
VI. VERIFICATION AND VALIDATION

A. QUALITY OF THE REPORTED PERFORMANCE INFORMATION

In FY 1999 concerns were expressed by the General Accounting Office (GAO) with regard to the quality of reported performance information used by NSF. To address these concerns in FY 2000, NSF engaged an external third party, Price-waterhouseCoopers LLP (PwC), to verify and validate selected FY 2000 GPRA performance data as well as the process used in collecting and compiling data and information. In their final reports, PwC concluded that NSF was reporting its GPRA measures with “sufficient accuracy such that any errors, should they exist, would not be significant enough to change the reader’s interpretation as to the Foundation’s success in meeting the supporting performance goal” Furthermore, PwC concluded that NSF “relies on sound business processes, system and application controls, and manual checks of system queries to confirm the accuracy of reported data. We believe that these processes are valid and verifiable.”

KPMG LLP, an independent certified public accounting firm, was selected by the NSF Inspector General to perform an audit of NSF’s FY 2000 financial statements. Their review included a review of the collection process and maintenance of data and information for NSF’s GPRA goals. NSF received an unqualified opinion stating that the principal financial statements were fairly stated in all material respects. The independent auditors did not report any material weaknesses in internal control or material noncompliance with laws or regulations.

All data are imperfect in some way. Establishing responsible and reasonable verification and validation procedures and understanding data limitations requires a balanced approach. NSF acknowledges the need to improve data systems for collecting and maintaining performance information and data as budget and time allow, and regards this as an evolutionary process which will continue to improve with time. NSF is comfortable with the quality of the data it uses in assessing the overall progress of the agency in meeting performance goals and makes use of the information it gains through performance reporting to improve policies, practices and management of the agency. Implementing GPRA has enabled NSF to gather information in a structured way, and to address issues in a more formal, focused way than in the past.

Because basic research and education projects rarely produce results in less than three to five years, it is difficult to compare the outcomes reported in the fiscal year with the funds that were obligated in that year. In some cases, the results of NSF support may not be recognized and reported for twenty years or more. Because the GPRA reporting schedule is annual, NSF conducts an annual assessment or evaluation of results submitted to the agency in the fiscal year, which is a retrospective evaluation, carried out by external experts. This retrospective evaluation makes use of the alternative form for reporting, to cover about thirty percent of NSF’s total portfolio in one year. This makes sense for NSF’s Outcome Goals, which are long-term goals, and are not expected to be achieved in a short time period. Nonetheless, we are

concerned that the current form of reporting goal achievement under GPRA does not convey the accomplishments of NSF or the full value of the NSF investment. To help the reader understand the level and accomplishment of performance for the Outcome Goals, examples are included to illustrate achievements reported during the fiscal year. While they may appear to be anecdotal, they can be traced to NSF-supported awards.

B. DATA VERIFICATION AND VALIDATION ACTIVITIES

For reporting goal achievement, all of NSF's goals are aggregated across the agency. To enable aggregation, NSF developed reporting templates in FY 1998, and data modules to collect data uniformly across the agency. These modules and templates were revised and refined in FY 2000 and were based on information gained in using the templates and systems in FY 1999. In FY 1999 NSF established a Data Quality Program to assess and improve the quality of data within the Foundation. NSF will continue to further refine data collection methods and systems to address areas in need of improvement as time and funds allow.

During FY 1999 NSF staff implemented a Data Quality Project for the quantitative Investment Process and Management goals. The objectives of the project were to:

1. Evaluate the quality of the data in the central databases.
2. Ensure the paper documents and the NSF central databases are synchronized.
3. Identify inconsistencies so that methods for correcting the cause of the inconsistencies can be developed.
4. Ascertain the causes of the data quality problems and develop systematic methods for correction.
5. Develop a comprehensive data dictionary.
6. Promulgate data quality policies and procedures NSF-wide.

In FY 2000, NSF increased the expected quality of information for the Outcome Goals in two ways:

1. NSF changed the two-level standard from *successful/minimally effective* to *successful*.
2. NSF required thorough justification for "grades."

NSF staff update and revise guidelines and reporting procedures for collecting data for the Outcome Goals annually. The Committee of Visitor (COV) guidelines were revised in FY 1999 and 2000 to incorporate the GPRA-related reporting requirements. COVs address a common set of questions for all programs reviewed in a fiscal year. Reporting guidelines were also developed for Advisory Committees to enable uniform, systematic aggregation of information. The results of using the new procedures helped to identify areas for improvement to the guidelines. These were incorporated for FY 2000 reporting and guidelines will be revised in FY

2001 based on experiences in FY 2000. The experience gained while conducting these assessments has also been used in revising the FY 2001 and FY 2002 Performance Plans and goals, and the updated Strategic Plan.

