

OERL Survey Design, Methodology, Administration and Results

Judith Fusco, Heidi Skolnik, Geneva Haertel, and Harold Javitz

SRI International

Nick Smith

Syracuse University

Elisabeth Thurston

Teachers College, Columbia University

Paper presented at the annual meeting of the American Educational Research Association, April, 2005, Montréal, Canada. Send correspondence to Geneva Haertel, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025. This research was supported by contracts from the National Science Foundation (REC-9912172 and NSB-0353574).

Do not cite or quote without permission.

OERL Survey Design, Methodology, Administration and Results

This paper documents the process employed to develop and administer a methodologically sound set of surveys to evaluate the Online Evaluation Resource Library (OERL) Web site. Results from the surveys are then presented. The survey work described in this paper was to serve as formative information for OERL and was not designed for summative purposes and our results serve to inform us about our users. The foci of the OERL surveys were on practical issues that are important to developing a Web site that functions as a resource for professionals. In the design of the surveys it is noted that OERL is a large, multi-faceted Web site that serves many purposes and audiences (Zalles, 2002). To more fully understand the many purposes for which the site can be and is used and the audiences that use it, three surveys were developed, aimed at three different audiences.

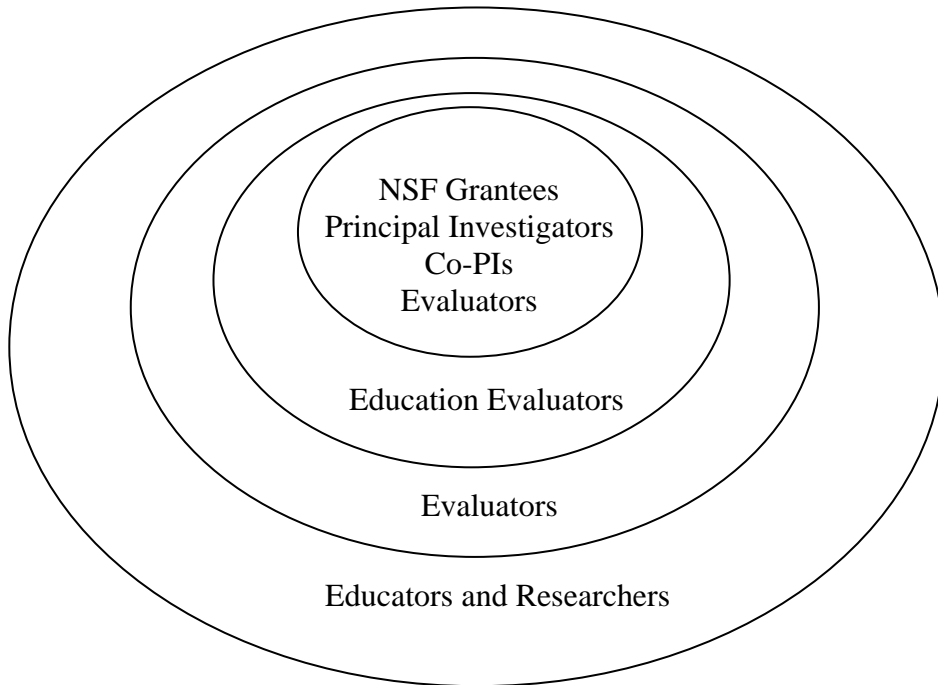
Defining the Survey Population

The original goal of the survey work was to understand the overall OERL audience, namely, who uses the Web site, how they use it, and their satisfaction with it. To focus the survey efforts, a diagram was developed that describes the possible users and audiences of the OERL Web site (see Figure 1). The concentric circles represent the various OERL audiences. The innermost circle, NSF Grantees, represents the group for whom the site was intended when it was conceived and developed: those participating in NSF-funded projects with an evaluative component. The second circle, a larger audience, is that of educational evaluators. Since OERL contains examples drawn from education evaluation projects, its resources are likely to be of use and interest to education evaluators, beyond those involved in NSF projects, including graduate students in the field. The third circle represents all evaluators. While all of OERL's resources are specific to education, the evaluation methodologies and resources presented in OERL may help evaluators in other fields and contexts. In addition, the Professional Development Modules are relevant to evaluators outside of education. Finally, the outermost circle represents educators and researchers who are not evaluators, but who need to develop a greater understanding of evaluation.

The conceptual diagram also was used to think about and to develop feedback procedures from OERL users, given the differences among them. When considering all the differences in the populations in the diagram, the decision was made to develop three surveys. The first survey focuses on current OERL users in general and could include users from each of the concentric circles in Figure 1. This survey was designed to understand the behaviors of users on the site, and to determine who is using the Web site.

Though understanding information about the current OERL user is an important step to enhancing the site, we also wanted to understand what non-users think. Posting a survey only on the Web site would not yield information about evaluation professionals for whom the site is intended but who have not discovered OERL, or those who may have tried to use OERL once or twice and then did not return. Understanding this non-user or non-returning audience was an important activity to pursue to help improve OERL. The two additional surveys that were developed were intended to help understand how we are reaching NSF grantees and educational evaluators, and their interest in and satisfaction with the Web site.

Figure 1. Conceptual Diagram of OERL Target Universe



The Survey Instruments

The three surveys employed in this study are the: (1) OERL User Survey, (2) NSF Grantee Survey, and 3) Education Evaluator Survey. Table 1 describes the features of each survey instrument, survey population universe, and sample of respondents

OERL User Survey

The first survey conceptualized and developed was the OERL User Survey. The OERL User Survey collected information on general OERL use and use of specific sections of the Web site (Plans, Reports, Instruments, Professional Development Modules, and Other Resources). In addition, information was gathered about the Web site's impact, usefulness, and customer satisfaction, as well as respondent background and evaluation experience. The user survey consisted of 64 items. The survey was administered online. Skip technology was employed in this online survey so that if a person had not visited a section of the Web site, they were not asked questions about it. The skip technology helped to prevent the survey from being too long for respondents who had not visited all sections of the site.

NSF Grantee Survey

The NSF Grantee Survey was designed to determine whether or not NSF Grantees, the original intended audience for OERL's evaluations resources, are using and benefiting from the OERL site. This survey gathered information on OERL and other Internet usage patterns, impact of the OERL Web site on current users, its potential perceived value to non-users, and customer satisfaction. The survey also gathered information about respondent background, and evaluation experience. The NSF Grantee Survey consisted of 30 close-ended- items and four open-ended items.

Education Evaluator Survey

The Education Evaluator Survey was designed to reach those members of the OERL site's broader target audience who do not currently use OERL. The topics covered included usage patterns related to Web-resources other than OERL, customer satisfaction, initial impressions, respondent background and evaluation experience. This survey consisted of 21 closed- and four open-ended items.

