Background on the Development of the EFF Assessment Framework, 1999–2000

In January 1999, as NIFL/EFF was conducting the third and final field review of the EFF Content Standards, we took the first steps toward developing an assessment framework for the Standards. We began by reviewing analogous efforts in England, Australia, and South Africa. We also commissioned a series of papers focused on issues important to assuring that programs and states could reliably and validly assess and report progress in relation to the EFF Standards. Intended primarily for internal use by the EFF team, these papers included

- a broad look at the issues of developing national assessments and national performance standards by Archie LaPointe of the Educational Testing Service,
- a "road map" for a process for developing performance standards for EFF by Regie Stites of SRI International,
- a review by Michelle Della Rosa of HumRRO and Joan Wills of the Institute for Educational Leadership (IEL) that looked more closely at the construct of each of the EFF Standards, and
- a paper by Sri Ananda of WestED intended to assist teachers in assessing student performance for in-program purposes using the EFF Standards. This paper, *How Instructors Can Support Adult Learners Through Performance-Based Assessment* is the only one prepared for general distribution.

At the same time, we invited our field development partners to work with us to clarify the various purposes a comprehensive assessment system for adult education needs to address[3] and to help us develop a set of Guiding Principles for Developing the EFF Assessment Framework (see page 4). Taken together, our Guiding Principles all pointed to using cognitive science research on the development of expertise as the theoretical underpinning for the EFF Assessment Framework. This research base enabled us to conceptualize a single continuum of increasingly skilled performance that includes all adult performance—from novice to expert. It also provided a starting point for defining a small number of key dimensions that distinguish performances along the continuum.

Defining the EFF Continuum of Performance was the first of three primary tasks toward fully developing the EFF Assessment Framework. The remaining tasks were

- to develop a continuum of performance for each standard, with levels that describe real-world competence and
- to identify and develop tools to assess performance for the range of assessment purposes.

[3]See Appendix B for our Purposes of Assessment chart, adapted from R.J. Mislevy (1994).

Defining the EFF Continuum of Performance

In order to assure that our model of performance reflected real teaching and learning situations, in the spring of 1999, we invited teachers at 10 EFF field development sites, who were already using the Standards to plan and guide instruction, to also document and evaluate learner performance.[4] Teacher reports on learner progress were analyzed by using the research on expertise to clarify which Dimensions of Performance were important to consider in rating how well an individual could use an EFF skill to carry out real-life tasks.

Through this review process, we identified four key dimensions that characterized progress in using a Standard along a continuum.[5] These Dimensions of Performance, validated through data collected by EFF field sites in 1999 and 2000, included the following:

1. Structure of Knowledge Base,
2. Fluency of Performance,
3. Independence of Performance, and
4. Range of Conditions for Performance.[6]

Our initial conception of this continuum was represented by a graphic that showed independence, fluency, and range of performance increasing as the knowledge base deepened (see Figure 1).

Developing a Continuum of Performance for Each EFF Standard

By the end of 1999, we were ready to begin the process of building performance continua based on these four dimensions. Teachers and tutors from 15 field sites in our five partner states used standardized performance templates and a standard data collection reporting protocol to collect and report data on student performance. The protocol guided teachers through a process of using the standards and dimensions to plan and carry out lessons centered on EFF "learning tasks" and to collect evidence of learner performance in relation to these tasks. Each teacher was responsible for reporting on only 2 of the 16 Standards—one Standard from the communication group and one from another group. EFF staff worked with all the sites to make sure that at least three teachers were collecting data on each standard—more in the case of the communication standards.

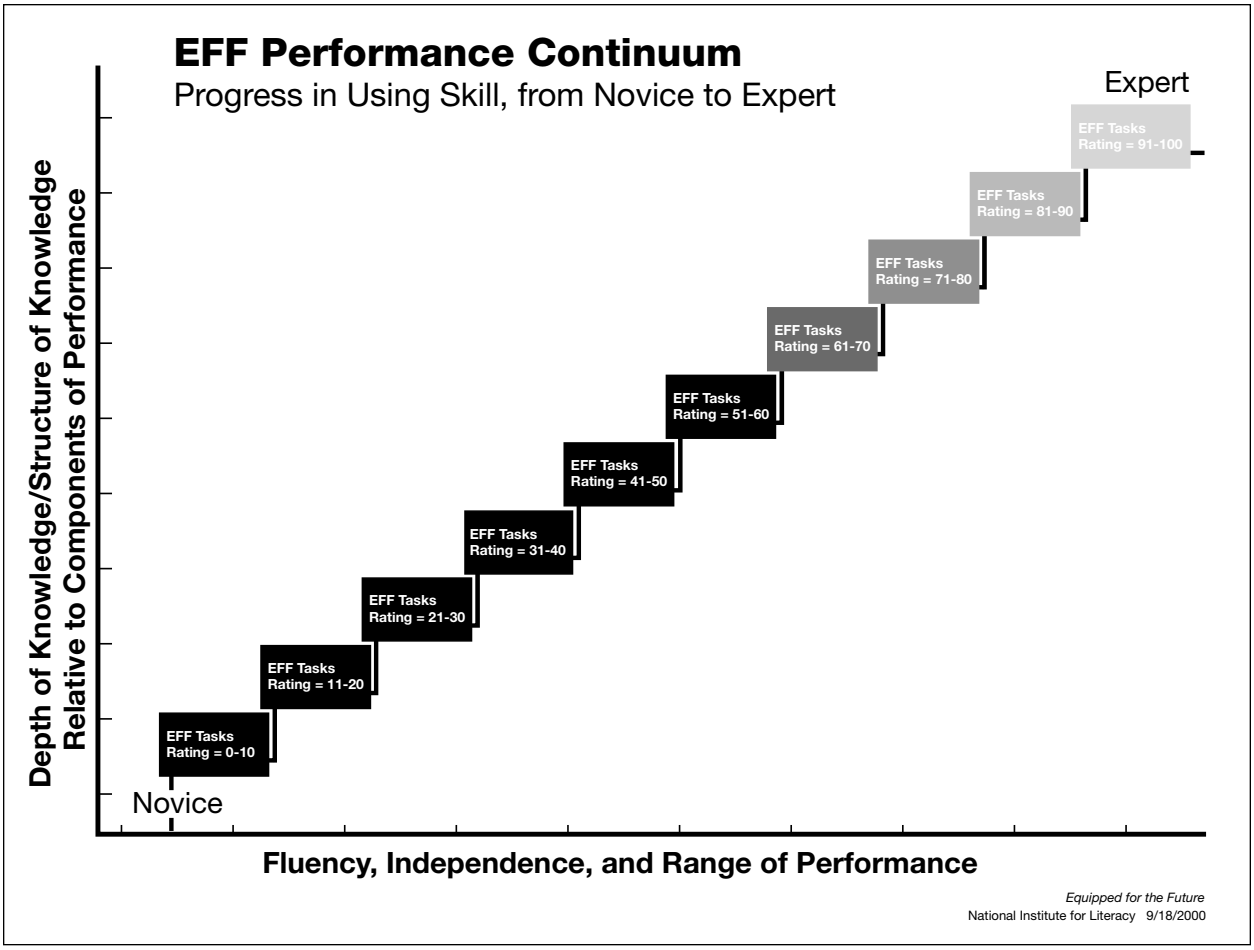
Review of this first round of data, followed by a midcourse technical assistance meeting with prac-

