

# Archived Information

## Interim Evaluation of the Northwest Regional Educational Laboratory

### **I. Brief Overview of Laboratory Review**

The Northwest Regional Educational Laboratory (NWREL) evaluation took place at NWREL's Portland, Oregon office April 5-9, 1999. I reviewed all the materials sent to panelists in advance and materials given to panelists on site. During the site visit, I met with several NWREL staff and clients.

### **II. Implementation and Management**

#### **A. To what extent is the REL doing what they were approved to do during their first three contract years?**

##### **1. Strengths**

It is useful to note NWREL's assertion in the Technical Proposal dated 8/8/95 that its plan covering the five-year period under consideration will focus on "supporting broad-based, comprehensive educational change to ensure excellence and equity for all students to attain high academic achievement." The Laboratory agreed to conduct research and development, dissemination, training, and technical assistance that results in tools and strategies to enable local and state practitioners and policy makers to better carry out a systemic approach to reform and to "scale up" reform efforts to encompass all schools and communities. Three priority areas were given: assessment and accountability, rural education, and school change. The list of tasks and activities is much too long to consider in this report (e.g., the listing for FY98 goes on for six pages).

As I will elaborate on later in this report, it was difficult to separate Laboratory activities conducted under the OERI contract from other Laboratory activities. The following data cover efforts from the entire Laboratory: In FY98 NWREL served 940 different agencies and organizations and 244 school districts. Participating in workshops, conferences, and consultations were about 21,000 individuals. There were no major concerns expressed in the ratings of the

Laboratory's performance turned in by OERI (apparently the ratings were a combination of ratings by the contracts-and-grants people and the program officer). According to NWREL, the Laboratory has always prided itself in submitting deliverables on schedule and has never had "major exceptions in audits."

Guiding the overall management of NWREL as an institution including the OERI portion of their contract is the *Northwest Regional Educational Laboratory Strategic Plan*, adopted by the Board of Directors in June 1996. Alfred Rasp, an independent reviewer, praised the quality of the plan. I concur with that praise and found the listing of strengths fully balanced by a listing of weaknesses for each of the seven major NWREL sub-systems. The subsequent *Strategic Plan Status Report*, dated March 26, 1999 (but received by the panel before that date), reported on accomplishments for each of the previous three fiscal years for each of the strategic initiatives (in the 1996 plan, the term "strategic initiative" was not used). The status report, however, did not report on the NWREL options for responding to twelve trends listed in the strategic plan. When the panel asked the Laboratory why the trends were not further delineated in terms of strengths and weaknesses, NWREL informed the panel that, based on advice from the team of Rasp, Popham, and Frierson, they have moved away from trends, which tend to be broadly based as well as frequently shifting, toward a more specific approach that should be more useful at the classroom level. However, the Laboratory's 1997 regional needs assessment noted that "The trends were again reviewed and refined in 1997."

The Laboratory leadership regularly conducts meetings to ensure timely progress as well as communication across programs.

## **2. Areas of needed improvement**

Given the emphasis on improving student achievement, NWREL needs to show progress in this area.

According to the Statement of Work that Laboratory proposals were to address, "Each Laboratory must also ensure that 25 percent of its resources are used to serve rural areas." I did

not see any written assertion in the documents that this requirement was met; however, we do know that the percentage of area that is rural for NWREL is the largest among the various RELs.

### **3. Recommendations for improvement**

If the thinking is such that it may take several years before effects for a particular intervention can be expected (e.g., the panel was told that OTE II effects were not necessarily expected to show up until as much as five years after initiation), such thinking should be presented in writing. For sure, such concerns should be expressed up front to potential adopters.