Table 1. Characteristics of Survey Instruments, Survey Populations, and Sample of Respondents

Survey Characteristics	OERL User Survey	NSF Grantee Survey	Educational Evaluator Survey
Purpose	To understand the overall OERL audience, including who is using the Web site, how they are using it, and how OERL can be improved	To understand how OERL is or is not meeting the needs of its original target audience—NSF grantees and their evaluators—including whether or not this group is using OERL and how those who are not using it might benefit from its use	To understand what evaluation professionals who have not become regular users of the OERL site perceive as the potential value of OERL
Target Audience	All OERL users	The audience for whom the OERL site was developed; recipients of NSF grants within selected program areas who do or do not use the OERL site as a resource	Education evaluators who do not currently use the OERL site as a resource
Universe	Unknown	1000	~2750
Sample Size	Unknown	493 (actual after bad addresses removed)	1335 (actual after bad addresses removed)
Minimum Number of Respondents	350	280	490
Topics Covered	OERL usage patterns, evaluation experience, usefulness of OERL components, impact of OERL use, familiarity with technology, other Web resource usage patterns, satisfaction with components of OERL, background	Web resource usage patterns, evaluation experience, impact of OERL use, familiarity with technology, other Web resources, satisfaction with OERL, background	Web resource usage patterns, evaluation experience, potential value of the site, and background

<p>Item Types</p>	<p>Closed-Ended</p> <p>25 Likert-scaled items (4-point + no opinion); 16 Yes / No; 1 Rank; 10 Categorical</p> <p>9 Mark all that apply</p> <p>Open-Ended</p> <p>3 open-ended</p>	<p>Closed-Ended</p> <p>6 Likert-scaled items (4-point + no opinion); 10 Yes / No; 1 Rank; 17 Categorical</p> <p>6 Mark all that apply</p> <p>Open-Ended</p> <p>4 open-ended</p>	<p>Closed-Ended</p> <p>6 Likert-scaled items (4-point + no opinion); 4 Yes / No; 1 Rank; 7 Categorical</p> <p>3 Mark all that apply</p> <p>Open-Ended</p> <p>4 open-ended</p>
<p>Administration</p>	<p>Invitation to take the survey posted on the OERL Web site; respondents take survey online</p>	<p>Survey and invitation to take survey mailed to Principal Investigators of NSF grants from selected program areas; respondents have the option of completing an online or paper-based survey</p>	<p>Survey and invitation to take survey mailed to randomly selected education evaluators from the American Evaluation Association (AEA) and the American Educational Research Association (AERA); respondents have the option of completing an online or paper-based survey</p>

Administration Procedures

Each of the three surveys required a unique administration process.

OERL User Survey

The OERL User Survey was administered online only. An appeal appeared on the OERL home page and in the banner of every page on the site encouraging visitors to take the OERL tour. The only qualification for responding to the OERL User Survey was that the respondents have used the OERL site. We did not require that the site had been used for a certain amount of time, but we did inquire about the respondent's use history with the site. For this survey, the universe size could not be estimated. On average, OERL receives approximately 2,900 visitors from unique IP addresses each month as recorded by the OERL Web server transaction logs. Though it is known how many different IP addresses have visited OERL, the actual number of users of the OERL site (e.g., population) is unknown (see below for a discussion about the limitations of Web server transaction logs). This is the case with all Web sites unless authentication (username and password) is used to determine visits (Bauer, 2000). Because the size of the population of OERL users is unknown, we collected data from as many users as possible.

The OERL User Survey, was administered completely online, and was accessible from the Web site for almost 10 months. In first three months the survey was online, there was no incentive for respondents to take the survey, and the response rate was very low (only 22 respondents in three months). The first survey appeal on the Web site read:

Have a minute?
Have feedback to share?
Please take our brief online survey.

We discussed the low response rate with our program officer, and with him, decided to offer an incentive of a book (the *NSF User Friendly Evaluation Handbook*) to motivate survey responses. The second appeal that appeared on our Web site was as follows:

WANT TO TRADE?

If you give us 10 minutes of your time, we will send you a free copy of the NSF USER FRIENDLY EVALUATION HANDBOOK!

We'll trade this valuable Handbook for your quick response to our survey about the quality of this OERL Web site. We will be able to improve this Web site based on your suggestions, and you will have a useful evaluation resource for those times when you are not near your computer. We both win!

After this incentive was implemented, the response rate increased to approximately 65 respondents a month (up from seven). Respondents could choose to sign up to receive the User Friendly Handbook after taking the survey. Their requests for the book were kept separate from their survey responses and there was no way to for sure match a person who requested the handbook to their survey responses.

NSF Grantee Survey

We selected a random sample of 510 grantees who had received awards in 2002 – 2003 from DUE (did not include interns) within the EHR Division of NSF. Both pre-contact and the number of subsequent contacts can influence response rate positively (Cook, 2000), so, each NSF grantee was sent: (1) an introductory postcard announcing that a survey would be sent; (2) one week later, a packet containing the survey, an explanatory letter, and an invitation to take the OERL online tour; and (3) two-weeks later, a reminder postcard; (4) a second packet containing the survey and a reminder letter to all of those who had not responded; (5) a 'third appeal' postcard mailed approximately three weeks after the second survey packet. To make responding to the survey as convenient as possible, users had the option of responding to an enclosed paper-based survey or responding via the Web. The invitation to take the OERL online tour, which was sent to each potential respondent along with the NSF Grantee Survey, was included in order to familiarize those NSF grantees who had not previously used OERL with the range of resources available on the site. To increase the survey response rate for this survey, in the second survey packet mailing, the offer for the *NSF User Friendly Evaluation Handbook* was

made to NSF Grantee survey respondents. The NSF Grantee Survey included questions about OERL's impact, for those respondents who had used it before, and its perceived potential value, for those respondents who had not used it previously.

Education Evaluator Survey

The target audience of this survey was education evaluators who do not currently use OERL. We believe the population of education evaluators would find OERL useful, but know not all have become a part of the OERL user community. Since we wanted to contact those who do not currently use the site, we could not use the Web site to reach them, so an alternate means to reach them had to be developed. We assumed that a large proportion of education evaluators are members of some type of professional organization, therefore names of education evaluators were randomly selected from the 2003 mailing lists of two professional organizations-- the American Evaluation Association (AEA) and Division H (Evaluation) of the American Educational Research Association (AERA).

The size of the population for education evaluators was estimated to be approximately 2,750. This was based on an average of the AERA Division H (Evaluation) membership, which is approximately 2,000 members, and the overall membership of AEA, which is 3,500. Estimating the population of educational evaluators at 2,750 may be an overestimate, since the overall number of members in AEA, not just those who are in educational evaluation, was used (AEA does not have a topical interest group specific to educational evaluation, so we were unable to get an actual number of their education evaluator members).

For the Education Evaluator Survey, an adequate response was achieved through the same process as the NSF Grantees survey: (1) an introductory postcard announcing that a survey would be sent; (2) one week later, a packet containing the survey, an explanatory letter, and an invitation to take the OERL online tour; and (3) two-weeks later, a reminder postcard; (4) a second packet containing the survey and a reminder letter to all of those who had not responded; (5) a 'third appeal' postcard mailed approximately three weeks after the second survey packet. In addition to an enclosed paper-based version, users also had the option of responding to the survey online. Like the NSF Grantee Survey packet, the Education Evaluator Survey packet mailed to potential respondents included an invitation to take the OERL online tour. The tour

was especially important for this group of respondents, so they could give feedback about how they perceived the site upon first being introduced to it. Again, like in the procedures for the NSF Grantee survey, we offered the NSF *User Friendly Evaluation Handbook* as an incentive to take the Education Evaluator survey in the second survey packet mailing.

OMB Clearance

We were required to undergo OMB Clearance on our surveys as per the Paperwork Reduction Act signed in 1995. As the OERL surveys are “customer satisfaction surveys,” we were eligible for clearance under NSF’s generic clearance from OMB. Customer satisfaction surveys are reviewed by OMB in an expedited fashion as feedback from customers is very important to the improvement process. After we received our clearance we began collecting data.

