

Report of the Committee of Visitors for the CISE Institutional
Infrastructure Program

November 14, 1995

Committee members

Jeannette Wing, Chair
Billy Buckles
Richard DeMillo
Jose Fortes
Roscoe Giles
Duncan Lawrie
Edward Stabler

This committee was convened to review the process used to solicit, review, recommend, and document awards in the CISE Institutional Infrastructure Program, administered under the Office of Cross-Disciplinary Activities (OCDA). There are four individual programs under the II Program: Research Infrastructure, Educational Infrastructure, Minority Infrastructure, and Postdoctoral Research Associates. The committee considered all four programs by examining the solicitations and a random selection of funded and declined proposals. The committee was presented with an overview of the four programs by John Cherniavsky (the Head of OCDA), Harry Hedges, and Tse Feng. Cherniavsky, Feng, and Carolyn Wardle provided additional input throughout the day.

Overall the COV review process was well-organized by NSF. The committee members had access to all proposal jackets to the four programs for the three years under review (1993-1995). Each jacket contains extensive documentation tracking the progress of each proposal.

Recommendation about the COV review process:

NSF should provide more information about the qualifications of the reviewers and the rationale for their selection. This information, provided by the program director, should be included in each proposal jacket. It could be brief, e.g., one-two lines about each of the reviewers and about the PD's justification for his or her choice of reviewers.

Recommendations that cut across all four programs:

1. NSF should continue to solicit explicitly and in a timely manner from the panel review members their opinions about the review process in which they participate. (This recommendation holds for all NSF panel-reviewed programs.)
2. NSF should continue to make an effort to include more members from industry and government research labs on their panels.
3. NSF should more systematically follow up on the funded parties after the award period ends to find out what worked and what did not. For example, are RI- and MI-awarded institutions self-sufficient? Where have NSF postdocs gone after they have finished? This follow-up feedback process can help establish some measures of success for each program and possible changes to their nature and administration.

In each section below, we discuss how we conducted the review, what we discovered, and any recommendations we have for improving the solicitation and review process.

RESEARCH INFRASTRUCTURE (RI) PROGRAM

We looked at nine proposals with the goal of evaluating the effectiveness and fairness of the processes used to decide funding and to determine the extent to which documented decisions were consistent with advertised RI goals and evaluation criteria. Three of the nine proposals were declined. One was funded through a cooperative agreement with ARPA. The remaining five proposals were funded in a manner that is normal for awards of this type.

Overall, we found good agreement between program goals and the output of the review and decision-making process. All of the successful proposals had merit that was adequately documented in the proposal jackets. For all but one of the declined proposals, there was adequate justification for turning down the proposal.

The decision-making process was organized into six phases. Three of these phases involved external reviewers: (1) initial review panel, (2) site review panel, (3) and final review panel. The three remaining phases of the process involved internal evaluations by NSF program personnel.

The COV felt that the information resulting from the initial panel was weak. Comments of individual reviewers tended to be telegraphic and sometimes cryptic. Panel summaries represented a consensus opinion of the review panels, but usually contained little substantive information beyond an overall separation of "must fund" "desirable to fund" , and "do not fund" proposals. It is unlikely that reviews of this quality would be useful to either the NSF program director or the PI.

Site review panels, on the other hand, were much more focused and complete in their evaluations. The practice of preparing narrative summaries of site visiting panel meetings is useful in providing complete documentation of the site reviewers' recommendations and should be continued. The final panel that reviews outputs of previous phases and ranks the surviving projects for eventual funding (or, more rarely, declination) appears to have a great impact on the eventual funding decision. Program directors usually respect these rankings of proposals. On the negative side, little useful narrative evaluations result from the deliberations of these final panels.

The quality of the documentation that supports individual decisions is highly variable. In one notable case, an unusual and volatile decision to reject a proposal that had been previously highly rated was taken with little supporting narrative discussion of why such drastic measures were necessary. The proposal (CDA-9303079) from the University of Illinois to operate a Computational Science and Engineering Research Center was highly rated by the initial review panel. After the site review, however, the program director suspended further evaluation of the proposal. After interviewing the program director on the merits of this case, it became clear that, during the site review, the PD became concerned that the proposal fell outside the intended scope of CISE and used this as a basis for rejecting the proposal prior to consideration by the final review panel. This decision by the program director is not supported by the narrative. The site reviewers expressed concerns about the quantity and quality of the proposed effort, and the choice of equipment and the overall integration of the activity with related activities at Illinois, but not about the whether the activity fell within the scope of the program. Our cause for concern is two-fold: one, that the PD redefined the scope of the program during the decision-making process; and two, that the rationale for the PD's decision was not documented in the narrative.

