIDs research programs engaged a number of successful strategies to attract outstanding researchers and develop a pipeline of strong researchers, and to draw proposals that would produce significant scientific results that would address this devastating societal problem. Other fields have engaged similar practices successfully. The US educational system and current student achievement is certainly at least as important a societal problem as AIDs and deserves no less attention from the research community. This requires a stable funding base of sufficient size as a necessary component.

Recruitment of people:
    While many outstanding scientists have submitted proposals for IERI, we have heard that some have not yet applied. Impediments to submitting include the perception that funding in this area has not been stable. The experience in other fields has been that specific recruitment efforts are needed. It would be important to talk with investigators who could contribute to IERI and learn their reasons for not submitting a proposal. This is especially important to do with math and science education researchers, groups who have not submitted many proposals thus far. Those reasons for not submitting proposals should be addressed in conferences or small group meetings to provide opportunities to recruit all who may be reticent. Recruiting a key person from a reluctant group is one effective tactic.

    Another approach is to recruit scientists from related fields. This can be done by inviting representatives from such groups to conferences that might bring them together with those scientists already involved with IERI to engage potential recruits with the attractiveness and importance of the focus of IERI by demonstrating compelling content in the conference. Stability of funding will require reassurance with this group as well.

Mid-career awards can also be used to persuade scientists from related fields to shift their current work to this topic. Specific training as part of this opportunity can be employed.

While it is beyond the scope of IERI, attention is also needed to the recruitment of students to this field. A steady pipeline of scientists is needed for a vital field. Scholarships and fellowships for students and postdoctoral scholars are a necessary means of recruitment in addition to conferences and other meetings.

Capacity Building/Skill Development:
We have observed that there are uneven skills among the awardees thus far, particularly in the research methods required by IERI. There are several approaches that have proven successful in other areas, particularly conferences or workshops designed for specific purposes and grant add-ons for particular expertise or approaches. Networking conferences with strategically designed agendas can teach participants about content areas and research methods. We were especially concerned about research design to test causality, assessment methods, and scalability. The new technical center should be able to help this effort.

Conceptual focus:
The focus of the IERI effort also requires continual attention. It is important that the work be cumulative and not meander. This requires continuity of leadership for the effort and consistency of vision and direction. We recognize that it may not be possible to have the same staff over a longer period of time but there needs to be attention to achieving consistency. The current team appears to be working hard to achieve this consistency and we saw progressively improvements in the program guidelines.

Finally, we recognize the difficulty of the task represented by this initiative. Interagency efforts are difficult to conceive, launch, and especially to implement. Yet, this effort must engage all relevant agencies. The importance of education is paramount and visibility given to the urgency and challenges here is also very high. The current status of the field of education research is uneven and certainly not at the level needed to show strong program results for children across all grade levels. The good news is that despite all the challenges, there is nothing that cannot be addressed given appropriate funding, effective initiative design, and strong implementation and oversight.

We applaud the effort and hope that our suggestions will help the initiative reach its goals.

B.6 NSF would appreciate your comments for improvement of the COV review process, format and report template. We found it to be a useful process in all respects.