Methodological Challenges of Using Online Surveys

In addition to the challenges and goals the OERL team had in developing the surveys, there are also methodological challenges in using online surveys. The following two sections describe issues that need to be considered and addressed when using online surveys. The first section considers challenges in using online surveys, and the following section discusses Web transaction log files, about which there are a myriad of misconceptions, and their limitations.

Administering surveys online has a variety of potential benefits, such as a reduction of costs, the elimination of data entry if the survey responses are compatible with existing software, and a simple way to pre-notify and follow-up with respondents. Even so, researchers have been slow to adopt technology for survey administration, with the major push for Web surveys coming from the computer-programming world (Dillman & Bowker, 2001). To date, the methodological features of Web surveys, including principles of Web survey design to reduce survey error, have not been carefully examined. Below we identify several sources of error that complicate the administration of the OERL User Survey.

Sources of Error in Web Surveys

Like traditional paper surveys, Web surveys have limitations. Unlike traditional surveys, these limitations have not been sufficiently addressed (Dillman and Bowker 2001). Dillman and Bowker (2001) identified four main sources of survey error, including coverage error, sampling error, measurement error, and non-response error. They discuss how those sources of error impact a Web survey specifically. It was our effort to reduce these four types of error that led to the development of three surveys, rather than just one, to determine the satisfaction of three different user groups with the OERL Web site. All four types of survey error, described below, need to be reduced in order for a Web survey to produce legitimate results.

Coverage Error. Coverage error, resulting from not giving everyone a known nonzero chance to be included in the survey, manifests itself in online surveys due to varying access to and experience with the Internet. For these reasons, Web surveys for professional or business groups tend to be more methodologically sound than those purporting to represent the general population (Shannon, Johnson, Searcy, & Lott, 2002). Thus, we judged that all three surveys, which were administered to professionals who are likely to be frequent Web users, would control coverage error reasonably well. In fact, on two of our surveys (the NSF Grantee and Education Evaluator survey), respondents could opt to take the paper-pencil version of the survey if they preferred.

Sampling Error. One of the main benefits of online surveys, as opposed to traditional paper surveys, is the relative ease with which one can reach a large population, fairly inexpensively. To avoid sampling error, representativeness, not just achieving a high response rate, needs to be taken into consideration. In random samples of defined populations, greater numbers of responses suggest greater survey accuracy. With online surveys, especially those in which the instrument is posted online and can be accessed by an undefined population, large numbers of volunteer respondents do not legitimize the use of inferential statistics (Dillman & Bowker, 2001). Thus, we have not attempted to generalize our results to the universe of OERL users.

Measurement Error. In addition to common causes of measurement error—poor item wording, unclear formatting—Web surveys face additional challenges as a result of the medium. In a

1999 review of Web surveys, Bowker found four common causes of measurement error based on the construction of Web surveys, including:

- Varying physical distances between points on attitude scales, due to changes in respondent's screen configuration or differences in browsers;
- Different background colors for points on an attitude scale, skewing the apparent distance between radio buttons;
- Differing construction procedures for wrap-around text, causing the distances between points to change;
- Inability to view the entire question or answer choices, based on screen size.

We designed our online surveys to minimize these sources of measurement error.

Nonresponse Error. Non-response error has been a problem for Web surveys. This can be explained, at least in part, because methods for increasing response rate for Web surveys comparable to methods employed for traditional paper-based surveys are just beginning to be explored. Like paper-based surveys, it appears that the number of contacts, both pre-survey and follow-up, can greatly increase response rate (Cook, Heath, & Thompson, 2000; Kittleson, 1997). Additionally, issues such as how long it takes a survey to load, personal questions, or complex displays, can also negatively influence response rates (Dillman, Tortora, Conrad, & Bowker, 1998; Solomon, 2001). Thus, we provided a banner advertising the OERL User Survey on the home page of the OERL site, as well as on other frequently hit pages, and offered an incentive (the *NSF User Friendly Handbook*) as a means of reducing non-response error. All three surveys contained few personal questions, loaded quickly, and did not employ overly complex displays. The surveys were formatted by an experienced Web design team, to reduce non-response error.

Limitations of Using Web Transaction Logs

We include this section, because of the misconceptions that exist about how easily and accurately Web transaction log files can document exactly who visits a Web site, how long they visit, and what they do. Limitations of Web transaction log files, in turn, affect the inferences that can be made about site usage. A Web transaction log is a file that a Web server keeps that contains a record of all the pages it served, what time it served them, to what machine it sent the

page. A Web transaction log can also tell you how people got to your Web site. It can tell you if the person did a Web search through a search engine, the terms they used, or if they clicked a link from another page on the Internet. However, there are many limitations. Many people imagine (wrongly) that a Web transaction log can tell everything about the person visiting your site, and that is not true.

The primary limitations of Web transaction logs are as follows:

- Identity of users: Even though IP addresses are captured by Web transaction logs, and even though IP addresses are unique, the identities of users of the site cannot be determined from most Web logs. Even when IP address look-ups are successful, IP addresses do not include the name(s) of the person using them. At best, one may find out the company the person works for, or the Internet Service Provider (ISP) used. Even the geographical location (city, state) of the person may not be resolved correctly (as, for example, when a company or large ISP provider associates their headquarters' geographical location with IP addresses that actually serve users at a different geographical location).
- Number of unique users: Even when an IP address corresponds to a single computer on a network, it may not be true that this corresponds to a single person. A computer may be used by one person, or it may be used by many people (e.g., a computer in a computer lab). Furthermore, Network Address Translation (NAT) and proxy servers may lead to a severe underestimation of the number of visitors coming from one site. With NAT, all users from one site show up as coming from one IP address. Laptop computers that connect from multiple locations may be counted more than one time as they may have more than one IP address, for example, one IP at home and one IP address at work. In addition, a user's IP address may switch as a user changes computers, moves a portable computer to different locations, or as a wireless network rebalances itself by changing the station requested by a stationary computer.
- How many, which and the order of pages viewed by a visitor: A personal computer can cache previously viewed files from Web sites, and if a person views a page from their cache of files, it is not recorded in the Web transaction log file. In addition to local caches, ISPs sometimes have cache sites for their users so that the user's connection

seems faster. Thus it is impossible to know how many or what pages were visited in a session. For example, in the OERL Web log files, a large percentage of sessions (approximately 60%) have only “one page” recorded as visited in the Web server log files with only 17% of those 1-page visits being visits to the home page. It is conceivable that many of the “sessions” which are recorded as one page are actually much longer than one page, but that the visitors visited only one page that was not already cached. Because of caching, it is also not possible to know the order of all pages (or path) a person visited in a session.

- How long each page was viewed: Some Web transaction log analyses calculate the lag time between pages served as an indicator of how long the user viewed a page. This is not a reliable indicator, as users could be engaged in another task (e.g., a phone call) during the time that the Web log indicates the person was looking at that page, or pages served from the cache might have been viewed in between the ones recorded in the log.
- Where the visitor entered and exited the Web site: The first and last pages recorded by the log are not necessarily the first and last pages viewed by the user. Some potential reasons for this are that the user visited pages from the cache at the beginning or end of the session or the user’s IP address changed mid-session.
- What the user actually viewed: This cannot be accurately known since individuals may not see or pay attention to some sections of a page and the page that is viewed may be obtained from the cache.
- Why the user visited the site: Sometimes log files are used to create “user scenarios” by attempting to extrapolate meaning from a user’s particular path. In addition to not knowing the reasoning behind the user’s visit to the page, this is problematic because of the other issues described above.
- When a session began and ended: Because of the presence of cached Web pages, it is not possible to know when a session begins or ends.

