

Metaevaluation of the Web-Based ATE Survey Evaluation System

David Hartmann
Georgios Loizides

Kercher Center for Social Research
Western Michigan University
November 13, 2001

Contents

Introduction / Executive Summary.....	1
Methodology.....	1
Conclusions.....	2
Program Evaluation Standards.....	3
Utility Standards.....	3
Feasibility Standards.....	6
Propriety Standards.....	9
Accuracy Standards.....	12
Checklist for Applying the Standards.....	18
Principles for Constructing Web Based Surveys.....	19
Checklist for Applying the Principles for Web Based Surveys.....	29
Individual Survey Sections.....	30
Conclusions and Recommendations.....	33
Bibliography.....	34

Introduction/Executive Summary

This report is a metaevaluation of the web-based survey system used by The Evaluation Center in their National Science Foundation (NSF) Advanced Technological Education (ATE) evaluation project in 2000 and 2001. Two sets of standards were applied to the web-based survey. The first is *The Program Evaluation Standards* (2nd ed.) from the Joint Committee on Standards for Educational Evaluation. This set incorporates evaluation standards in the areas of utility, feasibility, propriety, and accuracy. The second is a set of standards for Internet surveys derived from Don Dillman's "Mail and Internet Surveys: The Tailored Design Method" (2nd ed.). Other source material such as internal documents, email records, past reports by The Evaluation Center concerning the ATE survey, and other literature on web survey questionnaire design and administration was also referenced.

Methodology

In order to assess The Evaluation Center's web-based ATE survey, we examined, apart from the survey itself, the Center's reports on the impact and effectiveness of the ATE program (Status Reports 1-Overall description of ATE program and initial evaluation; Status Report 2-Survey 2000 findings and recommendations; Survey 2001 Report—draft and final), the 2000 Nonresponse Study, relevant sections of site visit reports, ATE web site information, implementation schedules, data handling/verification procedures, a helpful hints document available to survey respondents, emails sent during Survey 2001 open period, electronic versions of SPSS data files/data dictionaries, and feedback emails. Furthermore, we interviewed staff of The Evaluation Center on two

occasions in order to gather additional information. The web-based survey itself was accessed a number of times, using the two most widely used web-browsers in the market today; namely, Netscape (version 4.75) and Microsoft Internet Explorer (version 5.50). Upon accessing the ATE Survey site, we systematically followed all possible steps (all possible links) throughout the survey, using both browsers on each step.

Conclusions

Both by the Program Evaluation Standards and Dillman's web-survey construction guidelines, the ATE survey is an exemplary evaluation tool, which supported timely, thorough, and useful reporting to NSF and to the ATE programs. Careful and critical review of the survey's performance by Evaluation Center staff was also in evidence and resulted in substantial improvements in the second iteration. Despite this highly positive assessment, we do offer specific suggestions for improvement of the functionality and user-friendliness of the survey throughout this report, which are outlined on pages 19-28.

Program Evaluation Standards

What follows is an application of the Program Evaluation Standards to the ATE survey. Therefore, the structure of this section is organized around the four important attributes of an evaluation -utility, feasibility, propriety, and accuracy - as stated in "The Program Evaluation Standards: How to Assess Evaluations of Educational Programs," (2nd ed.) by The Joint Committee on Standards for Educational Evaluation, excerpts of which are used throughout this section. This narrative is summarized in the Joint Committee's recommended "Checklist for Applying the Standards" on p.18, Table 1.

UTILITY STANDARDS:

The utility standards are intended to ensure that an evaluation will serve the information needs of the intended users. Utility standards guide evaluations so that they will be informative, timely, and influential. They require evaluators to acquaint themselves with their audiences, define the audiences clearly, ascertain the audiences' information needs, plan evaluations to respond to these needs, and report the relevant information clearly and in a timely fashion. These standards are as follows:

1. Stakeholder Identification
2. Evaluator Credibility
3. Information Scope and Selection
4. Values Identification
5. Report Clarity
6. Report Timeliness and Dissemination
7. Evaluation Impact

Stakeholder Identification: The Standard was addressed.