#### **B. To what extent is the REL using a self-monitoring process to plan and adapt activities in response to feedback and customer needs?**

##### **1. Strengths**

Much effort has been made to collect feedback data from clients. NWREL's institutional evaluation reports (conducted internally) include the Laboratory's efforts to address needs uncovered in various evaluation studies (e.g., when feedback from board members indicated some concern in areas such as understanding NWREL's mission and selecting and orienting new Board members, the Executive Committee of the Board developed an action plan for improvement). Based on the regional needs assessment, a number of specific implications for NWREL's work were listed. According to NWREL, the needs assessment is used in conjunction with advisory committee guidance, internal and external evaluations, and close attention to constituent satisfaction to shape the direction of NWREL.

Biennial regional education needs assessments are used to get information on customers' needs. The most recent one available to the panel was the 1997 needs assessment. Also available for the needs assessment was a household survey conducted by Gilmore Research under contract to NWREL. Further informing the needs assessments were a series of seven state forums on educational needs in which 245 individuals participated in 1997.

A rigorous evaluation was conducted internally of the Onward to Excellence (OTE) Program in 33 schools in Mississippi. The findings have been reported as "good news, bad news." The bad news is that implementation was uneven. The good news, supposedly, is that

success was seen in those schools doing a good job of implementation (see the next section for concerns about this rendition of the findings).

## **2. Areas of needed improvement**

With a modest amount of additional effort, the methodology used to get feedback from users could be improved. The several self-monitoring efforts could be improved by (a) increasing the response rate (where the response rates were low, there was no indication of any additional effort to increase the rate); (b) ensuring representative samples as opposed to samples where clients are more likely to be favorable (collecting data only from clients requesting two or more products leaves out clients who may have been less satisfied and therefore did not request another product); and (c) including the collection of anonymous data, especially where that might produce more forthright answers.

NWREL quarterly sends out the Professional Activities Reporting System (PARS) Follow-up Survey to all clients who have contracted with NWREL for services or requested information or needs clarification services. According to the 1998 Annual Report of Institutional Performance, 367 PARS surveys were sent out with a return rate of 40%. The number of respondents providing feedback in the products and publications category was only 9. Although the mean ratings were high, the low sample size detracts from any confidence one might have in those figures. The fact that the survey is not anonymous may also detract from the validity of the feedback.

The response rates on the written regional needs assessment survey sent to educators in April 1997 ranged from a low of 14% of school board chairs in Idaho to a high of 40% of superintendents in Alaska. The overall response rate of 26% resulted in what NWREL called a sample "representative of each state and the region as a whole." It was not clear as to how NWREL came to that conclusion. I would suggest that low response rates such as 17% of the school board chairs responding may indicate the possibility of a non-representative sample.

In *Overall Laboratory Operation, Volume II* (prepared by NWREL for the panel), the results on the 1998 PARS include open-ended responses. There is a category called

“Compliments” but none called something like “Concerns,” even though some suggested changes are presented under other categories. At first glance readers may get the inaccurate impression that there were no negative comments.

The reporting of findings from the OTE evaluation done in Mississippi at first gave me the impression that good implementation of OTE clearly implied good results. It turns out that normal curve equivalent gains of about 2 to 7 were shown for the 11 sites that had good implementation with only the four highest implementers having gains closer to 7. Of several comparisons between OTE schools and comparison schools, all were not statistically significant (but the effect size was as large as .5), but four were negative (the comparison group’s means were higher than the means of the OTE group), one was positive, and one was neutral. The executive summary of the report has a more accurate rendition than the one we heard orally in group discussions with Laboratory staff: “Overall, student achievement showed a pattern of no change over time in OTE schools, and OTE schools did not outperform non-OTE schools while controlling for poverty and prior achievement. However, the results also demonstrated that OTE can result in improved student learning if the process is well implemented. The four highest-implementation schools showed somewhat more promising results than did the full sample of 33 OTE schools.”

In the vast amount of material we received, I was often not able to find explicit references to changes made in response to the feedback obtained on a number of formative evaluation efforts. For example, the Rasp report (6/97) to the Executive Cabinet has several explicit recommendations, such as shortening the needs assessment report and not having the data so dominated by professional educators or lay people close to education. The executive council was encouraged to define and establish desired performance measures and evaluation standards and to encourage staff to move toward them. Laboratory staff informed us orally of an integration of two programs that they felt addressed the recommendations. It was not clear if such an integration met the recommendations, and furthermore, such an assertion about how those recommendations were met should have been available in writing.