Given these limitations, it would be extremely difficult to gather data on the types of usage listed above. The best way to understand who is visiting a Web site—and the only way to determine the true number of users (Bauer, 2000)—is to require membership and authentication (username and password) on each visit. The best method for tracking page usage is to force each

page to dynamically update rather than be accessed from a cache (for example, by changing invisible characters on the page every few seconds). In addition, cookies could be used to distinguish multiple users on different computers sharing a single account and password. However, given the nature of OERL, we did not believe that requiring a username and password for accessing the information was appropriate. In addition, forcing dynamic updates could seriously degrade the user experience (by increasing the perceived amount of time necessary to load pages), and using cookies for tracking could be considered an invasion of privacy.

Recommendations for online surveys

Due to the limited nature of Web survey research, researchers choosing the Web as their mode of survey delivery face significant challenges in the development of a methodologically sound survey and administration plan (for more information on challenges of Web survey development, see the Shannon and Dillman papers). Our survey development process for the evaluation of user satisfaction with the Online Evaluation Research Library (OERL) encountered challenges, which were addressed through careful attention to sound methodological practices and multiple revisions of each survey. As a contribution to the survey development community, we have restated some of our challenges in reducing Web survey error into recommendations for others facing similar issues.

Coverage Error

Due to the target audience of our surveys, coverage error as a result of limited access or experience with the Internet is a small issue. Almost everyone who responded to our surveys, with the exception of those who have not used OERL, needed at least a basic familiarity with the Internet in order to have found and used the Web site. Even so, efforts to reduce, or to at least account for, coverage error were incorporated into the surveys.

Sampling Error

The majority of our challenges in developing methodologically sound surveys were in regards to our limited knowledge of our sample. For the OERL User Surveys, we did not attempt to generalize our findings to groups that we could not fully characterize.

Measurement Error

We revised each survey eight times to eliminate poorly worded or ambiguous items. In addition, the surveys were formatted by individuals with Web expertise in order to avoid common pitfalls associated with the online display of surveys.

Nonresponse Error

Our surveys were designed to ensure that we did not put an excess burden on respondents in terms of time or overly complex survey items and formats.

Web Transaction Logs

Because of the way our server and the OERL Web site was set up, we determined that we could not use our Web transaction log files for much more than quick indicators of a rough number of people visiting our web site. We also watch our transaction logs to see if there are trends occurring over time, of increases or decreases, or any significant change in the pattern we typically see.

What We Learned: Key Findings from the OERL Surveys

This remainder of the paper describes findings from the three surveys developed by the OERL team to better understand who the users of the OERL site are and how to better meet their needs. The three OERL Surveys were analyzed in November, 2004. The User Survey received 416 responses, the NSF Survey received 197 responses, and the Education Evaluator Survey received 433 responses.

Since the NSF Survey and the Education Evaluator Survey could be taken either online or with the paper-pencil version, they were first analyzed to determine if there were any differences in survey responses by mode of survey taking. For the NSF Survey, fifty-seven percent of the respondents ($n = 113$) completed the paper-pencil version, while the remaining 42.6% opted for the online version ($n = 84$). An analysis using cross-tabulations of survey taking mode (online versus paper-pencil) by all of the questions on the survey found no major systematic differences in responses across the two modes on any respondent background questions. A difference was found in that it was more likely for a person taking the online version to have taken the OERL

introductory tour than for a person taking the paper-pencil version. This finding was not surprising given that it would have been easy for online survey respondents to take the tour at the same time. In addition, we saw that people who took the paper-pencil version were more likely to have responded with “no opinion” on some of the questions asking them to rate the OERL resources, but most likely this is because fewer of the pencil-paper survey respondents took the tour and were not as familiar with OERL. For the Education Evaluator survey, sixty-five percent of the respondents (n = 283) completed the paper-pencil version, with the remaining 35% opting for the online version (n = 150). A preliminary analysis showed no systematic differences in survey response across the two methods.

We then analyzed all the common items from the three surveys by cross tabulation (item by Survey) together to determine if there were systematic differences on background e.g., level of evaluation experience, web usage, patterns of OERL use, and likeliness of using OERL in the future. The significant results from the common items among surveys are presented and discussed. In subsequent sections, significant results from the items on the individual survey are presented and discussed.

We found that level of evaluation experience the respondents of the different surveys did indeed differ ($P^2(6, N = 997) = 250.05, p = .00$) (See Table 2). Much of the difference is due to the respondents on the NSF survey who were Primary Investigators and not Evaluators. In the responses we see a larger percentage of “not an evaluator” from the NSF respondents. In addition, the NSF and User Survey respondents consider themselves less experienced than the Education Evaluator respondents with only 4% and 15% (respectively) responding “expert” in the NSF and User Survey versus 44% responding as “expert” in the Education Evaluator Survey. Respondents on the 3 surveys also differed significantly ($P^2(6, N = 967) = 197.9, p = .00$) in the number of years of evaluation experience they had. The years of experience followed a pattern similar to that reported for the respondents’ level of evaluation experience.

Table 2. Level of Evaluation Experience of Respondents by Survey.

Survey	level of evaluation experience								Total N
	Inexperienced		Proficient		Expert		not an evaluator		
	N	%	N	%	N	%	N	(%)	
ED Eval	24	5	196	46	185	44	14	4	419
NSF	72	37	79	41	9	.4	33	17	193
User	152	39	146	38	59	15	28	7	385
Total N	248	n/a	421	n/a	253	n/a	75	n/a	997

In addition to rating their level of evaluation experience higher, the Education Evaluator respondents had the highest percentage of respondents with formal training in evaluation (84%) compared to 24% for the NSF respondents and 56% for the User Survey ($\chi^2(2, N = 967) = 207.29, p = .00$).

In terms of how often the respondents used the web and related tools, no differences were found among the respondents of the three surveys. For all three surveys, approximately 94% of respondents said they used the WWW about daily, approximately 81% said they used search tools about daily, and approximately 34% said they participated in mailing lists, discussion groups, or newsgroups about daily. However, when asked about how often they access evaluation resources on the web, a difference among the survey respondents for the three surveys was found ($\chi^2(8, N = 973) = 91.74, p = .00$). It appears that the respondents to the Education Evaluator survey and the respondents to the User Survey access evaluation resources on the web in similar patterns, but the NSF respondents do not access them as often. Given that the respondents to the NSF Survey respond as “not evaluators” more often than respondents on the other two surveys, this is not surprising.

Table 3. Frequency of Accessing Evaluation Resources on the Web by Survey.

Survey	How often do you access evaluation resources on the web?										Total N
	Frequently		Fairly frequently		Fairly infrequently		Infrequently		Not at all		
	N	%	N	%	N	%	N	%	N	%	
ED Eval	67	16	118	28	101	24	107	25	32	7	425
NSF	14	7	14	7	38	20	79	41	49	25	194
User	70	20	87	26	80	23	76	21	41	12	354
Total N	151	N/a	219	n/a	219	N/a	262	n/a	122	n/a	973

Respondents on all three surveys ranked the same three NSF program areas as being of most interest to them. They were:

- 1.Underrepresented Populations
- 2.Curriculum Development
- 3.Teacher Education

In terms of outcomes, there was a significant difference between the survey taken and the likelihood of using OERL in the future($\chi^2(4, N = 995) = 20.43, p = .00$). Only 3% of everyone who took any of the surveys said no they would not plan to use OERL again in the future. The survey respondents on the User Survey said “Yes” (74%) rather than “Maybe” more often than the other respondents on the other two surveys said “Yes” (NSF, 57%; Education Evaluator, 64%). (See Table 4.) We believe this was due to being more familiar with OERL. A subsequent analysis further explores this (See page 31).