The Evaluation Center correctly identified two primary stakeholders: the sponsoring organization - the National Science Foundation (NSF), which has a need to know about the performance of its funded programs and has certain reporting requirements (e.g., Government Performance Results Act [GPRA]); and the Advanced Technological Education (ATE) *projects (i.e., projects and centers)* themselves, which have a responsibility to report but also an interest in minimal intrusions and efficiency of reporting. Secondary stakeholders' interests are also addressed, such as the interests of future grantees, whose interest is in learning from the experience of current programs, and the U.S. Congress, which is responsible for oversight of funds expended. The interests of all of these stakeholders are recognized in an evaluation system that stresses thorough timely reporting with a high response rate and in a form that allows rapid and easily understood dissemination of results.

Evaluator Credibility: The Standard was addressed.

The conducting of the evaluation by The Evaluation Center at Western Michigan University ensured both a trustworthy and a competent evaluation. The Evaluation Center is justly recognized as a national leader in educational evaluation in general and in the support of technological initiatives in particular.

Information Scope and Selection: The Standard was addressed.

Information collected in the Web survey was broadly selected, and addressed questions about the program, that are pertinent to the needs and interests of the

stakeholders (ATE - NSF). The potential burden of collecting so much information on program structure and performance is minimized by the modular design allowing temporal spacing and delegation of reporting responsibility.

Values Identification: The Standard was addressed.

The perspectives, procedures, and rationale used to interpret the findings were clearly described in the reports produced by The Evaluation Center, and in other documents produced by The Evaluation Center. The NSF-ATE guidelines embody the values that participants have agreed to address and are the basis for the information requested in all sections of the web-based survey.

Report Clarity: The Standard was addressed.

The Evaluation Center reports on the impact and effectiveness of the ATE program by annually publishing the results of its Web-based ATE evaluation. The series of reports are titled "Assessing the Impact and Effectiveness of the Advanced Technological Education (ATE) Program." All reports, up to date (including the nonresponse report) presented their findings in a clear and efficient manner.

Report Timeliness and Dissemination: The Standard was addressed.

Having the survey instrument on the Web diminishes the data collection time, and thus facilitates the timely publication and dissemination of evaluation reports. Reporting on the evaluation of the ATE program has indeed been very timely. It is characteristic

that the third report in the series had a data collection period from February 20-April 9, 2001 and a draft report published in July 2001.

The "Assessing the Impact and Effectiveness of the ATE Program" reports are sent by The Evaluation Center to NSF, but are also part of the Center's web page, which makes them accessible to all stakeholders. Evaluation Center staff has stated the intent of the Center is to also have this metaevaluation report available online.

Evaluation Impact: The Standard was addressed.

Although there were minor changes to the survey instrument between the time of the second (2000) and the third report (2001), these did not affect compatibility of results. The ATE evaluation is generally planned, conducted, and reported in ways that encourage follow-through by stakeholders.

FEASIBILITY STANDARDS:

Feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal. Feasibility standards recognize that evaluations usually are conducted in a natural, as opposed to a laboratory, setting and consume valuable resources. Therefore evaluation designs must be operable in field settings, and evaluations must not consume more resources, materials, personnel, or time than necessary to address the evaluation questions. The standards in this category are as follows:

1. Practical Procedures
2. Political Viability
3. Cost Effectiveness

Practical Procedures: The Standard was addressed.

The procedures involved in carrying out the web-based survey are practical, although perhaps more explicit and step-by-step help should be given to the survey participants (see “Principles fro Constructing Web Surveys” below). Furthermore, the use of a web-based survey instrument, together with the use of email invitations for participation, the use of passwords to ensure security, and follow-up phone calls to people who were invited to participate but did not yet log on to encourage them to participate, ensured impressively high response rates (73% for 2000 and 93% for 2001), as well as a short data collection time period. Changes in design and contact protocols after the first year appear to be directly responsible for the improved response rate.

Political Viability: The Standard was addressed.

"The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted."

The Evaluation Center carried out a nonresponse study for the year 2000 ATE survey, thus addressing the potential of nonresponse bias. The report is interesting in that it showed no significant response bias, but also in that it offered helpful feedback from "nonrepondents" on the reasons why they did not respond to the survey.