The 1996 Strategic Plan listed strengths and weaknesses under each of the major NWREL sub-systems (governance, policy planning and management, human resources, institutional development, asset management, interagency relationship, and research and development), but I could not find in the 1999 update any reference to addressing the aforementioned weaknesses. Again the Laboratory response was that the integration of two programs addressed those concerns. During the site visit, the Laboratory reported it had just done a mapping of recommendations related to the Strategic Plan and found that the Laboratory had addressed all the areas of concern.

### **3. Recommendations for improvement**

Use more rigorous evaluation designs, striving for unbiased, representative samples. Follow up with non-responders to try to increase return rates. This could be as simple as sending out another questionnaire to non-responders (Note: NWREL informed us that in a number of instances, they did use follow-up reminders but were still unable to obtain good return rates). Consider using incentives for responders.

In many cases I would expect the results to continue to be strongly positive. With a better response rate, those positive findings could then be used with a high level of confidence.

## **III. Quality**

### **To what extent is the REL developing high-quality products and services?**

#### **1. Strengths**

NWREL makes good use of comprehensive literature reviews, expert panels, and field-based experiences to guide the development of their products and programs. They are also assisting their clients in getting into thinking about education from a research base. For the 6-trait assessment model, a group of selected (by the Laboratory) teachers informed the panel that having a common (assessment) language has been most valuable to their classroom teaching. The enthusiasm of the group of seven educators who are using 6-trait assessment was impressive. NWREL is regarded as a good source of the latest research on what works.

Much of the feedback from users has been highly positive (e.g., for the teacher's guide to technology). In addition there have been numerous external awards such as the Golden Lamp Award for 1997 from EdPress and an award of excellence from the National School Public Relations Association for the Northwest Education magazine and the designation of OTE as one of ten "exemplary" programs produced by the regional educational Laboratories.

Dramatic changes in schools using OTE have been reported in areas such as Snoqualmie Valley, Washington and Bruce, Mississippi. Enthusiastic, convincing testimonies were presented to the panel by teachers and other OTE users. In one case, an educator credited OTE for helping 100% of the schools in his area become accredited. It should be noted that a formal formative evaluation of OTE II is not due until 5/31/99, and a summative evaluation is not due until 11/00.

## **2. Areas of needed improvement**

An institutional review board (IRB) at the Laboratory was not mentioned in the material we received; however, I was informed orally that the Laboratory does have some sort of review process in place, but in general, the Laboratory feels that the kind of research they do involving questionnaires does not need an IRB type of approach. It is my understanding that any student surveys or use of student school records conducted using USDE monies require written parental consent (Family Educational Rights and Privacy Act [FERPA] and Protection of Pupil Rights Amendment [PRPA]). For adults, informed consent is often not required for questionnaire-based research, but sometimes it is required, and NWREL also conducts research involving observations, so privacy concerns may arise.

There is a general feeling among users that programs such as 6-trait assessment of reading and OTE II are very effective; however, most of the evidence has not (yet) been based on rigorous studies.

## **3. Recommendations for improvement**

Unless there is a compelling reason against it, I recommend that NWREL put dates on all important publications. The lack of dates sometimes made it difficult for the panel to fully understand how particular documents fit into the larger picture. Users other than the panel also

have a right to know how recently documents were created; furthermore, dating seems to be a standard followed by virtually all commercial and academic publishers. Examples of undated documents include *6 plus one! TRAITS*, *Services to Task 2/3 Partner Sites*, and *Research You can Use to Improve Results* (prototype).

Another suggestion relates to NWREL's efforts to validate their 6-trait reading assessment model by examining the agreement between 6-trait testing results and commercial norm-referenced tests such as CTBS and ITBS. If indeed the 6-trait method is adding something substantial to the assessment procedure, then one would expect (even hope) to see some notable disagreement between 6-trait and ITBS results, for example.