Table 4. Use of OERL in the Future by Survey.

Survey	Will you use OERL in the future						Total N
	Yes		Maybe		No		
	N	%	N	%	N	%	
ED Eval	269	64	129	31	20	5	418
NSF	109	57	73	38	9	5	191
User	282	74	99	25	5	1	386
Total N	660	n/a	301	N/a	34	N/a	995

In the next three sections we will examine the individual surveys and the results from them in more detail.

OERL User Survey Results

The OERL User Survey was posted online on February 2nd, 2004. The data discussed in this report was collected between the initial posting data and October 29th, 2004 and includes responses from 416 respondents. The OERL User Survey is a 64-item online survey. After tabulating results for each of the items, we identified key findings that help us further characterize the users of the OERL Web site and their impressions.

Description of OERL User Survey Respondents

The User Survey respondents tend to be inexperienced (39%) or proficient (38%). (See Table 2.) This is less experienced than the respondents on the Education Evaluator Survey and about the same level as the respondents on the NSF Survey. The User Survey only has 7% of respondents who say they aren't evaluators compared to 17% among the NSF Survey respondents. A little over half (56%) of the User Survey respondents have formal evaluation training. Forty-four percent of respondents characterized the evaluation approach or methodology that they use as mix of qualitative and quantitative, and 32.7% stated that less than one-quarter of their job required the use of evaluation skills.

Highlights of OERL User Survey Results

- Survey respondents generally gave several reasons for coming to the OERL Web site. The top four reasons were:
 1. To get a general sense of types of instruments used in an evaluation (41%)
 2. To learn about evaluation in general (40%)
 3. To find an example of a complete evaluation plan or report (30%)
 4. To find a specific instrument to adapt to their needs (29%).
- Most OERL users learned of the Web site via a search engine such as Google (27%).
- For 75% of the survey respondents, the day they took the survey was the first time they had visited the site. From their responses to the rest of the survey questions (see below in

section “Satisfaction with OERL” and Table 6), we know that the first impressions of the site were generally positive.

- The average number of reasons for “coming to OERL today” was 2.4 and did vary by level of evaluation experience with experts reporting fewer reasons for why they came to OERL today than inexperienced, proficient, or non-evaluators ($F(3, 379) = 10.47, p = .000$).
- The average number of reasons for why an evaluator would return to OERL was 4.5, with experts having fewer reasons to return than inexperienced or proficient evaluators. Non-evaluators did not differ significantly from any of the other groups in their number of reasons for returning ($F(3, 367) = 6.01, p = .001$). A t-test on the difference score between the number of reasons a person gave for coming to OERL today and the number of reasons they might return (average difference = 2.1) revealed a significant increase for the entire sample ($t(371) = 18.26, p = .000$). A subsequent Analysis of Variance (ANOVA) was conducted on the difference scores for the four groups. It revealed no significance in the difference scores by groups based on levels of experience ($F(3, 365) = .55, p = .65$). So even though experienced evaluators initially have fewer reasons than less experienced evaluators, the extent to which respondents recognize the potential usefulness of the resources that OERL offers seems to be consistent across varying levels of expertise if the increase in number of reasons is an indicator of potential usefulness. (See Table 5 for means).

Table 5. Average number of reasons for coming to OERL, for visiting OERL again and the difference between them plus associated statistics.

Level of evaluation experience		Average reasons for coming to OERL today	Average reasons for visiting again	Difference between two averages
Inexperienced	Mean	2.8092	4.8707	2.0272
	N	152	147	147
	Std. Deviation	1.73294	2.01797	2.27236
Proficient	Mean	2.3517	4.5786	2.2446
	N	145	140	139
	Std. Deviation	1.74217	2.18914	2.28366
Expert	Mean	1.4138*	3.5172**	2.1404
	N	58	58	57
	Std. Deviation	.89901	1.84715	1.75701
Not an evaluator	Mean	2.7143	4.4231	1.6923
	N	28	26	26
	Std. Deviation	1.80241	2.13866	2.47821
Total	Mean	2.4178	4.5175	2.1030
	N	383	371	369
	Std. Deviation	1.70398	2.10959	2.21659

* Experts have fewer reasons for initially coming than all the other levels of experience.

** Experts have fewer reasons than the inexperienced or proficient evaluators, but do not differ from non-evaluators.

Future Use

- In terms of plans to use OERL again, 73% of survey respondents said “Yes” they plan to use OERL again and 25% said “Maybe” they would use OERL again. Only 1.2% (5) respondents said “No.” The User Survey had the highest rate of “Yes” respondents.

Further examination shows no difference in the rates of “Yes” or “Maybe” responses by levels of evaluation experience (not an evaluator, inexperienced, proficient, and experienced) ($F^2(3, N = 384) = .29, p = .96$). (“No” responses could not be included in the above analysis as there were only a total of five.) There was no association between rates of Internet use and whether a respondent planned to use OERL again.

- Survey respondents provided, on average, four reasons for using OERL again. The top reasons were as follows:

1. To find a specific instrument to adapt to their needs (73%)
2. To learn about a specific topic in evaluation (56%)
3. To find an example of a complete evaluation plan or final report (54%)
4. To get a general sense of the types of instruments used in an evaluation (50%).

Satisfaction with OERL

- Of the survey respondents, 74% agreed that information could be found easily on OERL; 73% of respondents agreed that the evaluation information they found on OERL was of high quality; 26% agreed that they used OERL more frequently than other online evaluation resources; and 50% agreed that they liked the “look and feel” of the OERL site better than that of other online evaluation resources.
- The table below (Table 6) shows survey respondents’ satisfaction ratings on the different sections (plans, instruments, reports, and professional development modules) of the OERL Web site. Respondents were asked to complete the survey questions about the different sections of the Web site only if they had visited that section. The response rates for the different sections do differ and are indicated below.

Table 6. A Summary of OERL User Survey Judgments of the OERL Plans, Instruments, Reports and Professional Development Modules

Survey Questions	Plans (N = ~199)	Reports (N = ~121)	Instruments (N = ~159)	Professional Development (PD) Modules (N = ~72)
Number who visited resource	47.8% (N = 199): visited the plans	29.1% (N = 121): visited the reports	38.2% (N = 159): visited the instruments	17.3% (N = 72): visited the PD) modules
Usefulness of Resource	85.4% (N = 170): rated the plans as fairly or extremely useful	84.3% (N = 102): rated the reports as fairly or extremely useful	88.1% (N = 140): rated the instruments as fairly or extremely useful	88.9% (N = 64): rated the PD modules as fairly or extremely useful
Quality of resource	81.9% (N = 163): rated the quality as excellent or good	86.7% (N = 105): rated the quality as excellent or good	84.3% (N = 134): rated the quality as excellent or good	87.5% (N = 63): rated the PD modules as excellent or good
Clarity of resource	N/A	N/A	N/A	90.3% (N = 65): rated the PD modules as very clear or clear
Ease of navigation	87.9% (N = 175): rated the navigation of plans as very easy or easy	84.2% (N = 102): rated the navigation of reports as very easy or easy	96.2% (N = 153): rated the navigation of instruments as very easy or easy	90.3% (N = 65): rated the navigation of PD modules as very easy or easy
Use of OERL in preparation of plan, report or evaluation instrument	74.3% (N = 148): indicated that they had <u>not</u> used OERL plans in proposal preparation	76.9% (N = 93): indicated that they had <u>not</u> used OERL reports in preparation of a plan or report	78% (N = 124): indicated that they had <u>not</u> used OERL in preparation of an evaluation instrument	N/A
Use of OERL in preparation of NSF plan or report	7% (N = 14): had used OERL to prepare an NSF proposal	9% (N = 11): had used OERL to prepare an NSF plan or report	8.2% (N = 13): had used OERL to prepare instruments for an NSF project	N/A