In addition, for FY 2000 NSF established parameters to define the acceptability and reliability of the qualitative information it uses. NSF defined the quality of the information it uses to insure uniform quality of results and applied stricter definitions of success in determining whether Outcome Goals had been met. NSF used a confidence limit to identify non-substantive information. Information falling outside the confidence limit was excluded from use.

The overall effect of applying these stricter definitions was to raise the expected performance level and reduce the aggregated success rate for NSF in FY 2000. However, the performance of the agency as a whole in FY 2000 was very much the same as in FY 1999 and positive trends are beginning to emerge. Many of the same issues identified by external groups in FY 1999 were identified in FY 2000. This is an interesting result in itself, since the Outcome Goals make use of judgement by different groups of external experts each year, and one might expect the result to be different if done by different groups, but this was not the case. Thus, this second year of reporting validated results obtained in the first year. A more complete picture will be obtained when results for the third year of reporting are known.

Information gathered from external sources for use in measuring performance related to the Outcome Goals is checked by NSF staff, reviewed by groups of external evaluators, and is subject to audit and tracking by association with grant numbers. In assessing its performance NSF makes use of reports generated by COVs who provide judgements. The scores and comments are compiled and aggregated to determine the success of the agency in meeting the Outcome Goals. This process was reviewed by PwC, who noted that the “approach NSF uses to assess its performance under these specific qualitative measures is reasonable”... and that in comparing NSF’s results with PwC’s results using the process established by NSF, “overall conclusions regarding program success or lack thereof in respect to individual goals remained largely unchanged.”

It is likely that NSF will continue to make use of external third parties on an appropriate schedule, to verify and validate data used in reporting performance goals as funds are available for this purpose.

C. TYPES AND SOURCES OF PERFORMANCE DATA AND INFORMATION

The data used in measuring performance are developed by and come from a variety of sources. Much of the data originate outside the agency, and quality is beyond the control of the NSF. Data come from administrative records, awardee reports, external committee reports, and internal data systems. Additional information can be found in the FY 2001 Performance Plan.

Quantitative data is used primarily in assessing the Investment Process and Management Goals. Most quantitative data used in assessing performance is collected using internal data systems and is reviewed by staff on a quarterly schedule.

Most of the qualitative information used in assessing Outcome Goal performance is provided to NSF by external evaluators (COVs) near the end of the fiscal year, and is reviewed by NSF senior management as it becomes available.

Collection of data is dependent on the type of data/information. Collection of data for all goals takes place throughout the year and is completed near the end of the fiscal year. Data are collected into reports for each goal by a group staff having reporting responsibility for the particular goal. The data obtained are reviewed on a continuing basis by senior management throughout the year, in order to observe whether the results are as expected, whether performance needs to be improved, whether targets need adjustment, or whether the information being obtained is useful to the agency. Data collection systems are also under constant observance and refinement.

D. DATA LIMITATIONS

Specific data limitation issues are discussed below. The NSF FY 2001 Performance Plan contains additional information on data sources and limitations.

This is the second year in which reports were collected, tabulated, and an assessment of NSF's performance was completed. Several data quality/limitation issues were identified in the first year of reporting. The agency worked to address these issues during FY 2000. The issues included: incomplete data collection systems related to some of the quantitative goals (such as the goals related to Facility Operations – Investment Goals 12-15); the need to improve report templates to ensure that the performance information provided by external groups is more complete and consistent for the qualitative Outcome Goals; and explanations for goals that were missed.

Steps were taken to improve the quality and value of performance data for the Outcome Goals. They included improved reporting templates for collecting program performance information from external committees by asking for more complete justifications for ratings. We note an improved quality and consistency of COV reports for FY 2000, but note they are still not optimal. NSF has modified the COV reporting template guidelines for FY 2001 to further improve consistency and completeness. NSF staff will work more closely with COV members to ensure improved reporting. This will aid NSF in aggregating qualitative information for measuring progress in achieving the Outcome Goals.

NSF employs an alternative form for determining progress made in achieving its Outcome Goals for research and education. In FY 1999 NSF made use of the alternative form using the two-standard approach required by the Act (*successful* or *minimally effective*). In doing so, NSF learned that there was little to be gained in using *minimally effective* standard, and that in many

instances it was confusing to the evaluators. Therefore, for FY 2000, NSF defined one standard only: the *successful* standard. The effect of this change was to increase the level of expected performance for the Outcome Goals.

