

2005 NSDL Evaluation Practices Survey: Summary of Results & Recommendations

Sarah Giersch, Jim Dorward and Mimi Recker

Introduction

The NSDL Education Impact and Evaluation Standing Committee's (EIESC) charter contains the following objective: *to promote a culture of user-centered design through sharing instruments and expertise and through capacity building*. This idea has been operationalized through EIESC activities (e.g., Evaluation Practices Survey, 2002; Education Impact Workshop, 2003; Webmetrics Workshop, 2004) and by workshops organized by NSDL projects and other standing committees (Participant Interaction in Digital Libraries, 2004). These activities have been motivated by committee chairs and project directors recognizing the necessity and value of enabling NSDL projects to describe their efforts and results quantitatively and qualitatively in order to improve the design, development and use of educational digital libraries. Building the capacity to conduct useful evaluations becomes even more important as NSDL (and NSDL-funded projects) seek to justify their reach and impact to NSF and to other funding agencies.

When the EIESC met in June during JCDL 2005, attendees agreed that not enough was known about the current state of NSDL projects' evaluation activities and needs to provide direction for upcoming EIESC activities. A taskforce (Sarah Giersch, Jim Dorward and Mimi Recker) agreed to conduct a survey regarding NSDL projects' evaluation capacity before the NSDL Annual Meeting in November 2005. This report summarizes the results from the 2005 NSDL Evaluation Practices Survey and offers recommendations to the EIESC, NSDL Core Integration and NSF on how to meet the evaluation needs of NSDL projects going forward. The EIESC website contains the 2002 and 2005 surveys and results (<http://eduimpact.com.nsdlib.org>).

Methodology

The Evaluation Practices Survey (EPS) taskforce decided to re-use the survey instrument developed and implemented in 2002 (see EIESC website) because this would allow the taskforce to compare 2005 results with 2002 baseline data. A detailed comparison of data is not within the scope of this document. However, major differences are noted in the context of specific 2005 results. In the interest of time, the taskforce decided to distribute the survey online (via Web Online Surveys, <http://web-online-surveys.com>). In 2002, the survey was distributed on paper at the NSDL Annual Meeting, resulting in a very high response rate but a delay of several months before analysis and reports were available for action. The purposes of the 2005 NSDL Evaluation Practices Survey were

- 1) To determine the current evaluation practices and planned activities used in NSDL projects
- 2) To determine the needs of the NSDL community regarding evaluation tools, services, and related activities

The target audience for the 2005 NSDL Evaluation Practices Survey was NSDL Principal Investigators (PI). The audience was notified of the survey by email via the

NSDL Voting listserv, which contains addresses for every PI that has received NSDL funding from 2000 to the present. The survey was available online from September 21 – October 14, 2005, and one follow-up reminder was sent. One result per project was requested; there were no duplicate results. No technical problems were reported with the survey form.

Summary

There are 158 unique email addresses on the NSDL Voting listserv; 38 survey responses were received, for a response rate of 24%. Many of the questions had “check all that apply” instructions, so the number of responses is often greater than the survey response rate.

Although there is no longer a Collections Track in NSF’s NSDL program, 36% of respondents self-identified their projects as Collections, followed closely by Services, 31%. And while the Pathways Track is relatively new (introduced in FY 2004), seven of the nine Pathways grants responded. Four respondents indicated that they had received funding under more than one track.

Evaluation Goals

It is difficult to know how to interpret projects’ evaluation goals since respondents were given the option to “check all that apply” and clear definitions for each option were not included. Process-oriented goals (see Table 1) received the majority of votes. The next three categories (Educational Impact, Usability / Accessibility, and Usage Tracking), do not rank as highly but did receive 10 or more votes.

Table 1: Top Four Evaluation Goals

<i>Options</i>	<i>Number of responses</i>	<i>Response ratio</i>
Process-Oriented (e.g., community building, educational outreach, collection building processes)	21	30%
Educational Impact	14	20%
Usability and/or Accessibility (e.g., per ADA or W3C guidelines)	12	17%
Usage tracking (via web metrics, profiles, etc.)	10	14%

The details respondents provided about their evaluation goals only somewhat match their responses (see Table 1) as opposed to the 2002 EPS, where respondents selected goals from a list comparable to Table 1 which closely matched their stated goals. One possibility is that the options presented, while relevant in 2002, may not reflect the current state of NSDL projects’ evaluation activities in 2005. Respondents self-described their project goals in varying levels of detail, from highly-specific objectives and key questions to general statements (e.g., “Justify my existence”). Goals for Collections projects included aligning content retrieval tools (e.g., website, search interface) and content to users’ needs. Services goals reflect the formative process by which projects are making tools for digital libraries (see Figure 1 for sample comments). Reflecting their “vertical market” approach, Pathways projects’ goals were a combination of Collections and Services goals (e.g., aligning content and building relevant tools).