Draft: Do not cite without permission

Impact of OERL resources on professional activities	<p>42.2% (N = 84): increased their understanding of how to organize an evaluation plan</p> <p>28.1% (N = 56): Used OERL as a refresher for design methodology</p> <p>25.6% (N = 51): improved their evaluation plans</p> <p>24.1% (N = 48): no impact</p>	<p>33.9% (N = 41): increased their understanding of evaluation</p> <p>33.1% (N = 40): presented alternative ways to display data in reports</p> <p>32.2% (N = 39): increased their understanding of how to organize an evaluation report</p> <p>28.9% (N = 35): acted as a refresher for design methodology</p> <p>28.1% (N = 34): no impact</p>	<p>44.7% (N = 71): introduced new ideas for instruments</p> <p>33.3% (N = 53): acted as refresher for aligning goals to instruments</p> <p>23.3% (N = 37): improved the quality (reliability and validity) of their instruments</p> <p>27% (N = 43): no impact</p>	<p><i>Designing Evaluation Module (N = 35)</i></p> <p>57.1% (N = 20): acted as a refresher for evaluation methodologies</p> <p>45.7% (N = 16): improved their evaluation designs</p> <p>17.1% (N = 6): no impact</p>
				<p><i>Questionnaire Module (N = 29)</i></p> <p>55.2% (N = 16): improved quality of written questionnaires</p> <p>55.2% (N = 16): helped them understand when a questionnaire should be used</p> <p>55.2% (N = 16): helped them write better items</p> <p>51.7% (N = 15): increased awareness of importance of piloting a questionnaire</p> <p>31% (N = 9): helped administer questionnaires more effectively</p> <p>20.7% (N = 6): no impact</p>

Draft: Do not cite without permission

<p>Adequacy of topic coverage</p>	<p>Evaluation Theory: 58.3% (N = 116): very adequately or adequately covered</p> <p>Quant/Qual Methods: 67.8% (N = 135): very adequately or adequately covered</p> <p>Stakeholder Involvement: 56.3% (N = 112): very adequately or adequately covered</p> <p>Research Design: 67.8% (N = 135): very adequately or adequately covered</p> <p>Use of Instruments: 66.3% (N = 132): very adequately or adequately covered</p> <p>Data Collection Procedures: 63.8% (N = 127): very adequately or adequately covered</p> <p>Timeline: 53.8% (N = 107): very adequately or adequately covered.</p>	<p>Executive Summary: 74.4% (N = 90): very adequately or adequately covered</p> <p>Project Description: 75.2% (N = 91): very adequately or adequately covered</p> <p>Stakeholder Involvement: 61.2% (N = 74): very adequately or adequately covered</p> <p>Evaluation Theory: 66.1% (N = 80): very adequately or adequately covered</p> <p>Evaluation Design: 71.1% (N = 86): very adequately or adequately covered</p> <p>Instrument Development and Use: 73.6% (N = 89): very adequately or adequately covered</p> <p>Quantitative Methods: 70.2% (N = 85): very adequately or adequately covered;</p> <p>Qualitative Methods: 68.6% (N = 83): very adequately or adequately covered</p>	<p>N/A</p>	<p><i>Designing Evaluation Module</i></p> <p>Stakeholder Involvement: 74.3% (N = 26): very adequately or adequately covered</p> <p>Evaluation Methodologies: 88.6% (N = 31): very adequately or adequately covered</p> <p>Deciding What Data to Collect: 85.7% (N = 30): very adequately or adequately covered</p> <p>Sampling Method: 80% (N = 28): very adequately or adequately covered</p> <p>Deciding what Conclusions to Draw From Evidence: 77.1% (N = 27): very adequately or adequately covered</p>
------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Draft: Do not cite without permission

<p>Adequacy of topic coverage (cont.)</p>		<p>Presenting Results: 68.6% (N = 83): very adequately or adequately covered</p>		<p><i>Questionnaire Module</i></p> <p>Determining When to Use a Questionnaire: 96.6% (N = 28): very adequately or adequately covered</p> <p>Common Errors in Question Writing: 96.6% (N = 28): very adequately or adequately covered</p> <p>Different Question Formats: 96.6% (N = 28): very adequately or adequately covered</p> <p>Structuring Questionnaires: 93.1% (N = 27): very adequately or adequately covered</p> <p>Pilot Testing: 86.2% (N = 25): very adequately or adequately covered</p> <p>Issues of Confidentiality: 82.8% (N = 24): very adequately or adequately covered</p> <p>Preparation of Data for Analysis: 86.2% (N = 25): very adequately or adequately covered</p> <p>Ways to Increase Participation: 82.8% (N = 24): very adequately or adequately covered</p>
<p>OERL improves quality of plans, reports, or instruments</p>	<p>83.4% (N = 166): strongly or moderately agree that OERL could improve their plans</p>	<p>85.6% (N = 104): strongly or moderately agree that OERL could improve their reports</p>	<p>88.1% (N = 140): strongly or moderately agree that OERL could improve their instruments</p>	<p>N/A</p>

N/A= not applicable.

Key Findings from NSF Grantee Survey

The OERL NSF Grantee Survey was sent to a random selection of DUE grant recipients who received their awards during the years 2002 – 2003 (and did not include interns). On May 26, we mailed the first mailing to all names listed in the NSF sample, (n = 510). Seventeen were returned as undeliverable, for a total sample size of 493. The first mailing was a postcard announcing that the survey packet would be sent. The first mailing of the survey packet, containing the cover letter and survey, was mailed a week after the introductory postcard. A reminder postcard was mailed about two weeks after the survey packet. A second survey packet was sent to all non-respondents. To increase the survey response rate for this survey, in the second survey packet mailing, the offer for the *NSF User Friendly Evaluation Handbook* was made to NSF Grantee survey respondents. A ‘third appeal’ postcard was then mailed about three weeks after the second survey packet.

The survey consisted of 34 items. This survey was administered as a paper-pencil survey and was also available online for those recipients who preferred to complete it electronically. (The Web address for the online version was included along with the paper version of the survey.) We received a total of 197 completed surveys. After adjusting the sample size for undeliverable names/addresses and accounting for the survey respondents who took the wrong online survey, the response rate was 42.6%.¹ After tabulating results for each of the items, we identified key findings from the survey to help us understand whether or not NSF Grantees, the original intended audience for OERL’s evaluations resources, are using and benefiting from the OERL site.

Description of NSF Grantee Survey Respondents

Principal investigators and evaluators made up the bulk of this survey’s respondents, with 56.9% being affiliated with a university. Of the 197 respondents completing this

¹ Included in the numerator of the response rate are 13 respondents who, although from the NSF sample, mistakenly responded to the OERL User Survey that was posted on the OERL Website during the same period that the NSF Grantees survey was being conducted.

survey, 83.2% were principal investigators and 6.6% were evaluators of NSF funded projects.