The COV detected no visible bias in the decision-making process. The reviewers were selected to represent the research community, although industrial researchers appeared to us to be underrepresented (it was

explained later that industrial reviewers from equipment manufacturers tended have irresolvable conflicts of interest). Similarly, the results from the period examined by the COV (1993-1995) seem to indicate that program is supporting a broad range of institutions with diverse research goals. We found, for example, about as many top-ranked programs as those with developing research programs.

There was some discussion of how well the awards were supporting previously defined program goals. One COV member, for instance, noted that parallel and distributed architectures dominated the awards. Upon closer examination of the research goals of the successful proposals it was clear that most participants had a rather clear idea of exactly what kind of research would be enabled by funding the infrastructure. In general, the Committee was pleased to see that funding decisions supported programmatic goals.

Recommendations:

1. The amount of technical detail from the initial review panel should be increased. Initial reviewers should provide detailed narrative reviews as input to NSF program management, to submitting organizations, and to later review panels. More effort should be expended to capture the deliberation of the initial review panel. The existing practice encouraging of terse "yes/no" responses to questions that could provide useful feedback should be modified appropriately.
2. Program directors should be discouraged from re-interpreting program goal statements while proposals are under active review.
3. NSF should renew its efforts to attract industrial researchers to its pool of reviewers for the RI programs.

Finally, a minority opinion was expressed regarding the three phases of external review (i.e., initial panel, site visit, final panel):

Eliminate the final review panel. If the initial review panel gave more substantive reviews and the number of site visits were kept low, for example, one or two above the number of those likely to be funded, then the program director would have enough input to make the final selection without the need for a final review panel. The advantages to this two-stage review process would be to streamline the review process and to place less of a burden on the research community.

EDUCATIONAL INFRASTRUCTURE (EI) PROGRAM

After examining seven proposals (four of which were awarded and three declined), we found that the review process for the EI program worked extremely well with remarkably consistent reviews across several panels. This finding was in contrast to the 1992 COV report which found reviews to be inconsistent. We attribute this turnaround to:

- A restatement of the goals in the program guide that makes the review criteria more evident; and
- Panels that consisted almost entirely of eminent computer scientists.

The Educational Infrastructure Announcement clearly states as the first and fourth criteria for evaluation that the proposed study/experiment have national impact and that there be a coherent plan for national dissemination. These seem to be the criteria that most affected reviewers.

Proposals that met all but one criteria did not receive a recommendation for funding. For example, Princeton (Sedgewick) did not propose a curricular objective with national impact and Columbia (Zukowski) had no dissemination plan. Other factors entered into some declination recommendations such as lack of faculty expertise in a key

discipline (Louisiana Tech) as well as overlap of plans with a competing proposal (Louisiana Tech versus USF).

The proposals that the panels recommended for funding appeared to meet the program criteria to the letter. For example, CUNY (Arnow) proposed holding a national workshop to coincide with the conclusion of their project to incorporate distributed processing units in three early courses. In the Rice (Cartwright) proposal the plan to assimilate two programming language paradigms, algebraic and physical world models, early in the curriculum and continue it throughout was closely related to the PI's research area providing multiple, well-developed avenues of dissemination.

Surveying the awards over 1993, 1994, and 1995, the Committee found that a broad range of topical areas were supported -- CAD, multimedia, databases, and software engineering were a few not already noted in the paragraphs above. In short, there was little duplication which is a positive feature of the program. While this factor is not entirely controllable (e.g., the proposal pressure in some areas may be greater than others), the Committee views it as a positive aspect of the program.

A minor problem noted by the committee was the withdrawal of proposals as not within the scope. This was reported at 8% for 1994 which translates to only two of 24 proposals. Can it be determined why PIs are submitting proposals inappropriate to the scope? While the goals of the program are loosely stated at the beginning of the program guide, the evaluation criteria that follow are reasonably clear.

Recommendation:

This program seems to work well from a fairness and consistency standpoint. The committee recommends that the Foundation continue to seek the highest-qualified reviewers possible to maintain the record of the last three years.

MINORITY INFRASTRUCTURE (MI) PROGRAM

The committee examined seven proposals. (None were planning proposals.) Three were funded and four were not. One proposal (CUNY-Hunter, unfunded) was the result of a planning proposal from the previous year. We also examined and discussed in detail the present description of this program--its goals and evaluation criteria.