Figure 1: Project Evaluation Goals: Sample Comments

- 1) Collections: Is the site useful and used by the community for which it was designed?
- 2) Services: [My projects'] two goals are 1) evaluate our tools to make sure they work like they are supposed to, and inform what improvements to make 2) evaluate the usability of the tool with actual users to make sure the tool does what they would like it to do.

Some of the evaluation goals reflected topic-specific objectives (e.g., tool building) *and* broader impact objectives. While this is most likely a product of NSF solicitation requirements, it is interesting to review the "bigger picture" goals of respondents (see Figure 2). The most common measures of impact that NSDL projects mention using are changes in knowledge and practice among teachers and students. Aside from the Education Impact workshop (2003), no formal work has been undertaken since then to identify work across NSDL to measure impact. It might be useful to learn more about individual projects' plans and progress on these broader goals for the purpose of informing EIESC, Core integration and NSF on the progress of

Figure 2: NSDL Projects' Impact-related Goals

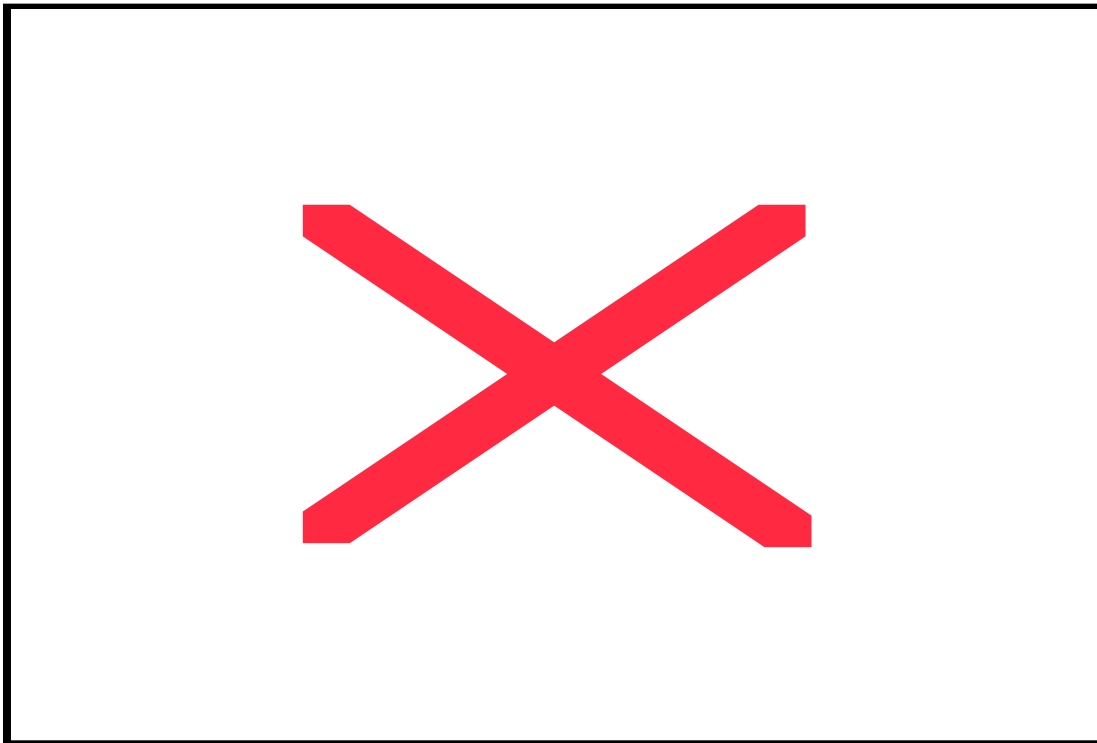
- 1) Evaluate... the effectiveness of SIMPLE science in helping students achieve specific educational standards.
- 2) Evaluate impact in 1) promoting stronger connections between research and education, 2) integration into the community and related disciplines, 3) positively influencing learning
- 3) [Evaluate] effectiveness in terms of the enhancement of teacher knowledge of the specific content and concepts and student achievement in selected science concepts.
- 4) K-12 teachers' attitudes, content knowledge and skills will be assessed to measure the effectiveness of the Digital Library.
- 5) Impact of NSDL through incorporation into traditional academic library services.
- 6) Study of collaborative learning in a digital library.
- 7) To learn whether we can build a community of practice around the study of a specific scientific field, which includes both professional and lay participants
- 8) Whether the educational interventions that designed in the software may (or may not) support students with their investigative practices.

measuring NSDL's education impact.