The majority (87.3%) of survey respondents had been principal investigators of an NSF funded project. A smaller number of respondents (48.7%) had designed and implemented an evaluation for an NSF project. Of all the respondents, 74% had no formal training in evaluation. Despite the large number lacking formal training, 40.1% described themselves as being proficient (participated in several evaluations), but only 4.6% said they were an expert evaluator. None of the respondents to this survey reported having 10 or more years of evaluation experience, making the respondents to the grantee survey the least experienced in terms of hands-on evaluation experience.

We learned that 72% of survey respondents had never visited the OERL site until they participated in the survey. Because of this, we closely examined responses related to barriers to use of a Web-based resource. In terms of Internet use, we found that 93% of NSF Grantee survey respondents access the World Wide Web about every day, 82% used search tools such as Yahoo or ERIC about every day, and 32% participated in mailing lists, discussion boards or newsgroups about every day. Given that many of the survey respondents were frequent users of the Internet and Web-based search tools, we concluded that using a Web-based evaluation resource such as OERL should not pose a challenge for them. Further strengthening this conclusion was the response of 14.2% of the respondents indicating that they already accessed evaluation resources on the Web frequently or fairly frequently. (Note that only 6.6% actually identify themselves as evaluators of NSF funded projects.)

We also learned that 80% of respondents said that OERL has resources that would be of both use and interest to them. Ninety-two percent indicated they might or would definitely use OERL again. Further analyses revealed that those who took the OERL tour (Web pages devoted to explaining OERL) said, “Yes” (versus “Maybe”) they would use OERL again more frequently than respondents who did not take the tour (68% versus

55%). These findings indicated that once visitors became more familiar with OERL, they responded with more interest in using it again ($\chi^2(2, N = 190) = 16.29, p = .000$).²

- The survey respondents identified several ways they had heard about OERL. The most frequent ways are listed below
 1. From NSF Program Officers (33.2%)
 2. From e-mail inviting the individual to participate in this survey (24.4%)
 3. From a brochure or print mailing (18.5%)
 4. From a professional conference (5.0%)
 5. From a colleague (4.6%).
- Of the 48 survey respondents who had previously visited OERL,
 1. Ten had used OERL in the preparation of a proposal
 2. Five had used OERL in the preparation of a report
 3. Fifteen had used OERL in the preparation of an evaluation instrument.
- Three percent of survey respondents had submitted evaluation resources to be reviewed for potential inclusion on the OERL site.
- The top four reasons survey respondents named for using OERL again were:
 1. To find an example of an evaluation plan or report (71%)
 2. To get a general sense of the types of instruments used in an evaluation (67.5%)
 3. To find a specific instrument to adapt to their needs (66%)
 4. To learn about a specific topic in evaluation (47.7%).
- Survey respondents indicated some of the potential impacts that OERL could have on their projects:
 1. Improve their understanding of evaluation activities in general (76%)

² Though we found a significant difference on the variable of “Do you plan to use OERL again” when cross-tabulated with how the respondent took the survey (online versus paper-pencil) in the initial tests (p. 18) that were conducted to determine if survey taking mode caused systematic differences, we do not believe that it influenced the finding here. Further tests (we did a cross-tabulation between “Did you take the tour” and “Would you use OERL in the future” and added in the variable for “how did you take the survey” as a control variable) revealed that for both online and paper-pencil versions significantly more respondents said “Yes” to the question of “Would you use OERL in the future” if they took the tour of OERL. For the online survey, the results are ($\chi^2(2, N = 80) = 9.4, p = .009$) and for the paper-pencil survey, ($\chi^2(2, N = 110) = 6.5, p = .04$). Therefore, we do not believe that the mode in which the respondent took the survey caused any systematic differences in response rates for future OERL use.

2. Improve the way they plan future projects (74%)
3. Improve communications between the evaluative and substantive components of the project (64%).

Even though NSF Grantees were not familiar with OERL before learning of it through the survey, they reported that OERL could potentially help their projects and their responses to OERL were generally positive. In addition, those who took the tour were more likely to say “Yes” than “Maybe” to the question of whether they planned to use OERL in the future. It seems that as people become more familiar with OERL, they are more inclined to use it again.

Key Findings from the OERL Education Evaluator Survey

The sampling frame for the Education Evaluator Survey consisted of members of the American Education Research Association (AERA) who either currently or previously had participated in the evaluation division of the organization and members of the American Evaluation Association (AEA) who had participated in AEA's educational evaluation topical interest group. Three samples were randomly drawn from the combined lists, consisting of $n = 589$, $n = 106$, and $n = 798$ (respectively), for a total of 1493 potential respondents. After omitting those names found to be non-deliverable, the total sample size was 1335.

Data collection took place in three waves (a wave per sample) between May and September 2004. For each set of names, the survey was announced with an introductory postcard, mailed in advance of the actual survey packet. The first mailing of the packet, containing the cover letter and survey, was mailed a week after the introductory postcard. A reminder postcard was mailed about two weeks after the survey packet. A second survey packet was sent to all non-respondents, with an offer of the NSF *User Friendly Evaluation Handbook* as an incentive to increase survey responses, and lastly, a 'third appeal' postcard was mailed about three weeks after the second survey packet.

Respondents had the choice of responding to either the paper-pencil version of the survey received via U.S. mail, or to an online version of the same instrument, posted on a commercial survey distribution Web site. (The Web site address for the online version was included along with the paper version of the survey.) We received a total of 433 completed surveys; after adjustment for non-deliverable addresses (above), this constitutes a response rate of 36.25 percent.³

Description of Education Evaluator Survey Respondents

A primary affiliation with a university or college was reported by about 47% of the respondents. Fifteen percent were primarily affiliated with a nonprofit organization.

³ Included in the numerator of the response rate are 51 respondents who, although from the Education Evaluator sample, mistakenly responded to the OERL User Survey that was posted on the OERL website during the same period that the Education Evaluator survey was being conducted.

Nine percent were self-employed evaluators; 9% held positions with K-12 education organizations or school districts, with private companies (7%), or government agencies (3.5%). The remaining respondents listed other professions (3.5%) or were full-time students (3.5%).

The sample of respondents almost unanimously reported being engaged in some form of program evaluation relating to education. Only about 3 – 4% (depending on the survey item) indicated that they were not an evaluator or had no evaluation experience. Most respondents (82%) had received formal training in evaluation, and about 52% had more than 10 years of evaluation experience. Thus, of the three samples investigated, this respondent group contained those with the most hands-on experience with education evaluation.

OERL Use

The OERL survey served as an introduction to OERL for many of the respondents; and, parallel to the responses from NSF grantees, many first-time users indicated that they would be likely use the resource in the future. About 56% of the respondents indicated that they had not visited OERL before they had received the survey appeal. Close to 13% on the other hand, had used OERL previously and continued to make frequent or occasional use of its resources. When all respondents, regardless of their experience with OERL, were asked if they intended to make use of OERL in the future, 62% indicated “Yes” they would, with an additional 30% indicating “Maybe.” From this survey, we had hoped to gain insight as to why it seems we have so many one-time users of our Web site. Unfortunately, we did not find an answer as most respondents (over 90%) to this survey (and the other two surveys) indicated that they might or would use OERL again.

To become familiar with OERL resources, respondents either took the online tour (57% did so) or explored the layers of the Web site on their own. About 15% spent between 16 and 30 minutes investigating the Web site; about 10% spent more than 30 minutes. The remainder spent 15 minutes or less before responding to the survey.