Despite the lack of documentable bias found in the nonresponse study, it may still be the case that - due to lacking help on how to proceed - novice, or first time computer

users may be adversely affected more than other groups of users. “True novice users are assumed to know little of the task or interface concepts. By contrast, first-time users are professionals who know the task concepts, but have shallow knowledge of the interface concepts. Both groups of users may arrive with anxiety about using computers that inhibits learning. Overcoming these limitations is a serious challenge to the designer of the interface, including instructions, dialog boxes, and online help.” (Shneiderman, 1998; p.69). The Evaluations Center’s attention to such issues in the second year (e.g., availability of a hardcopy option to those respondents with technical issues, helpful hints document) is commended and bodes well for future iterations of the survey.

Cost Effectiveness: The Standard was addressed.

Although the web-based survey instrument had high initial costs in order to properly develop the appropriate computer program and remove all glitches from it, its maintenance should be very low cost. Furthermore, the use of a web-based survey instrument renders data collection practically cost-free, while producing highly valuable information (Shneiderman, 1998; p.133. Also see Birnbaum, 2000). Therefore, in the long run, the present set-up is a cost effective means of evaluating ATE. It is also important to note that The Evaluation Center attempted to find an “off the shelf” survey system at the start of the project and assessed known products. The state of the art at that time simply would not support the functionality (e.g., delegating responsibility for sections of the survey in a secure and tractable environment) that they felt was essential to the feasibility and utility of the system.

PROPRIETY STANDARDS:

Propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results. Propriety standards reflect the fact that evaluations affect many people in a variety of ways. These standards are intended to facilitate protection of the rights of individuals affected by an evaluation. They promote sensitivity to and warn against unlawful, unscrupulous, unethical, and inept actions by those who conduct evaluations. The standards in this category are as follows:

1. Service Orientation
2. Formal Agreements
3. Rights of Human Subjects
4. Human Interactions
5. Complete and Fair Assessment
6. Disclosure of Findings
7. Conflict of Interest
8. Fiscal Responsibility

Service Orientation: The Standard was addressed.

The ATE evaluation was designed and is carried out to assess the impact and effectiveness of the Advanced Technological Education (ATE) Program, in order to assist organizations to address and effectively serve the needs of future participants. Both the content of the survey and response advantages in terms of timeliness and response rate met the needs of NSF and the funded ATE programs.

Formal Agreements: The Standard was addressed.

Obligations of the formal parties (Evaluation Center, Kercher Center, NSF) to the evaluation were agreed to in writing (what is to be done, how, by whom, when) so that these parties are obligated to adhere to all conditions of the agreement. Although there were no formal agreements drawn with ATE program grantees, participants in the ATE survey received timely and extensive notification concerning their participation (through email, mail, and telephone follow-ups). They were also introduced to reporting expectations by NSF staff at the grantees conference at the beginning of the second year, thereby further encouraging full compliance.

Rights of Human Subjects: The Standard was not applicable.

The ATE survey gathers feedback from organizers of ATE *projects* concerning their *projects* and report the aggregate findings back to NSF. Thus, the ATE evaluation does not involve “human subjects” in the formal sense of the term; respondents are collaborators sharing non-personal information rather than research subjects asked to disclose personal information. Therefore, the use of (informed) consent forms, as well as other ethical considerations - such as confidentiality - do not apply to this project; this standard is not applicable.

Human Interactions : The Standard was addressed.

There is nothing in the ATE evaluation survey that runs the risk of threatening or harming the human dignity and worth of survey participants or others that may be affected by the survey.

Complete and Fair Assessment : The Standard was addressed.

The ATE evaluation survey is constructed so as to be complete and fair in its examination and recording of strengths and weaknesses of the various ATE *projects* in that the various sections comprising the questionnaire are designed to follow NSF-ATE guidelines and response categories of individual questions are thorough. Open-ended questions are also used to allow additional information to be shared.

Disclosure of Findings : The Standard was addressed.

The Evaluation Center published its findings of the ATE evaluation survey in three status reports up to date, titled “Assessing the Impact and Effectiveness of the Advanced Technological Education (ATE) Program.” The first report dealt with “the Nature of the ATE Program,” the second reported on the “Findings from a Survey of ATE Projects and Centers,” and the third reported on “the Status of ATE Projects and Centers.