The previous COV recommended a clearer description of the goals and criteria for this program and the Program Announcement was subsequently revised. We found there is still some room for interpretation in the guidelines, but there should be some flexibility in a program like this and so the present amount is appropriate. One particular concern arose regarding the declination of a proposal due to the reviewers' perception that it contained unrelated projects instead of a common focus. The RI program has the requirement for a focused proposal. This requirement for coherent focus may not be realistic for minority institutions which are unlikely to have faculty covering a large spectrum of topics with significant overlap. The NSF document describing this program does not mention the need for a common focus. It would be helpful to clarify whether such focus is or is not a requirement.

In general, we believe the evaluation process is a good one. But we did express some concern that the eminence of the reviewers seemed significantly less in this program than, for example, the EI program.

We discussed the current balance between risk and payoff, and there was some agreement that several of the proposals (FonDuLac and Texas El Paso) involved significant risk, but that at the same time, the

payoff could be significant if the programs are successful. We believe that the present balance of risk versus payoff is appropriate and that this program should continue to fund innovative proposals for improving the participation of minorities in CISE education and research.

The efficacy of the planning grants should be examined at some point. Very few have been awarded so far. Of the one awarded in each of 1993, 1994, and 1995, the ones awarded in 1993 and 1994 resulted in full proposals in 1994 and 1995, and both of these full proposals were unsuccessful. One might conclude that the planning grants are not effective, but the numbers are too small to be predictive at this point. In 1994 and 1995, 12 and 9 planning grants were submitted, with only one funded each year. Full proposals generally did not come from institutions whose planning grants were not funded. One could conclude that more planning grants would result in more full proposals. This might be reasonable given that these institutions typically have higher teaching loads and reduced emphasis on grant proposal activities; some release time might be necessary to allow sufficient time to permit preparation of an effective full proposal. Additionally the broad range of quality in the proposals we examined would argue for some process which would result in better proposals. Perhaps some form of mentoring or partnering would help in this respect.

Recommendations:

The nature of the MI program and the diverse criteria for proposal evaluation make it imperative that:

1. It use highly qualified reviewers.
2. It adopt a methodology and criteria for measuring success. This program has been in operation long enough to expect some degree of success. Yet there is presently neither measures of success nor a way to inform the community of successful versus unsuccessful practices. We recommend that a effort be made to define measures of success, and to apply these measures.
3. It provides an evolutionary path that evaluates the success of infrastructure projects, disseminates successful models, and adapts the program to changing needs. Specifically, we recommend that a workshop would be useful, both to disseminate successful models, but also to influence the evolution of the program. The result should be an NSF report that can be distributed nationally.

POSTDOCTORAL RESEARCH ASSOCIATES (Postdoc) PROGRAM

We considered in detail six proposals. These represented one accepted and one declined proposal for each of the three years (1993-1995) under consideration.

We felt that the candidates supported were strong and likely to be successful in their fields. We also concluded that there have been problems in clearly communicating to the applicants and the reviewers the goals and eligibility criteria of the program with respect to "breadth" of the training program.

The 1993 award to Stanford (9309389) exemplified what the program should be supporting: a strong candidate broadening her experience, moving from one strong research group to another at a different institution with a distinct research/academic goal.

Other cases were less clear. In another award to Stanford (95405058), the candidate was to be a postdoc with her thesis advisor at the same institution. A reviewer commented that this "doesn't meet the spirit

of the program," but the strength of the research, together (perhaps) with the argument that the candidate's research emphasis shifted, made an award possible. In the case of the MIT proposal (9404937), which presented a strong candidate working in the area of robotics, the proposal emphasized the strength of the advisor's research program over the training plan and received a negative reviewer comment that the research had been proposed already in another context. This proposal was declined.

What we infer from these examples is that two aspects of the program are not being communicated effectively: (1) the focus should be on the training of the postdoctoral candidate rather than exclusively on the research being conducted; and (2) the experience should be broadening, for example, by a substantial change in focus or direction of research and a change of advisor and institution; this requirement was unclearly stated early in the program. Both aspects are now stated in the program announcement, but need to be emphasized in practice.

Recommendation:

The program announcement should include a clear, definitive list of the kinds of activities eligible for support and a list of those ineligible.

For example, the first list should encourage applicants who desire:

- to enter computer science from a different discipline, e.g., going from medicine to computer science;
- to change the nature or style of their research, e.g., going from theoretical to experimental computer science;
- to change areas within computer science, e.g., going from artificial intelligence to software engineering;
- to change areas within a field in computer science, e.g., going from the architectural-level to the operating system-level of parallel computing.

The second list should discourage applicants who propose:

- to continue working in the same area as their thesis topic;
- to continue working with their thesis advisor.