Among all respondents, the following findings emerged:

- The most important evaluation topics among the respondents were:
 1. Evaluation design (very important to 64% of respondents),
 2. Quantitative methods (very important to 58% of respondents, and
 3. Statistical methods (very important to 50% of respondents).
- Relative to the three topic areas listed above, qualitative design, stakeholder involvement, and evaluation theory were less important.
- Among the same set of six evaluation topics, those judged to be most adequately addressed by OERL content were:
 1. Evaluation design (50% thought coverage adequate or better),
 2. Quantitative methods (41% thought coverage adequate or better),
 3. Stakeholder involvement (41% thought coverage adequate or better), and
 4. Qualitative methods (40% thought coverage adequate or better).

(Note that 42-50% of the respondents marked “no opinion” about how these topics were covered on OERL, due in part to their limited exposure to the Web site to this point.)

- The reasons most commonly given for visiting OERL in the future were
 1. to find a specific instrument (76% would visit for this reason),
 2. to find an example of an evaluation plan or report (70%), and
 3. to get a general sense of instruments (56%).
- Overall, respondents thought that OERL would be most useful to those with relatively less evaluation experience. In a ‘check all that apply’ context, 69% thought OERL would be useful to inexperienced evaluators, 61% thought it useful to those who are not evaluators but are conducting evaluations, and 63% thought it useful to proficient evaluators. A smaller percentage, 29% thought the Website would be useful to expert evaluators.

Bivariate Findings from the Education Evaluator Survey

Bivariate analyses (cross-tabulations and comparisons of means) were conducted to see if intent to use OERL in the future varied by type of evaluator, level of experience, or

by other characteristics, such as familiarity with the resources available in the various sections of the Web site.

While the most respondents in the Education Evaluator sample experienced their first exposure to OERL in the context of responding to the survey, there was variation in the amount of time that respondents spent exploring the Web site before responding to the survey itself. Among those who had taken the tour, 71% indicated that they would use OERL in the future, compared to 64% of those who had not taken the tour ($\chi^2(3, 411) = 13.72, p < .001$). In addition, among those who'd spent more than 30 minutes exploring OERL, 85% said that they would visit the site in the future.

Of seven reasons listed to visit OERL in the future the mean number of reasons indicated was 3.60 (standard deviation 1.91). This mean number of reasons did not vary by level of experience of an evaluator (inexperienced, proficient, expert, not an evaluator), by methodological orientation (mostly quantitative, mostly qualitative, mixture of both), or by types of evaluations typically conducted by the respondent (e.g., for K-12 programs, higher education).

We tested a set of research questions that examined whether the rating of OERL's resources in terms of interest and usefulness varied by type of evaluator, level of experience, etc. Given the overall lack of variation across the five-point Likert-scale items (with most respondents agreeing that the Web site was of interest and useful), we did not detect differences by group.

We also tested whether the coverage of six evaluation topics (evaluation theory, statistical methods, stakeholder involvement, qualitative methods, quantitative methods, and evaluation design) was rated differently by various types of evaluators. There were no significant differences in the ratings of the coverage of quantitative methods or statistical methods based on methodological orientation (quantitative, qualitative, mixture of both), nor in the rating of the coverage of qualitative methods.

Twenty-one percent of the Education Evaluator respondents had designed and/or implemented at least one evaluation for a National Science Foundation project in the past.

Of these, the largest proportion (60%) conducted evaluations for higher education organizations. Those who had conducted evaluations for NSF projects were significantly more likely than others to say that they would use OERL in the future to find a specific instrument ($\chi^2(1, 415) = 2.91, p < .056$). On the other hand, they were less likely to say that they would use OERL to get a general sense of instruments ($\chi^2(1, 415) = 4.383, p < .03$) or to find an example of an evaluation plan or report ($\chi^2(1, 415) = 6.977, p < .01$).

Lastly, an analysis was conducted to determine if evaluators with 0 to 4 years of evaluation experience would make use of the features of OERL in the same proportions as evaluators with 5 or more years of experience. These findings are presented in Table 7 on the following page and suggest that OERL resources have different uses for evaluators at different career stages. For example, the resources relating to statistical analysis seem to be most useful to evaluators in the earlier stages of their careers. While it seems that OERL resources can be seen by more experienced evaluators as a teaching tool or as networking tool.

Table 7. Reasons to Use OERL in the Future, By Years of Evaluation Experience

Reason to use OERL in the future		All	Years of Evaluation Experience		Chi-square	df	P-value
			0 – 4 years	5 or more years			
To find a specific instrument	%	78.1%	74.0%	79.4%	1.32	1	.269
To get a general sense of instruments	%	57.2%	63.0%	55.5%	1.78	1	.204
To learn about a specific topic in evaluation	%	53.4%	60.0%	51.4	2.27	1	.137
To get information to teach others about evaluation	%	40.9%	32%	43.6%	4.256	1	.047*
To connect with another evaluator	%	23.0%	15.0%	25.5%	4.782	1	.030*
To find an example of an evaluation plan or report	%	72.4%	78.0%	70.7%	2.026	1	.161
To get information on how to analyze data	%	39.2%	51.0%	35.5%	7.67	1	.007*

Total n = 421

0 – 4 years of experience, n = 100

5 or more years of experience, n = 321

* Statistically significant difference between those with 0 – 4 years of experience and 5 or more years of experience

References

- Bauer, K. (2000). Who goes there? Measuring library Web site usage. *Online Incorporated*, 25–31.
- Bowker, D. (1999). Constructing the client-computer interface: Guidelines for design and implementation of Web-based surveys. *Summary Report 99-15*, Social and Economic Sciences Research Center. Washington State University, Pullman, WA.
- Cook, C., Heath, F., & Thompson, R.L. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. *Educational and Psychological Measurement*, 60 (6), 821–836.
- Dillman, D.A., & Bowker, D.K. (2001). The Web questionnaire challenge to survey methodologists. In U. Reips & M. Bosnjak (Eds.), *Dimensions of Internet Science*. Lengerick, Germany: Pabst Science Publishers.
- Dillman, D. A., Tortora, R.D., Conrad, J. & Bowker, D. (1998, August). *Influence of plain vs. fancy design on response rates for Web surveys*. Presented at Joint Statistical Meetings, Dallas, TX.
- Kittleson, M. (1997). Determining effective follow-up of e-mail surveys. *American Journal of Health Behavior*, 21, 193–196.
- Shannon, D.M., Johnson, T.E., Searcy, S., & Lott, A. (2002). Using electronic surveys: Advice from survey professionals. *Practical Assessment, Research & Evaluation*, 8, (1). Retrieved October 7, 2003 from <http://edresearch.org/pare/getvn.asp?v=8&N=1>
- Solomon, D.J. (2001). Conducting Web-based surveys. *Practical Assessment, Research & Evaluation*, 7 (19). Retrieved October 7, 2003, from <http://ericae.net/pare/getvn.asp?v=7&n=19>
- Vehovar, V., Zenel, B., Lozart, K., & Zaletel, M. (2002). Nonrespondents in Web surveys. In R.M. Groves, D.A. Dillman, J. Eltinge, & R. Little, *Survey Nonresponse* (pp. 229–243). New York: Wiley-Interscience.
- Zalles, D. (2002). *Web-based performance support for evaluation*. Presented at the Educational Technology Conference of the Society for Applied Learning Technologies.