#### **IV. Utility**

##### **A. To what extent are the products and services provided by the Laboratory useful to and used by customers?**

###### **1. Strengths**

There have been several users who have reported that they regard specific Laboratory products very highly; for example, a panel of teachers and other educators using the 6-trait reading assessment model (and the 6-trait writing assessment model), an evaluator in a Washington State school district, and users submitting various ratings on questionnaires such as PARS. High percentages of respondents gave "good" or "excellent" ratings to the various NWREL products and services. Almost two thirds reported making some work-related changes as a result of NWREL's services. More than a third said they noticed changes in student achievement, behavior, or attitudes. Examples of highly positive comments made to me directly about NWREL: "...6-trait assessment has made more difference in our district than any other initiative since I arrived." "Their website for writing is excellent and the printed materials are in wide use here." "They (NWREL) are cheap if not free, and what's more, they are competent and easy to deal with. A truly rare combination."

The numbers are sometimes impressive; for example, 100,000 copies of *Effective School*

*Practice* have been distributed. A GAO study found that OTE was the school-change resource most familiar to schools.

By trying to be current and poised to help schools on short notice, the Laboratory has sometimes been able to serve as a resource much needed by schools under pressure to respond to mandates and initiatives (e.g., moving to standards and standards-based assessment, improving reading, and comprehensive school reform).

Work is available through a variety of modes such as print, web sites, videos, and in-person workshops.

## **2. Areas of needed improvement**

The evaluator who highly praised the 6-trait assessment model had not heard of OTE. In at least one situation, the Laboratory has changed from sending out a catalog to responding only to requests. In order for good products to be used (and therefore to be useful), they have to first get into the hands of potential users.

## **3. Recommendations for improvement**

The Laboratory should examine its revised practice of responding to requests rather than aggressively and systematically informing potential users about the availability of its products. Perhaps a research study could help the Laboratory determine how to maximize dissemination (and therefore, use) of its products.

### **B. To what extent is the REL focused on customer needs?**

#### **1. Strengths**

NWREL regularly collects feedback from customers; for example, via PARS. Most of the feedback has been highly positive. Several individuals reported that the Laboratory really met their needs.

By sponsoring many workshops and conferences throughout the country, NWREL has provided the level of support needed in order for major school changes to take place. By conducting regular needs assessments, NWREL remains informed of the type of support schools need and want.

The Laboratory develops and pilot tests its products in real school settings (e.g., the 6-trait assessment of reading model), thus enhancing the likelihood that the products will be useful to teachers and other educators.

## **2. Areas of needed improvement**

Needs assessments seem to consist almost exclusively of responses to questionnaires from clients or potential clients.

## **3. Recommendations for improvement**

Consider conducting more comprehensive needs assessments that are based on more than user feedback.

# **V. Outcomes and Impact**

## **A. To what extent is the REL's work contributing to improved student success, particularly in intensive implementation sites?**

### **1. Strengths**

There are some case studies that show substantial student success, largely credited to Laboratory assistance (e.g., Bruce (Mississippi) schools, Snoqualie Valley School District, Glenfair Elementary School, and Fall City Elementary School). Some of the studies include data collected over several years.

The Laboratory has published numerous articles in national journals and has made several presentations at national conferences. In addition it has widely distributed its works, especially to its region, where single copies of products are often free, with a modest cost for multiple copies or for copies sent outside the region. Many works are also available by downloading from the Laboratory's web site. Much of the Laboratory's work addresses current national as well as regional needs (e.g., reading, school improvement, and standards-based assessment).

### **2. Areas of needed improvement**

When it is appropriate, the Laboratory needs to conduct rigorous studies to evaluate the impact its products or services have on students.

### **3. Recommendations for improvement**

The Laboratory also should make explicit what it feels is the appropriate time to try to measure impact on students.