Conflict of Interest: The Standard was addressed.

No conflict of interest is discernable. Careful attention to the Program Evaluation Standards and, in particular, to frequent and clear reporting to all stakeholders helped ensure this standard.

Fiscal Responsibility: The Standard was not reviewed.

Although the ATE web-based survey had high initial costs, in the long run, the use of a web-based survey will be efficient as maintenance and application of web-based surveys carry minimal costs. Since we were not asked to conduct a financial audit or review, no further information on fiscal responsibility was requested by the metaevaluators or provided by The Evaluation Center.

ACCURACY STANDARDS:

Accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated. Accuracy standards determine whether an evaluation has produced sound information. The evaluation of a program must be comprehensive; that is, the evaluators should have considered as many of the program's identifiable features as practical and should have gathered data on those particular features judged important for assessing the program's worth or merit. Moreover, the information must be technically adequate, and the judgments rendered must be linked logically to the data.

The standards in this category are as follows:

1. Program Documentation
2. Context Analysis
3. Described Purposes and Procedures
4. Defensible Information Sources
5. Valid Information
6. Reliable Information
7. Systematic Information
8. Analysis of Quantitative Information
9. Analysis of Qualitative Information
10. Justified Conclusions
11. Impartial Reporting
12. Metaevaluation

Program Documentation: The Standard was addressed.

The ATE program, which was the subject of the evaluation survey was described and documented clearly and accurately in The Evaluation Center's reports, so that the program is clearly identified.

Context Analysis: The Standard was addressed.

The context in which the program exists should be examined in enough detail, so that its likely influences on the program can be identified. The programmatic and pedagogical context of ATE *projects* was well understood by The Evaluation Center and underlies the survey content and design. In sum, the appropriateness and thoroughness of the ATE survey is indicative of a thorough knowledge of the operations and context of ATE programs.

Described Purposes and Procedures: The Standard was addressed.

The purposes and procedures of the evaluation were monitored by The Evaluation Center team, and described in enough detail, as part of the Center's reports, and other documents provided to the Kercher Center evaluation team so that they can be identified and assessed.

In particular, the protocol used to enhance response rate and to produce valid results was described in detail in Appendix A of the 2001 report (p. 57).

Defensible Information Sources: The Standard was partially addressed.

The Evaluation Center used a number of sources for designing and carrying out its evaluation of the ATE program. These include, apart from the web-based survey itself, a thorough literature review on web-based survey design and feedback from participants to the pilot program, which preceded the formal data collection. These sources of information were described in enough detail, so that the adequacy of the information can easily be assessed. That said, it must be noted that the survey does not require, nor does it allow for, the provision of documentation to support responses by survey participants (i.e., the survey is a self-report). This is a standard problem with survey response information and was addressed in this project through site visits to program sites. The site visit phase was reviewed separately and independently from this analysis of the web-based survey.

Valid Information: The Standard was addressed.

The information gathering procedure (ATE web-based survey) was developed and implemented in a way so as to assure that the interpretation arrived at is valid for the intended use. The survey was well designed (see especially below on web-based standards) and requested information that program staff could reasonably be expected to have available. Site visits were used as a validity check.

Reliable Information: The Standard was addressed.

The information gathering procedures (ATE web-based survey) was developed and implemented in a way so as to assure that the information obtained is sufficiently reliable for the intended use. Again, site visits allowed a reliability check.

Systematic Information: The Standard was addressed.

The information collected, processed, and reported in an evaluation should be systematically reviewed and errors corrected. The Evaluation Center does extensive quality monitoring and item analysis. Their response to potential problems (most notably with nonresponse) is thorough and well grounded in relevant literatures.

Analysis of Quantitative Information: The Standard was addressed.

Quantitative information in the ATE evaluation was appropriately and systematically analyzed so that evaluation questions are effectively answered. Responses to close-ended questions were analyzed using SPSS. Presentations were clear in the

reports and followed accepted standards, based on our examination of the SPSS files and procedures used for importing data and analyzing this data.

Analysis of Qualitative Information: The Standard was addressed.