Evaluation Resources Used

Regarding the number of project FTE's devoted to evaluation, 76% of projects indicated that evaluation receives half-time or less of FTE attention. Collections, Services, Pathways and respondents with more than one project comprise the majority of that number (see Graph 1). Targeted Research projects devote more FTE's to evaluation. 62% of projects indicated that they use, or intend to use, an independent evaluator, though this number could not be linked with any significance to the type of project or FTE's devoted to evaluation.

Graph 1: Approximate Number of NSDL Projects' FTEs Devoted to Evaluation (by funding track)



36% of respondents indicated that they utilized in-office evaluation expertise, and 22% indicated they consulted with a colleague or friend. Via the "other" option, 6 respondents indicated they used an independent evaluator, confirming responses to question #5. 42% of the respondents used a combination of the resources listed (in-office expertise, colleague, web, other campus department).

Evaluation Resource Needs

Respondents preferred for NSDL to provide evaluation information first via the web (46%), then via the Annual Meeting (28%) and workshops (26%). However, 84% of respondents indicated they would attend a session or workshop on evaluation at the NSDL Annual Meeting. Table 2 provides a ranked list of workshop topics, where the high numbers indicate a high level of interest. In addition, respondents overwhelmingly expressed interest in knowing about other projects' evaluation strategies, tools and results through various means (e.g., NSF-wide directory; central clearinghouse). Reflecting projects' evolution towards maturity, there were also several requests for assistance in finding suitable groups of users for usability and other types of testing.

Table 2: Interest level in Workshop Topics

<i>Options</i>	<i>Number of responses</i>	<i>Response ratio</i>
System and web log analysis	24	13%
Usability testing	24	12%
Survey data	18	9%
Data analysis resources	18	9%
Specific information about designing and using different data collection methods	15	7%
Heuristic evaluation	14	7%
Tools for designing evaluation plans (templates, checklists, examples, budgets)	14	7%
Focus groups	13	6%
Performance assessment	12	6%
Observation of learners or participants	11	5%
Student/user journals or logs	11	5%
Interviews	10	5%
Classroom observation	10	5%
Student work samples	6	3%
Other	2	1%
1. Results of other project evaluations		
2. Human subjects research information		

This concludes the summary of the 2005 Evaluation Practices Survey. The full results can be found on the EIESC website (<http://eduimpact.comm.nsdlib.org>).

Conclusion & Recommendations

Since the response rate (24%) for the 2005 Evaluation Practices Survey is low, it is difficult to predict whether the survey results are representative of all NSDL projects. From a quick review and comparison of 2002 and 2005 survey results, it can be inferred that evaluation needs indicated on the 2005 EPS reflect the maturation of NSDL projects. Two indicators of maturity are: projects need specific methodologies and tools to guide their evaluations (rather than general evaluation information) and projects need specific user groups with which to conduct evaluations (implying that initial prototyping is complete and projects are ready to expand and refine their content, tools and services for broader audiences). The lack of access to user groups may depend on the projects' target audience (e.g., middle level science teachers vs. undergraduate earth science faculty).

- 1) *Recommendation:* The EIESC should extend the Evaluation Exchange program (only conducted at face-to-face meetings) to include a moderated, online exchange or forum dedicated to posting evaluation methodologies, tools, results and questions.
- 2) *Recommendation:* The EIESC, in conjunction with NSDL Core Integration and NSF, should convene a taskforce to examine strategies for reaching various user groups in order to facilitate conducting user-centered evaluations (e.g., a brokering service or leveraging the audiences of other NSF programs).

While some of the topics in Table 2 (Interest in Workshop Topics) are related to formative evaluations for content and tool development (e.g., usability testing, heuristic evaluation), other topics are related to evaluating education impact (e.g., student journals, participant observation). However, topics about measures of impact were of less interest to respondents. This could be a result of overall project goals or different stages of project development. It was not within the scope of this survey to identify projects' activities to measure impact. Perhaps such an activity should be undertaken, though, given the emphasis within NSF to demonstrate broader impact and the emphasis during NSDL's development on demonstrating education impact.

- 3) *Recommendation:* EIESC and Core Integration, with input from NSF, should identify, for project tracks and/or types, evaluation questions designed to elicit impact information and should identify methodologies and tools to support NSDL projects' evaluation activities.