#### **B. To what extent does the Laboratory assist states and localities to implement comprehensive school improvement strategies?**

##### **1. Strengths**

The volume of work conducted by NWREL is impressive. In FY97 the Laboratory had about 220 formal contracts with schools, districts, educational agencies, and communities. The Laboratory as a whole conducted 3,057 activities in which 25,419 individuals participated (approximately 13,000 were teachers or other local school staff). There were several reports of clients citing the Laboratory as being instrumental in getting them to implement their comprehensive school improvement efforts.

Scaling up efforts are a regular part of the Laboratory's efforts. The Laboratory has most often shown a willingness to make their products accessible to a wide audience. It has also sponsored a number of major conferences geared toward enhancing school improvement.

Once the Laboratory gets involved with schools, it maintains support through implementation and follow-up. It has provided focus to schools' efforts to employ best practices; for example, it has made available literature reviews as well as booklets designed to enhance a best-practice approach. The Laboratory's web site has high usage as evidenced by more than a million hits annually.

##### **2. Areas of needed improvement**

None noted.

##### **3. Recommendations for improvement**

None noted.

**C. To what extent has the REL made progress in establishing a regional and national reputation in its specialty area?**

**1. Strengths**

NWREL is currently providing services or having its products and programs used in places widely spread out geographically throughout the U.S. In FY97 more than 13,000 educators participated in NWREL-sponsored events. The Laboratory has cooperative working relationships on specific projects with nationally prominent institutions or agencies such as CRESST at UCLA, the University of Michigan, and the International Reading Association.

The Laboratory and partners such as ASCD and NCTE have published numerous Laboratory-developed materials. NWREL is well represented in professional journals and presentations at professional conferences. As noted elsewhere, a number of Laboratory efforts have received national recognition.

**2. Areas of needed improvement**

While recognizing the realities of what is feasible when it comes to developing working arrangements with schools or districts, it does appear that a notable part of the distribution is largely circumstantial rather than one of logic or planning. For example, the large number of Mississippi sites using OTE and OTE II is mainly due to corporate assistance from Weyerhaeuser, who in turn is especially concerned because of their wanting to ensure a work force of adequately trained high school graduates.

It was somewhat surprising to hear from an Oregon elementary principal who reported that his school heard about OTE II by just happening to hear about it from another school at a meeting.

**3. Recommendations for improvement**

The Laboratory might consider (re)investigating how to best inform those districts and educators that apparently are still unaware of its services and products.

### **Other comments**

Laboratory administration has asserted that NWREL conducts its own business consistent with the approaches and philosophy they are advocating for the schools. The Laboratory's claim of ethnic diversity among its staff has been backed up by statistics showing a reasonably representative proportion of members of the ethnic groups from the surrounding regions. Its gender representation is of possible concern. Even though the top two executive positions are held by women, according to the 1997 institutional evaluation report, 73 percent of the long-term staff were female, as were 88 percent of the technical and clerical, 69 percent of the mid-level professional, and 42 percent of the executive and administrative staff. Corresponding results for the second quarter of FY99 are 81 percent of the technical and clerical, 65 percent of the mid-level professional, and 45 percent of the executive and administrative staff. Having made a point of noting that its ethnic composition was reasonably close to that of the region, the Laboratory should discuss whether its non-representative gender composition needs addressing.

Causing some difficulty for the panelists were the large number of documents and NWREL's use of a large number of categories or terms, sometimes in different renditions. For example, the Strategic Plan has seven areas that are referred to in different documents as strategic functions, strategic initiatives, and strategic areas.

Another logistical area of difficulty was the co-mingling of information such that it was often not obvious to the panel whether the information given applied to the OERI portion of the Laboratory's funding or whether it applied to the Laboratory as a whole.

Despite the aforementioned concerns, it was clear that the Laboratory fully cooperated in the evaluation. Without such excellent cooperation, the evaluation would have been almost impossible to carry out within the agreed upon parameters.