Qualitative information in an evaluation should be appropriately and systematically analyzed so that evaluation questions are effectively answered. The qualitative information from the survey was clearly presented and analyzed. Based on our examination of the Excel files and procedures used for importing data and analyzing this data, presentations were clear in the reports and followed accepted standards.

Justified Conclusions : The Standard was addressed.

The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them. The evaluation reports were very careful to report results completely and in a clear manner. Conclusions were largely descriptive and well supported by the quality of the data.

Impartial Reporting :The Standard was addressed.

Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings. As above, reporting was straightforward, comprehensive, and non-judgmental.

Metaevaluation: The Standard was addressed.

The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

The Evaluation Center decided to contract out a metaevaluation in order to assess the success of the ATE evaluation survey. This metaevaluation of the web-based survey was assigned to the Kercher Center for Social Research. It also contracted to have the site visit phase, a check of the survey, metaevaluated.

Checklist for Applying the Standards

To interpret the information provided in this form, the reader needs to refer to the full text of the Standards as they appear in "The Program Evaluation Standards by the Joint Committee on Standards for Educational Evaluation. A summary of the Standards appears in narrative form in the Applying the Standards section of this report. The Standards were consulted and used as indicated in the table below:

Table 1: Checklist for applying the Standards

	The Standard was addressed	The Standard was partially addressed	The standard was not addressed	The Standard was not applicable
U1 Stakeholder Identification	X			
U2 Evaluator Credibility	X			
U3 Information Scope and Selection	X			
U4 Values Identification	X			
U5 Report Clarity	X			
U6 Report Timelines and Dissemination	X			
U7 Evaluation Impact	X			
F1 Practical Procedures	X			
F2 Political Viability	X			
F3 Cost Effectiveness	X			
P1 Service Orientation	X			
P2 Formal Agreements	X			
P3 Rights of Human Subjects				X
P4 Human Interactions	X			
P5 Complete and Fair Assessment	X			
P6 Disclosure of Findings	X			
P7 Conflict of Interest	X			
P8 Fiscal Responsibility				
A1 Program Documentation	X			
A2 Context Analysis	X			
A3 Described Purposes and Procedures	X			
A4 Defensible Information Sources		X		
A5 Valid Information	X			
A6 Reliable Information	X			
A7 Systematic Information	X			
A8 Analysis of Quantitative Information	X			
A9 Analysis of Qualitative Information	X			
A10 Justified Conclusions	X			
A11 Impartial Reporting	X			
A12 Metaevaluation	X			

Principles for Constructing Web Based Surveys

What follows is an application of the Internet Survey Guidelines, as developed by Don Dillman (2000, pp. 376-400), in his "Mail and Internet Surveys: The Tailored Design Method." Therefore, the structure of this section is organized around Dillman's guidelines, excerpts of which are used throughout this section. This narrative is summarized on p. 29, Table 2.

Principle 11.10: The guideline was partially addressed.

"Introduce the Web questionnaire with a welcome screen that is motivational, emphasizes the ease of responding, and instructs respondents about how to proceed to the next page."

- Although a welcome page (actually two: the login and the welcome page) is provided, its message could be more motivational. Furthermore, more instructions on how to proceed to the next page are needed as detailed below.

Login Page

The introduction on the welcome page could be clearer. In particular, albeit simplistic, instructions on how to proceed could be included; for example, "Please enter your project's NSF number, your username, and password, as mentioned in the email sent to you on... and then click on the Login button to proceed to the main page..."

- Please login does not follow the proper "tab" sequence for those who are used to using the keyboard for moving from field to field. In particular, after entering NSF #, the tab key should take the user to the next field (username) and not to the

"login" button as it does now. This only happened in Explorer and not on Netscape. Therefore, perhaps there should be a notice on the login page recommending the use of Netscape.

Welcome Page

- Instructions under "Please bookmark this page" could be larger. Also bookmark notice could be on previous page (the "login" page) since, that's where the bookmark will take the user next time he/she uses it.

"logout", "Help", and "Survey Info"

- These links that are now part of the Welcome Page could become standard buttons on ALL pages (maybe a frame or part of the top banner).
- There should be a confirmation dialog window asking the user to confirm their wish to log out, and informing them that they will have to go through the process of logging back in again after they log out.