[4]See Sondra G. Stein (2000) *Equipped for the Future Content Standards: What Adults Need to Know and Be Able to Do in the 21st Century.* (Washington, DC: National Institute for Literacy), See Appendix B, for a full description of the documentation protocol.

[5]Primary sources were Wittrock and Baker (1988) *Testing and Cognition* (Englewood Cliffs: Prentice Hall), Bransford et al., *How People Learn: Brain, Mind Experience and School* (Washington DC: National Academy Press), and a wide range of technical reports from the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA.

[6]For more information on the EFF Dimensions see *Equipped for the Future Content Standards* pp. 59-60 and Peggy M McGuire (2000) "A Performance Framework for Teaching and Learning With the Equipped for the Future (EFF) Content Standards," *Adventures in Assessment*, Vol.12, Winter, 2000, pp. 28-43 (Boston: SABES/World Education).

Figure 1. EFF Performance Continuum



titioners in the spring of 2000, enabled us to improve the technical quality of our data collection methods in three areas crucial for the work of the current EFF/NRS Data Collection Project:

1. **Sharpening the focus on the "construct" or conceptual model of each EFF Standard.** Data from teachers helped us understand that a well-structured Performance Task must focus equally on the full Standard and on the Dimensions of Performance. In their effort to create tasks that provided opportunities for students to develop along all four dimensions of performance, teachers at first lost focus on the Standard that was the target of their effort. They created interesting, real-world performance tasks, which often did not, in fact, provide evidence of use of the Standard on which they intended to focus. We clarified that the construct for each Standard is composed of the Components of Performance of the Standard (i.e., the full Standard statement) as modified by the four Dimensions of Performance.
2. **Sharpening the distinction between the requirements of a task created to elicit evidence of performance and the descriptions of learner performance on this task.** To help clarify this critical distinction, we worked with our field partners to develop criteria for both defining a *well-structured Performance Task* and generating rich descriptions of learner performance. Placement of a performance-task description on the continuum does not vary from learner to learner and is determined by the complexity and context of the task, including the knowledge required to perform it. The placement of descriptions of

learner performance on the continuum does vary from learner to learner and is determined by the knowledge, strategies, fluency, and independence demonstrated in performance on a task.

3. **Increasing interrater reliability on the placement of tasks and learner performance on the continuum.** Working through the issues described above gave teachers a set of more objective criteria for deciding where to place tasks and descriptions of learner performance on a continuum. In the mid-course technical assistance meeting, we practiced coming to consensus on placing tasks on the continuum and developed increased ability to judge both tasks and performance in light of explicit, common criteria.

Through this collaborative work with field sites, we refined data collection procedures to assure that teachers and tutors paid close attention to whether the learning tasks they were developing were well structured and that descriptions of learner performance on an EFF Performance Task focused clearly on *how well* a student can use the targeted skill (as defined by the Standard) in carrying out a task. To assure that we were collecting data based on the same construct of a specific Standard, we created templates that provided teachers with a common language for describing tasks and learner performance relative to the four dimensions. These refinements in our data collection practices enabled us to be certain that the data we were collecting could be aggregated to build a valid continuum of performance for that Standard.

By the summer of 2000, the EFF Assessment Team had built a strong foundation for the EFF/NRS Data Collection Project. We had defined the EFF Purposes of Assessment (Appendix B), identified the four Dimensions of Performance upon which to build a cognitive model of adult performance, and begun the process of observing student performance of EFF Standards by using a performance framework and template based on these four dimensions. We had revised our data collection protocols based on feedback from our partner teachers/researchers, and we had prepared a guide for EFF field researchers that focused much more clearly on the importance of basing EFF performance data on *well-structured Performance Tasks.* The rich body of data and data collection tools we had developed prepared us to create a partnership with the Division of Adult Education and Literacy of the U.S. Department of Education's Office of Vocational and Adult Education and to continue the process of building the continua with the clear intention of aligning the EFF assessment system with the U.S. Department of Education's National Reporting System.