When NSF collects performance information it asks COVs to indicate when data is not adequate or available to evaluate progress toward meeting an Outcome Goal. NSF found in FY 1999 and FY 2000 that external evaluators did not always have adequate information available to judge each program in use of the merit review criteria; in achieving increased participation from under-represented groups; and, in achieving science and math skills for *all* Americans. NSF management is reviewing means that will help NSF staff to provide this information for FY 2001 assessments. In some instances, data is difficult to obtain. An examples is complete data describing the participation of under-represented groups, which is voluntary.

NSF is also reviewing the wording of goals to correct issues which created difficulty when aggregating results in FY 1999 and FY 2000. For example, Outcome Goal 3 combines achieving increased diversity with achieving a globally-oriented workforce. While NSF was judged successful in achieving a globally-oriented workforce in most programs, it was judged less successful in achieving increased diversity. Consequently, NSF is not able to indicate success across the entire agency in FY 2000 for this goal, although some aspects of the goal were realized by programs, in particular those programs with funds targeted directly to meet these goals.

A similar situation arose in evaluating Outcome Goal 4, which targets improved math and science skills for *all* Americans. NSF programs were successful when they had clearly invested funds to support activities relevant to achieving this goal. It was less apparent to external groups whether success had been achieved for programs not designed to specifically address this goal, and the resulting COV reports did not provide clear evidence of success at the aggregate level. Therefore, for this goal, we are unable to indicate successful performance for the agency.

A new format has been adopted for NSF's goals in FY 2001 which we hope will help to alleviate some of these issues: the five Outcomes Goals have been organized under three headings, each with independent indicators. This will aid assessment by COVs to address the indicators separately when they are relevant to the program being evaluated.

Another limitation noted was for Investment Process Goal 8, Maintaining Openness in the System. It was found that the identification of new PIs was inaccurate on occasion, and steps were taken to identify such individuals in the NSF PI system more carefully. For Investment Process Goals 12-15 on Facility Oversight, the reporting system was revised and implemented in FY 2000. This system was provided to facility managers located at universities who must use the NSF developed system to report data that support this goal. There was lack of agreement in FY 1999 on how the required data were defined, which led to different interpretations. This deficiency was addressed in FY 2000. However, facility managers are still gaining experience in collecting and providing information needed for reporting these goals.

For the quantitative Management Goal 3 - Staff Diversity - a reported data limitation for this goal is that an applicant by law cannot be required to provide gender and ethnicity information.

Thus, it is certain that the results for this measure are not accurate. Goals which require voluntary self-reporting are also subject to being incomplete or inaccurate.

Finally, in reporting results for all goals in FY 1999, NSF did not always have a complete understanding of why some goals were not met. For FY 2000 NSF revised report templates for collecting information across directorates and offices within NSF. Each reporting organization within NSF was asked to provide a summary of performance at a lower organizational level, and to provide explanations when agency goals were not met along with plans to meet those goals. This provided the agency with more complete information on goal performance in FY 2000 and has helped to identify several important obstacles critical to achieving some goals. This information is being used to develop implementation strategies for meeting targets in future fiscal years.

E. OTHER ISSUES - TIMING

One of the most significant challenges for NSF is that results of research and education investments do not appear annually or on schedule. NSF faces other timing issues in preparing this report. Such timing issues may be shared by other agencies. One timing issue is related to NSF's reliance on external committees to conduct assessments after the close of the fiscal year. Materials are prepared in advance, but there is a narrow window of time between the end of the fiscal year and the start of the calendar year for the assessment to be conducted by external groups and the results to be finalized and written by the external committees. The committee reports must be submitted to the agency, and reviewed by the agency. NSF relies on the availability and cooperation of the external community and their ability to deliver their assessments on a tight schedule.

In addition, the timing and phasing of the annual plan, collection of information, and data for reporting have been difficult to coordinate with the budget process. To optimize goals and plans for the new fiscal year, NSF must review progress from the prior fiscal year, and make revisions to the annual plan for the upcoming year. However, the Performance Plan is typically due well before the results of the prior fiscal year are known. This creates an awkward situation, in that an early Performance Plan may need significant revisions to best serve the agency.

In FY 1999 and FY 2000 NSF found that the time needed to collect and review the annual performance data, and incorporate changes into the FY 2001 and FY 2002 annual performance plans in a way which we believe benefits the process and lead to desired results, was insufficient to meet the current schedule set by law. We are reviewing staffing and procedural mechanisms to accelerate the process for aggregating performance results. However, this is unlikely to yield an agency result at the right time to develop a plan appropriate for the upcoming year.