Change Delegation of PI Responsibilities

- The introduction is too "wordy" (too long a narrative). Information included in the introduction is somewhat overlapping, includes some redundancy, and includes some conflicting information. Overall, the introduction is not easy to understand, and not very user friendly. For example, instructions on reassigning particular sections vs. reassigning the whole project are not clear.

- If one attempts to delegate responsibilities, and makes an incorrect entry, a message pops-up and says: "Please fill your correct email address." Still it is not one's own email address that the person responsible fills-out, but the email of the person to be delegated responsibility.
- Although one cannot duplicate the email of the already existing PI, one can choose one from the drop-down menu.
- It is useful that the survey will not allow the user to continue if not all fields are filled out.

Survey Info Page

- Information under "Survey Info" ("About this survey") is exactly the same as the content of the login page (including the links).

Browsercheck Page

- The "browsercheck page" link should open a new window. This would enable the user to view the information required to carry out the check (username and password) while typing them on the separate window. If this is not acceptable, then it would be helpful if the "browsercheck page" included a "back to the survey" link. Note that going back to the survey after checking one's browser for compatibility is NOT a click away. One must actually be experienced enough to be able to bypass the last page visited in order to go back to the survey (by double-clicking on the browser's "back" button or by bypassing the last page visited through the browser's back button's drop-down menu).

Principle 11.11: The guideline was addressed.

"Provide a PIN number for limiting access only to people in the sample."

- A PIN number is provided, thereby limiting participation only to the people in the sample.

Principle 11.12: The guideline was addressed.

"Choose for the first question an item that is likely to be interesting to most respondents, easily answered, and fully visible on the welcome screen of the questionnaire."

- Indeed, the questionnaire's first section addresses "Basic Information". Response fields are clear and drop down menus include a "please select an option" message.

Principle 11.13: The guideline was addressed.

"Present each question in a conventional format similar to that normally used on paper self-administered questionnaires."

- Although features allowed by Internet language are used in the web-based survey, to maximize efficiency, the general format of the survey resembles the format normally used on paper (self-administered) questionnaires.

Principle 11.14: The guideline was addressed.

"Restrain the use of color so that figure/ground consistency and readability are maintained, navigational flow is unimpeded, and measurement properties of questions are maintained."

- The ATE survey uses a beige background color that does not affect readability.

The use of color on the web-based survey instrument is light so that readability is maintained, monitoring of progress is facilitated, and navigation is easy and user friendly. In particular, the use of a light beige color for background maintains readability, while the use of tri-colored progress gauges on the main "Survey Questions" page helps users monitor their progress, without impeding the navigational flow of the instrument.

Principle 11.15: The guideline was partially addressed.

"Avoid differences in the visual appearance of questions that result from different screen configurations, operating systems, browsers, partial screen displays, and wrap-around text."

- Although overall the survey instrument functions well with the two most popular web-browsers, Netscape and Microsoft Explorer, minor differences are evident. In general, the instrument functions perfectly in Netscape, with a couple of noticeable glitches in Microsoft Explorer. For example, Logging in with Microsoft Internet Explorer at times may be problematic, particularly when the user makes a mistake in entering his/her username and password. When this happens, the system may continue giving the user an error message (wrong

password), even after the proper correction is made on the username and/or password. In short, it seems that at times the program does not allow refreshing the login information on Microsoft Internet Explorer.

- Column width and text format are well within the most commonly used monitor specifications. A disclaimer on preferred browsers and screen specifications is presently on the Help page, whereas it could be more helpful if it was on the welcome page as well (or a shorter version of it). Furthermore, as it stands today, the survey instrument works better with Netscape than with Microsoft Explorer, although the differences are minimal and do not significantly affect the functionality of the survey process. Therefore, perhaps Netscape should be presented as the preferred browser.
- In Microsoft Explorer, the "briefly describe" questions have different widths in different sections of the survey.
- Monitor settings: the survey recommends 800x600 or higher but 800x600 presents problems such as the need for horizontal scrolling with the effect that the questions disappears by the time the user reaches the appropriate response category (e.g., Collaboration Section q. 3, PI Overview q. 1) and/or vertical scrolling that eventually obstructs the response category labels. The best screen setting seems to be 1024x768.
- It might be useful to have a note at the bottom of the login page advising as to the optimum web-browser and screen configuration. As it is now, this information (and more) exists on the help page (FAQ).

Principle 11.16: The guideline was partially addressed.

"Provide specific instructions on how to take each necessary computer action for responding to the questionnaire, and give other necessary instructions at the point where they are needed."

- As mentioned in other parts of this report, this seems to be an issue where there is some room for improvement. Specific instructions on how to proceed through each step of the survey are not always provided, and when they are, they are not always clear, or simple enough, especially for participants who are not very fluent in computer usage or windows navigation. Two further suggestions seem appropriate concerning the present state of the "Help/FAQ Page." Firstly, it would be best if the survey web site could provide "Context-Sensitive Help" instead of merely a collection of "Frequently Asked Questions;" secondly, a help page organized around a subject index could be more useful than FAQ, although perhaps both would be best. However, to address this shortcoming, The Evaluation Center added a Helpful Hints document on its ATE web page in year 2. This document is organized around a subject/function index, and survey respondents were encouraged to print it out and use it as a "how to" guide as they were working with the survey. While this document may be helpful, it does add another step for the survey respondent.
- Dillman insists on instructions of the type "click here" on links, and careful instructions on how to proceed on each step. This is NOT done consistently.
- "Links" on certain pages (i.e. welcome/introduction page, and Change delegation of PI responsibilities pages) are too obscure (small) and could be

standardized (maybe create a standardized place on the page and place them there consistently throughout the site. Still, it is very helpful that the links are "external". Floating windows with help and definitions are very helpful.

Principle 11.17: The guideline was addressed.

"Use drop-down boxes sparingly, consider the mode implications, and identify each with a 'click here' instruction."

- Drop-down boxes are sparingly used (except in the "Materials Development" section, where they are appropriately used more extensively). Where they are used, each is identified with a "please select an option" title.

Principle 11.18: The guideline was addressed.

"Do not require respondents to provide an answer to each question before being allowed to answer subsequent ones."

- Respondents are not required to answer to each question before being allowed to answer subsequent ones except on the "Change Delegation of PI Responsibilities" page, where the program will correctly not allow the user to complete the change unless all required fields are completed. Furthermore, a "does not apply" response is included where appropriate.

Principle 11.19: The guideline was addressed.

"Provide skip directions in a way that encourages marking of answers and being able to access the next applicable question."

- Whole sections are able to be independently accessed and completed.

Principle 11.20: The guideline was addressed.

"Construct Web questionnaires so they scroll from question to question unless order effects are a major concern, or when telephone and Web survey results are being combined."

- The survey instrument is constructed so it scrolls from question to question within each section of the questionnaire.

Principle 11.21: The guideline was addressed.

"When the number of answer choices exceeds the number that can be displayed in a single column on one screen, consider double-banking with an appropriate grouping device to link them together."

- Double banking is used where appropriate.

Principle 11.22: The guideline was addressed.

"Use graphical symbols or words that convey a sense of where the respondent is in the completion process, but avoid those that require significant increases in computer resources."

- Tri-colored (Red = No Answers Saved; Yellow = Some Answers Saved; Green = Section Completed) progress gauges are placed on the "Survey Questions" main page, in a way that does not impede the flow of navigation. Furthermore, these gauges do not seem to require significant computer resources. A fourth gauge (black) marks sections that do not apply.

Principle 11.23: The guideline was addressed.

"Exercise restraint in the use of question structures that have known measurement problems on paper questionnaires, such as check-all-that-apply and open-ended questions."

- The survey includes open-ended questions but for reasonable purposes to allow reporting of innovative program-specific information.

Checklist for Applying the Principles for Web-Based Surveys

To interpret the information provided in this form, the reader needs to refer to the full text of the guidelines as they appear in Don Dillman's (2000) (2nd ed.) "Mail and Internet Surveys: The Tailored Design Method", a summary appears in narrative form in the Principles for Constructing Web Surveys section of this report.

The guidelines were consulted and used as indicated in the table below:

Table 2: Checklist for applying the internet survey guidelines

	The Guideline was addressed	The Guideline was partially addressed	The Guideline was not addressed	The Guideline was not applicable
Principle 11.10		X		
Principle 11.11	X			
Principle 11.12	X			
Principle 11.13	X			
Principle 11.14	X			
Principle 11.15		X		
Principle 11.16		X		
Principle 11.17	X			
Principle 11.18	X			
Principle 11.19	X			
Principle 11.20	X			
Principle 11.21	X			
Principle 11.22	X			
Principle 11.23	X			

Individual Survey Sections

General Comments

- Although at times the site seems not very user friendly, or lacking proper instructions, or just cluttered with information and links, the questionnaire itself is quite user friendly, with sparingly used drop-down menus, and context sensitive links. Introduction with definitions and instructions in each section of the questionnaire is very useful and easy to follow, as is the "reassignment" section, and the "this section does not apply" button.
- It may be useful to have a PRINT button (or at least printing instructions) in the beginning and end of each questionnaire section.

Initial Survey Questions Page (GO TO SURVEY QUESTIONS)

- Instructions on "questions" page are not well placed nor are they well structured.
- Java script with help on "Closing Survey" is very helpful and not intrusive.

However, we did not close the survey. From our interviews, we understand that the directions and reaction of the system upon closing do not mesh and an additional email with more explicit directions was needed to assist some respondents in properly closing their surveys.

MONITORING

- Instead of a small introduction, the first line under this heading states that "confidentiality of responses to this section will be provided to the extent allowed by law." This needs clarification and needs to be accompanied by a small introduction to the section (as in other sections).

- Although Dillman recommends it, we understand from Evaluation Center staff that the “Not Applicable” (“NA”) option was purposely left out in this section as all projects are expected to have some level of interaction with NSF and to file FastLane reports.
- "Back" buttons are placed both on the top and on the bottom of the "answer" pages; this is very helpful, as the user has to scroll down to reach the end of each page. On the "view" pages though, "back" buttons are only placed on the top of each page. It would be helpful to place them on the bottom as well.

COLLABORATION

"answer mode"

- Question three includes many rows and so the user must scroll down and then the column headers get hidden. They could be included on top of each of the four sections (of rows), as in Question four.
- Question 5: This question addresses what you consider to be the most important products and/or results of your collaborative efforts - does this imply that the user should be able to choose more than one?

MATERIALS DEVELOPMENT

- Question three includes many rows and so the user must scroll down and then the column headers get hidden. They could be included on the bottom of the question as well as on the top.

PROGRAM IMPROVEMENT (ALL programs)

- Question 4: Which of the following options does your specified center/project program offer (check all that apply)? - The question mark should be placed after the parenthesis.

THE PRINCIPAL INVESTIGATOR'S OVERVIEW OF THE

CENTER/PROJECT

- Question 3a. Did you collect information to determine workforce needs (i.e. a workforce needs assessment) before submitting your proposal to NSF – The question needs a question mark.
- Different open-ended response fields have different widths in Microsoft Explorer.
- Ethnicity/race question includes racial categories that are outdated, like “White” and “Black”.

CONCLUSIONS AND RECOMMENDATIONS

By Dillman's principles or any other standards for survey/web-survey design with which we are familiar, The Evaluation Center's web-based survey is an exemplary product that is used appropriately and is subjected to regular and rigorous scrutiny and improvement. We have nevertheless offered specific suggestions for consideration at several points. Most of these are suggestions for minor modifications in the form (not content) of the survey as well as ideas for better instructions in each step of the survey. In combination these might help to improve the "user-friendly" feel of the survey and therefore combat the perception of the survey as too complex and lengthy (see Nonresponse Study for the Year 2000 Survey of ATE Projects and Centers). Finally, browser specific behaviors suggest one might modify the browser/screen disclaimer and/or the help page to address these issues and show the disclaimer on the welcome page as well as the help page.

BIBLIOGRAPHY

Birnbaum, Michael H. (ed., 2000). Psychological Experiments on the Internet. Academic Press.

Dillman, Don (2000, 2nd ed.). Mail and Internet Surveys: The Tailored Design Method. John Wiley & Sons.

The Joint Committee on Standards for Educational Evaluation (1994) (2nd ed.). The Program Evaluation Standards: How to Assess Evaluations of Educational Programs. Sage.

Shneiderman, Ben (1998, 3rd ed.). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley.