

Guidebook
for the Video Program

**PORTFOLIO ASSESSMENT AND
HIGH TECHNOLOGY**

National Center for Research on Evaluation,
Standards, and Student Testing

Table of Contents

	<i>About Our Centers</i>	<i>iii</i>
Part I:	<i>Introduction to Portfolio Assessment</i>	<i>1</i>
Part II:	<i>Getting Started.....</i>	<i>3</i>
Part III:	<i>Portfolio Contents.....</i>	<i>5</i>
Part IV:	<i>Assessing the Portfolios</i>	<i>6</i>
Part V:	<i>Questions and Answers About Portfolios.....</i>	<i>13</i>
Part VI:	<i>Resources for Teachers, Schools and Districts</i>	<i>16</i>
Part VII:	<i>Glossary of Useful Assessment Terms</i>	<i>20</i>

About Our Centers...

UCLA Center for the Study of Evaluation & National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

For over 28 years, the UCLA Center for the Study of Evaluation (CSE) has been at the forefront of efforts to improve the quality of education in America through systematic evaluation practices. CSE has helped pioneer valid and sensitive evaluation and assessment techniques and has promoted vigorously the use of evaluation for more reasoned decision making.

Located within the UCLA Graduate School of Education & Information Studies, CSE is a unique center devoted to educational research, development, training, and dissemination. The result of a national competition, CSE was designated the national center for research in educational evaluation in 1966. This charge was renewed in 1990 when CSE successfully competed for the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), receiving a five-year, \$14 million grant from the Office of Educational Research and Improvement (OERI).

CRESST Initiatives

CRESST projects are conducted in collaboration with colleagues from the University of Colorado at Boulder, University of California, Santa Barbara, National Opinion Research Center at the University of Chicago, Learning Research and Development Center at the University of Pittsburgh, the University of Southern California, and The RAND Corporation.

CRESST Program One, *Building the Infrastructure for Improved Assessment*, seeks to assure that the R&D efforts of CRESST and the many other players currently working on assessment reform will have optimal impact on educational policy and practice. Program Two, *Designs for Learning-Based Assessments: Prototypes and Models*, is creating new prototypes for assessing student performance and new models for analyzing and validating assessment results.

Program Three, *Collaborative Development and Improvement of Assessments in Practice*, also addresses the creation of alternative assessments. However, this developmental work is conducted collaboratively with state and/or local constituencies, subject to their time and resource constraints, and targeted to their goals and objectives.

CSE/CRESST Dissemination

CSE/CRESST produces a variety of dissemination materials. A quarterly newsletter, *CRESST Line*, contains up-to-date news on current CRESST assessment research. *Evaluation Comment*, a semiannual research-based publication, and *CRESST Line* are distributed free to a mailing list of over 11,000 education policy makers, researchers, and practitioners throughout the United States and overseas. For more information about CSE or CRESST, please contact:

Ronald Dietel, Director of Communications
TEL (310) 206-1532, FAX (310) 825-3883
CSE/CRESST
UCLA Graduate School of Education &
Information Studies
405 Hilgard Avenue, 1320 Moore Hall
Los Angeles, CA 90024-1522

The work reported herein was supported under the Educational Research and Development Center Program cooperative agreement number R117G10027 and CFDA catalog number 84.117G as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in any CRESST publication or product do not reflect the position or policies of the Office of Educational Research and Improvement or the U.S. Department of Education.

Part I: *Introduction to Portfolio Assessment*

Great Expectations for Portfolio Assessment

I feel that portfolios are a very useful tool in assessing student writing, for seeing progress, and what students enjoy about their writing.

Phyllis Lewcock
Stevens Creek Elementary School

When students ask me, as they had in the past, why do I need to learn this, I have a real answer now. I tell them that the kinds of skills that they are developing through the portfolio [program] are skills that are going to be serving them for the rest of their lives.

Rita Kirby
Ithaca (MI) Public Schools

Such enthusiastic statements exemplify the type of support portfolios have received from teachers, school districts, and states over the past few years. Perhaps because portfolio assessment is designed to be integrated with instruction, it has become one of most appealing forms of performance assessment. As discussed in *Portfolio Assessment and High Technology*, the potential benefits of portfolio assessment compared to traditional standardized testing are many:

- By showing what they can *do* through their portfolios, students demonstrate skills and competencies for teachers, parents, policy makers and even potential employers. Portfolios provide valuable information useful in evaluating quality of education and quality of student achievement.
- Focused on the products of classroom instruction, portfolio assessment can be integrated with instruction, not added on.
- Portfolios target the assessment of complex thinking, deep understanding, and application of knowledge—rather than more isolated skills and knowledge such as memorization of dates, facts, and formulas.
- Because portfolios offer students a wide variety of ways to demonstrate what they know and can do, students are encouraged to become reflective learners responsible for their own growth.
- Portfolios offer teachers opportunities to understand what their students are learning, and support their efforts to design appropriate instruction that can increase their students' achievement.

Are such great expectations warranted? Is there evidence that portfolio assessment supports student achievement, motivation to learn, or the skills necessary to get a well-paying job in the year 2000? Research from the Center for Research on Evaluation, Standards, and Student Testing (CRESST) is providing some answers.

The First Question: How Will You Use the Results From Portfolio Assessment?

Webster's defines "portfolio" as "a hinged cover or flexible case for carrying loose papers, pictures, or pamphlets." Indeed, for many of us, the word "portfolio" conjures up an image of a photographer or commercial artist schlepping around his or her very best work to potential clients.

For educators, portfolios used to assess a child's achievement have a similar meaning: The child assembles a collection of work that can reveal what he or she knows and can do. But portfolio assessment in the educational setting may well encompass additional uses of the portfolio—e.g., decisions based on student performance, effectiveness of instruction, or even the quality of teachers and schools. In other words, portfolio assessment has many possible "uses."

Any assessment "use" has consequences for students, teachers, or schools, and therefore it is critical to know the intended uses of portfolio assessment results.

Consider the ways that portfolios could be used to assess the performance of students or teachers. Will you use student portfolios to evaluate a student's strengths and weaknesses? to assign a grade to a portfolio on a periodic basis and use it to supplement existing grading? as a sole indicator of student achievement for a semester or year? Will teachers or the entire school be judged on the results of portfolio assessment? Each of these questions reflects a specific use as well as consequences for either students, or teachers, or schools.

Reflect on how portfolios might be used to enhance professional development. In Vermont and California, for example, researchers have shown that participation in a statewide portfolio assessment program has enhanced teachers' knowledge of new state frameworks and increased the likelihood that they have implemented the frameworks in their classrooms.

A first step for teachers or schools then is to define clearly the *use* of the portfolio assessment. Without this initial decision, teachers and schools may waste valuable energies and resources. You cannot design the details of portfolio assessment—specifications of the portfolio contents, rubrics for assessing those contents—without firm decisions "up front" regarding the ways that the resulting scores will be used.

Part II: *Getting Started*

First Steps

As mentioned in *Portfolio Assessment and High Technology*, just collecting children's work into portfolios limits the usefulness of portfolios for assessment. The full potential for portfolios to assess children's competence and progress can only be met if there is agreement on:

1. **standards**—what students should know and be able to do;
2. **contents**—what should go into the portfolio to demonstrate what children know and can do; and
3. **assessment**—how the portfolio material should be assessed.

We discuss each of these topics in the remainder of this guidebook.

Standards: What Should Students Know and Be Able to Do?

The federal government has encouraged the development of national standards in a wide variety of subject areas, from mathematics to vocational education. To decide what students at your school should “know and be able to do,” you and your colleagues might well start with a review of the voluntary national standards as well as any available state or local curriculum frameworks. If you are working in an interdisciplinary area, review all frameworks pertinent to your instructional programs. The Regional Laboratory for Educational Improvement of the Northeast & Islands has developed the *Status Report on Voluntary National Standards in Education* outlining 13 standards-setting efforts, including status and contact information. (See Part VI: Resources for Teachers, Schools and Districts.)

Local resources within your school, district and community are invaluable to the process of developing, adapting and implementing standards. In *Portfolio Assessment and High Technology*, for example, CRESST researchers Maryl Gearhart and Shelby Wolf worked with Stevens Creek Elementary School teachers to brainstorm examples of good writing and create standards based on those examples. Gearhart and Wolf provide some valuable pointers on how to get started with standards whether from national, state or local sources.

Good assessment starts with well thought-out standards. Of course we're focusing here on portfolio assessment, but the decision on the type of assessment should wait until after you have established a clear understanding of what students should know and be able to do. You may find at this point that portfolios are a reasonable way to assess student learning or that other types of assessment are more suitable.

Start With Standards

1. Develop a specific strategy for the creation of standards in each subject area. Ask yourselves some basic questions: Where are our greatest strengths in subject matter expertise? Where are we weak?

We recommend that you start in your strongest areas first. Once you have created standards in one area, the next ones will become successively easier. Remember that implementation of standards depends on solid understandings of subject matter.

2. Link your efforts to your local goals and resources. Ask:
 - What are our goals for school improvement in each subject area?
 - What topics and concepts need to be covered in what grade levels?
 - What are the school resources available, and what grants might be possible?
 - What is a reasonable time frame? Is there a professional development person available to assist us?
3. Build support: Who is available and willing to work in which areas? If you are a teacher, enlist support from your principal, parents and any stakeholders who have an interest in your efforts. If you are a principal, you will need at least a good percentage of your teachers who are willing to work through the process.
4. Discuss students' work. A useful beginning point at your first meetings is to review and analyze samples of students' work. Not only will everyone gain an appreciation for what others are doing, but consensus will be built around what kids *should* be doing. Contribute actively to the discussion, and make constructive comments.
5. Build a library of resources in each subject area and make them available to all teachers at the school. You'll find our own suggestions in Part V of this guidebook.

After a few meetings with other teachers, you are likely to accumulate a pretty overwhelming amount of material related to standards. At this point you will need to prioritize your standards. Ask yourself what kids *must* know and what would be *useful* for them to know. Ask yourself what *you* can feasibly achieve with the resources currently available to you. Set high goals, but don't feel that you will be able to cover all the standards you have discussed.

We strongly recommend that you include parents and other important stakeholders in the standards-setting process. Parents, for example, are likely to have strong opinions, positive or negative, about instructional changes. Including parents in your planning will improve the likelihood that your program will have a broad range of support.

Part III: *Portfolio Contents*

Portfolio Contents: What Will We Assess?

Because portfolios are not standardized, it becomes a challenge to use them to assess learning. How can Ms. Juarez compare one writing portfolio that contains mostly poetry with another that contains mostly letters? There are no set criteria for what must go into a portfolio, and, indeed, imposing such a requirement might destroy teacher enthusiasm and support for portfolio assessment. Nevertheless, it is essential that you establish criteria for portfolio contents, or the usefulness of your portfolio scores will become questionable.

What should be included in the portfolios? How to decide? Your goal is to assess your students' progress toward the standards you have developed. Therefore, portfolio tasks must be designed to reveal your students' competency with a particular standard. In addition, the assignments should be designed in a way that the products of students' work—the material that will be placed in the portfolio—will reveal what students know and can do. A student's final report, for example, may not tell you much about the process of planning and carrying out the research.

At this point, you may discover the benefits of fairly well-specified standards. If your standard says, for example, "*students should know and be able to use the experimental method,*" then the task is fairly well defined. You then design an opportunity for scientific experimentation in a way that will ensure assessable work appropriate for the science portfolio. Your students might record their initial hypotheses, the procedures they devised, their findings from each procedure, and their conclusions. They might also reflect in writing on what was challenging to understand in the process.

If your standard says "*students should be able to integrate prior and newly acquired knowledge of important historical periods into an effective essay,*" you have even more guidance in the design of your assignments. "Prior" knowledge might be the content that you recently covered related to the Civil War; "newly acquired" could be original source materials related to the issue of slavery; and "essay" refers to the type of performance expected. The preceding example, in fact, represents a performance assessment developed by CRESST that requires students to read and understand a portion of the famous Lincoln/Douglas debates from the pre-Civil War era, draw on prior knowledge and recent instruction about this historical period, and write an extended essay demonstrating their integration and understanding of the material.

The assignments that you design will provide opportunities for students to work toward your standards as well as opportunities to use portfolios to assess their progress. Do consider providing students multiple opportunities to engage in work related to each standard, so that the portfolio can reveal students' progress over time.

Part IV: *Assessing the Portfolios*

Assessment: How Will We Assess the Portfolios?

Without standards and without scoring, portfolios remain simply collections of student work. Scoring criteria enable teachers and school administrators to answer the questions “How are our students doing? Are students achieving high standards? Is their work improving over time?”

Good scoring criteria:

- Clarify instructional goals for teachers, parents and, most importantly, students;
- Enhance fairness, because students know exactly how they will be assessed and what learning outcomes are most important;
- Help teachers or other raters to be accurate, unbiased and consistent in scoring.

Scoring Criteria Are Tied to Use

Before you can develop scoring criteria, you must once again revisit the *intended use* of your portfolio scores. Will the results be used to measure how well students perform to a specific *level* of achievement, such as “basic,” “proficient,” or “advanced?” In this case, your criteria need to specify exactly what performance will represent each achievement level. Or will you use the results to measure student *progress* over a period of time, say during the academic year? In this case, you need to design criteria to assess changes in student skills or knowledge.

Your decision about use will also shape your decision about the type of portfolio rubric. Should the rubric be holistic—producing a single score—or analytic—producing scores on several distinctive scales? You may feel that a holistic rubric is adequate if your goal is to provide a general index of students’ overall competence or progress. If you are striving to represent distinctive skills and knowledge, however, you may decide to construct several scales.

Criteria Example

The Narrative Rubric on the following page may assist you in developing your own criteria. Originally developed for assessment of writing samples, this

rubric contains analytic scales representing various components of performance, e.g., theme, character, setting, plot and communication. The rubric also represents student development, the stages through which most children seem to progress as they master the subject or skills.

We think you'll find the Narrative Rubric adaptable to domains other than writing.

Narrative Rubric



Theme

explicit ↔ implicit
didactic ↔ revealing

- Not present or not developed through other narrative elements
- Meaning centered in a series of list-like statements (“I like my mom. And I like my dad. And I like my...”) or in the coherence of the action itself (“He blew up the plane. Pow!”)
- Beginning statement of theme—often explicit and didactic (“The mean witch chased the children and she shouldn’t have done that.”); occasionally the theme, though well stated, does not fit the story
- Beginning revelation of theme on both explicit and implicit levels through the more subtle things characters say and do (“He put his arm around the dog and held him close. ‘You’re my best pal,’ he whispered.”)
- Beginning use of secondary themes, often tied to overarching theme, but sometimes tangential; main theme increasingly revealed through discovery rather than delivery, though explicit thematic statements still predominate
- Overarching theme multi-layered and complex; secondary themes integrally related to primary theme or themes; both explicit and implicit revelations of theme work in harmony (“You can’t do that to my sister!” Lou cried, moving to shield Tasha with her body.)



Character

flat ↔ round
static ↔ dynamic

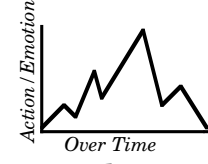
- One or two flat, static characters with little relationship between characters; either objective (action speaks for itself) or first person (author as “I”) point of view
- Some rounding, usually in physical description; relationship between characters is action-driven; objective point of view is common
- Continued rounding in physical description, particularly stereotypical features (“wart on the end of her nose”); beginning rounding in feeling, often through straightforward vocabulary (“She was sad, glad, mad.”)
- Beginning insights into the motivation and intention that drives the feeling and the action of main characters often through limited omniscient point of view; beginning dynamic features (of change and growth)
- Further rounding (in feeling and motivation); dynamic features appear in the central characters and in the relationships between characters; move to omniscient point of view (getting into the minds of characters)
- Round, dynamic major characters through rich description of affect, intention, and motivation; growth occurs as a result of complex interactions between characters; most characters contribute to the development of the narrative; purposeful choice of point of view



Setting

backdrop ↔ essential
simple ↔ multi-functional

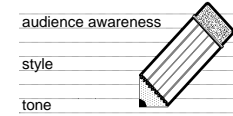
- Backdrop setting with little or no indication of time and place (“There was a little girl. She liked candy.”)
- Skeletal indication of time and place often held in past time (“once there was...”); little relationship to other narrative elements
- Beginning relationship between setting and other narrative elements (futuristic setting to accommodate aliens and spaceships); beginning symbolic functions of setting (often stereotypical images—forest as scary place)
- Setting becomes more essential to the development of the story in explicit ways: characters may remark on the setting or the time and place may be integral to the plot
- Setting may serve more than one function and the relationship between functions is more implicit and symbolic—for example, setting may be linked symbolically to character mood (“She hid in the grass, clutching the sharp, dry spikes, waiting.”)
- Setting fully integrated with the characters, action, and theme of the story; role of setting is multifunctional—setting mood, revealing character and conflict, serving as metaphor



Plot

simple ↔ complex
static ↔ conflict

- One or two events with little or no conflict (“Once there was a cat. The cat liked milk.”)
- Beginning sequence of events, but occasional out-of-sync occurrences; events without problem, problem without resolution, or little emotional response
- Single, linear episode with clear beginning, middle, and end; the episode contains four critical elements of problem, emotional response, action, and outcome
- Plot increases in complexity with more than one episode; each episode contains problem, emotional response, action, outcome; beginning relationship between episodes
- Stronger relationship between episodes (with the resolution in one leading to a problem in the next); beginning manipulation of the sequence through foreshadowing, and subplots
- Overarching problem and resolution supported by multiple, episodes; rich variety of techniques (building suspense, foreshadowing, flashbacks, denouement) to manipulate sequence



Communication

context-bound ↔ reader-considerate
literal ↔ symbolic

- Writing bound to context (You have to be there) and often dependent on drawing and talk to clarify the meaning; minimal style and tone
- Beginning awareness of reader considerations; straightforward style and tone focused on getting the information out; first attempts at dialogue begin
- Writer begins to make use of explanations and transitions (“because” and “so”); literal style centers on description (“sunny day”); tone explicit
- Increased information and explanation for the reader (linking ideas as well as episodes); words more carefully selected to suit the narrative’s purpose (particularly through increased use of detail in imagery)
- Some experimentation with symbolism (particularly figurative language) which shows reader considerations on both explicit and implicit levels; style shows increasing variety (alliteration, word play, rhythm, etc.) and tone is more implicit
- Careful crafting of choices in story structure as well as vocabulary demonstrate considerate orchestration of all the available resources; judicious experimentation with variety of stylistic forms which are often symbolic in nature and illuminate the other narrative elements

Achievement Levels and Criteria

To build shared understandings of the new assessments with colleagues, parents and students, it will be useful for you to include examples of student work representing each achievement level of your scoring rubric. The writing rubric below and on the following pages provides guidance. Note that the descriptions at each level of the scale help to communicate clearly what it is that students are able to do. The authors of this developmental rubric were especially concerned to describe what students *can* do at each level and to avoid slipping into descriptions of what students cannot do. Positive descriptors will help to clarify your standards to students and parents.

Theme

Rubric level	Example
Not present or not developed through other narrative elements	<p><u>A Frog and Toad Adventure</u></p> <p>Frog and Toad and snake there in a forest. They get bit by a snake friendship. (1st grader)</p>
Meaning centered in a series of list-like statements (“I like my mom. And I like my dad. And I like my....”) or in the coherence of the action itself (“He blew up the plane. Pow!”)	<p><u>Friends</u></p> <p>Adam and Larry don’t like each other. One day Adam got hurt by tripping on a rock and Larry helped Adam and then they were friends. They like to go to school together. They like to play together. They like to read books together. They were best friends. (2nd grader)</p>
Beginning statement of theme—often explicit and didactic (“The mean witch chased the children and she shouldn’t have done that.”); occasionally the theme, though well-stated, does not fit the story	<p><u>The Easter Play</u></p> <p>It was a stormy Saturday and I (Karen) was trying to get my little sister (Holly) to wear a gypsy costume. Then my big brother (Larry) came in and started making fun of the costume because I had made it myself.</p> <p>“How would you like it if you made a costume and somebody made fun of yours?” “Huh?” Well that fixed him. (2nd grader)</p>

Rubric level	Example
<p>Beginning revelation of theme on both explicit and implicit levels through the more subtle things characters say and do (“He put his arm around the dog and held him close. ‘You’re my best pal,’ he whispered.”)</p>	<p><u>Wilbur Spends a day with Fern</u></p> <p>...Fern flipped over the old milk pail and started talking to Wilbur.</p> <p>“Wilbur, I have missed you a lot so I decided to spend a day with you.” She put him on a leash and took him to the park. Wilbur couldn’t believe that this was really happening. He was so happy he didn’t watch where he was going, and he fell into the little duck pond. When he lifted his head he had a lily pad on his head. Fern took Wilbur out of the pond and dried him off in the grass. They decided to have lunch. They sat under a tree and watched the leaves fall off the tree and blow in the wind.</p> <p>Soon the day was over and Fern had to take Wilbur back to his pen. It was hard to see Fern go but now Wilbur knew she loved him. (4th grader)</p>
<p>Beginning use of secondary themes, often tied to overarching theme, but sometimes tangential; main theme increasingly revealed through discovery rather than delivery, though explicit thematic statements still predominate</p>	<p><u>My Sister, Kristi</u></p> <p>We stayed there all night and then in the morning I went into check on her, and then I looked if there was a heart beat. There was nothing. I screamed for the doctor. Then Mom and Dad rushed into the room, the doctor told me that she was dead, and it looked like she had been dead for a while. I couldn’t believe it, and I wouldn’t believe it! I thought it was all my fault, letting Mike spike the punch. His parents were the only ones that had an alcohol problem. I was ready to faint. I still couldn’t understand why it had to happen to my sister. I missed her already, even though she just left me.</p> <p>...I learned a BIG lesson from the mistake that Kristi made. It’s a mistake that I will never make in my life and I hope no one ever will. I know that I can tell my friends that getting drunk and driving, and disobeying your parents when they trusted you, can ruin or risk your life forever. Even if it’s not your fault. I think that you should really take your responsibilities seriously. You should never do drugs, smoke, or drink alcohol. That is something that can get you in a lot of trouble, or can really risk your life forever. That’s something that I won’t do, because I really do want to live for as long as I can....And if you want to commit suicide, talk about it. (6th grader)</p>

Rubric level	Example
Overarching theme multi-layered and complex; secondary themes integrally related to primary theme or themes; both explicit and implicit revelations of theme work in harmony (“You can't do that to my sister!”, Lou cried, moving to shield Tasha with her body.)	No example available.

Assigning Scores

If the purpose of your portfolio assessment makes it important to assign a traditional numerical sequence, e.g., 1 to 3, 1 to 5, 1 to 6, to your rubric levels, then of course that can be done. There is no magic formula for the number of levels. As previously mentioned, some rubrics for portfolios or performance assessments have only 3 levels (basic, proficient or advanced), and other rubrics have 4, 5 or even 6. Of course, the greater the number of levels, the greater may be your capacity to differentiate levels of student achievement.

Another advantage to using many levels is that it will let you set a very high standard. Indeed, note that there are no examples for the highest level in the theme rubric above because none of the students in our studies has yet written a narrative that complex!

Creating well-differentiated and clearly communicated criteria is a time-consuming process. But once you have mastered the process with one rubric, designing additional rubrics should become much easier. Patience and persistence pay off.

Reliable Scoring Leads to Fairness

To ensure fairness in the rating of portfolios, we believe that multiple raters should be involved in your scoring process. Even if you are the sole user of your students' portfolio results, having a second teacher score the portfolios will help to ensure that your scores reflect the competence of the work in the portfolio. This colleague should be someone with excellent content knowledge and someone who is willing to disagree with you. Naturally, you should be willing to review portfolios from your colleague's students.

Another way to check the reliability of your scoring is to re-score the same portfolio after a set period of time, perhaps two months or more after the first scoring, so that you are not as likely to remember the previous score you gave

to a portfolio. Compare your two sets of scores and evaluate whether or not you have been consistent over time.

For portfolios where students may be compared on a schoolwide basis, a more formal process should be established. In these cases, a group of teachers may periodically come together to score portfolios, perhaps at the end of every semester. Anchor papers or samples of student work from the portfolios should be used by teachers to reach common agreement on the “fit” between student work and each level of your rubric. Although such an undertaking is no small accomplishment, teachers in Vermont and other states have said that formally rating portfolios has been one of the best professional development opportunities they have experienced.

What About Bias?

How will you know if the portfolio tasks are free from bias? We suggest that you share the portfolio tasks with a variety of teachers who have diverse cultural backgrounds, as well as track how students from different backgrounds perform on the individual tasks. A task, for example, that is heavily dependent on language may unfairly disadvantage Limited English Proficient students. If you find that these types of students are not performing well on some specific tasks, you will want to reassess the fairness of those items.

Part V: *Questions and Answers About Portfolios*

Portfolio Assessment Questions We *Know* You're Asking!

In our CRESST research, we have found that nearly all teachers encounter challenges as they implement portfolio assessment into their classrooms and schools. Some of these issues are demonstrated in *Portfolio Assessment and High Technology*. Let's discuss some of the questions now.

- *Will teachers, parents and other students be permitted to assist children in the creation of their portfolios? If so, how will we differentiate the student's work and effort from that of others?*

This is one of the most perplexing questions related to portfolio assessment. As yet, no definitive answer exists. Student learning will be enhanced from the guidance of others and from the reactions of others to the work in their portfolios, but then—as one Vermont teacher put it—“whose work is it?” One possible solution may be to designate certain tasks within a portfolio to represent “individual student work *without* input from others” and then use assessment of these tasks to make higher stakes decisions. Meanwhile, the full portfolio could represent “student work *with* input from others,” thus maintaining the benefits of portfolios for dialogue, ongoing assessment, and student reflection.

- *I found the portfolio model in Portfolio Assessment and High Technology quite intriguing. Should my students have both a “working” portfolio and a “showcase” portfolio?*

The answer depends on your use of the portfolio for assessment and your recognition that the terms “working” and “showcase,” though quite common, may refer to different portfolio models. The teachers at Stevens Creek School, for example, simply ask students to keep all of their work in their “working portfolio” as a management strategy. In other projects, a “working portfolio” may be more akin to a “process portfolio,” a portfolio that contains the various phases of projects, and so can be used for assessment of a student's growth from initial idea to draft and then final product.

The “showcase portfolio” is student-selected, and thus it encourages students to reflect on their best work and share those accomplishments with others. Stevens Creek elementary teachers ask students to present their showcase portfolios to their next teacher, to provide the new teacher with examples of their achievements the prior year.

- *I've heard so many suggestions for what the portfolio should contain! Should the portfolio contain "all" student work or just "best" pieces? A letter of introduction by the student? A student resume or autobiography, a table of contents, a journal of self-reflection, self-evaluation? Special achievements?*

Your answer to these questions may be "none," "some," or "all of these!" You will decide what to include based on what evidence you need for your particular assessment. If, for example, you are assessing how students approach a writing project and carry it through, then drafts and other process materials are essential; a student's reflection on her process could provide additional insight; and, you may want two or more sets to track students' progress over time.

For a portfolio that will be passed to the next grade, inclusion of cover letters from both you and the student could be very helpful to the next teacher. Whatever your portfolio purpose, for handling and storage considerations, you will likely develop a program that targets just the critical evidence for your assessment and allows you to weed out less relevant material.

We suggest that teachers who are interested in or currently working with portfolios meet in small groups to address the above questions and others that arise.

A Summary of Our Recommendations

Over the past four years, CRESST research has garnered some excellent insights from working with teachers on a number of portfolio projects. The same issues come up time and again, and, consequently, we close with a summary of the issues that we hope will help you think through the portfolio process.

- Ask yourself "what is the purpose of the portfolio?" How will you *use* the results? The higher the stakes, the more attention you will have to pay to key technical issues such as the reliability of your ratings.
- Portfolios should be much more than miscellaneous collections of student work. When carefully coordinated with instruction, portfolio assessment provides a powerful learning experience for both students and teachers.
- Developing effective standards, instruction and assessment starts with content knowledge, content knowledge and more content knowledge. You do need to understand your subject area deeply to design substantive and effective standards and assessments.

- Agreement on standards cannot be assumed. Teachers bring to both the instructional and assessment process vastly different experiences and opinions. A trained professional from a school district or teacher support organization may be able to coordinate diverse views during the initial stages of the standard-setting and scoring process.
- Despite all the challenges of design and implementation, available evidence suggests that the implementation of portfolio assessment has positive consequences for teaching and learning. Teachers appear to raise their expectations for students' performance and to spend more time on thinking and problem-solving tasks for students.
- Teachers need a great deal of time to develop, implement and score portfolio assessments. Although teachers may be able to adapt some tasks from other sources, creating or adapting good performance assessment tasks is challenging.

In Conclusion

The expectations for portfolio assessment are indeed great. Early research suggests that portfolio assessment, when well designed and adequately funded, does lead to increases in student skills, achievement and motivation to learn. We suggest that you learn as much as you can both from the resources on the following pages and by talking to other teachers who have successfully implemented portfolio assessment in their own classrooms and schools.

Part VI: Resources for Teachers, Schools and Districts

Many of the following resources are available from CRESST. Call Kim Hurst at (310) 206-1532 or send her an e-mail message at “kim@cse.ucla.edu” for more information. Or write to CRESST at CRESST/UCLA, 10880 Wilshire Boulevard, Suite 700, Los Angeles, CA 90024-4108.

You may also find that your local school district, county or state department of education, or certain nonprofit educational support groups can provide expertise and assistance. Additionally, the American Federation of Teachers and National Education Association have resources available. Contact Eugenia Kemble at the American Federation of Teachers, Educational Issues Department, 555 New Jersey Avenue, NW, Washington, DC 20001, (800) 238-1133; or Rod Riffel at the National Education Association, Educational Policy and Professional Practice, 1201 16th Street, NW, Washington, DC 20036, (202) 822-7384.

Selected Books on Assessment

A Practical Guide to Alternative Assessment

Herman, J.L., Aschbacher, P.R., & Winters, L. (1992). Washington, DC: Association for Supervision and Curriculum Development.

One of the most widely used books on performance assessment. Contains chapters on rethinking assessment, linking assessment and instruction, determining assessment purpose, selecting assessment tasks, setting criteria, ensuring reliable scoring, and using alternative assessment for decision making.

Modern Educational Measurement: A Practitioner's Perspective

Popham, J.W. (1990). 2nd ed. Englewood Cliffs, NJ: Prentice Hall.

A valuable reference for teachers who would like to learn more about assessment in general, including standardized assessments.

Assessing Student Performance: Exploring the Purpose and Limits of Assessment

Wiggins, G.P. (1993). Jossey-Bass education series, Vol. 1. San Francisco, CA: Jossey-Bass.

In Teachers' Hands: Investigating the Practices of Classroom Assessment

Stiggins, R.J., & Conklin, N.F. (1992). Albany: State University of New York Press.

Reports and Handbooks

The Evolution of a Portfolio Program: The Impact and Quality of the Vermont Program in Its Second Year

Koretz, D., Stecher, B., Klein, S., & McCaffrey, D. (1994). (CSE Tech. Rep. No. 385). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Vermont was the first state to implement portfolios on a statewide basis. This CRESST commissioned study by The RAND Corporation discusses changes that occurred to instruction as a result of this portfolio mandate as well as key implementation issues.

CRESST Performance Assessment Models: Assessing Content Area Explanations

Baker, E.L., Aschbacher, P.R., Niemi, D., & Sato, E. (1992). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

A concise model for developing alternative assessments. Contains useful examples of CRESST assessment materials, an effective scoring rubric applicable to multiple topics, and benchmark papers.

Performance-Based Assessment and What Teachers Need

Higuchi, C. (1993). (CSE Tech. Rep. No. 363). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Early attempts to quickly implement performance assessment oftentimes failed to provide teachers with appropriate professional development opportunities. As an early user of performance assessments, Charlotte Higuchi documents in this report many of the needs she had in her own implementation efforts.

Complex, Performance-Based Assessment: Expectations and Validation Criteria

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). (CSE Tech. Rep. No. 331). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

One of the first major research articles that established validity criteria for performance-based assessment. Professors Baker and Linn are the CRESST co-directors, and Professor Dunbar teaches at the University of Iowa.

Writing What You Read: A Guidebook for the Assessment of Children's Narratives

Wolf, S., & Gearhart, M. (1994). (CSE Resource Paper No. 10). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Designed to help teachers think about the important role of assessment in guiding students' narrative writing, this guidebook should help teachers draw close connections between curriculum, instruction and assessment. Provides specific examples for the scoring of student writing, including a narrative feedback form and narrative rubric to help teachers evaluate students' understandings and ability to learn.

Other Useful Publications and Articles

Assessment: Authenticity, Context, and Validity

Wiggins, G. (1993, November). *Phi Delta Kappan*, 75(3), 200-214.

Portfolio News

Cooper, W. (Ed.). (1994, Winter). [Entire Issue]. *Portfolio News*, 5(2).

A good resource for portfolios. Contains thoughtful articles on both theory and implementation. Also documents what is happening in some major portfolio programs across the United States.

The ERIC Review: Special Issue on Performance-Based Assessment

Boston, C.L. (Ed.). (1994, Winter). [Special Issue]. *The ERIC Review*, 3(1).

A useful overview of performance assessment and a guide to valuable resources and organizations.

Status Report on Voluntary National Standards in Education

Harvey, G. (Executive Director). (1994, October). Available from The Regional Laboratory for the Educational Improvement of the Northeast & Islands, 300 Brickstone Square, Suite 950, Andover, MA 01810.

A complete listing of current national standards development projects, their status, and appropriate contact information.

Using Performance Assessment

Brandt, R.S. (Ed.). (1992, May). *Educational Leadership*, 49(8).

An entire issue dedicated to performance assessment. Includes an article by CRESST Associate Director Joan Herman, "What Research Tells Us About Good Assessment."

Databases/Tool Kits

Alternative Assessments in Practice (AAIP) Database

Containing listings from over 250 developers of new assessments, the *Alternative Assessments in Practice* database is of special use to teachers, school district administrators, assessment developers and others interested in new methods for assessing student growth. Released in an easy-to-use Macintosh HyperCard format, the database contains detailed information about each assessment, including subject matter and skills measured, assessment type and purpose, scoring characteristics, and availability of the assessment. The AAIP database, including a 50-page user's manual, is available through the CRESST offices for \$15.00.

***Facilitating Systemic Change in Science and Mathematics Education:
A Toolkit for Professional Developers (In Development)***

This upcoming resource toolkit for teachers will contain information on assessment and effective teaching practices. Contact The Regional Laboratory for Educational Improvement of the Northeast and Islands, 300 Brickstone Square, Suite 950, Andover, MA 01810, (508) 470-0098.

On the Information Highway

CRESST Gopher Server

Available through the CRESST gopher server are articles from the *CRESST Line* newsletter, listings and descriptions of CRESST technical reports, and the CRESST *Alternative Assessments in Practice* database. Computer users with Internet access and a gopher client program may access the CRESST server by starting a new gopher and pointing at: **gopher.cse.ucla.edu**.

CRESST World Wide Web Server

The CRESST World Wide Web (WWW) server provides access to CRESST assessment research in an exciting visual format. Advanced text formatting, graphics, audio, and video make the CRESST World Wide Web server another valuable product for K-12 educators. Internet users with the appropriate software may connect to the CRESST World Wide Web at: **<http://www.cse.ucla.ed/>**.

From the Department of Education

Useful internet servers are maintained by the Department of Education. Connect to their gopher server at: **gopher.ed.gov**; or their World Wide Web server at: **<http://inet.ed.gov/>**.

Northwest Regional Educational Laboratory (NWREL)

This gopher server contains its own alternative assessment database and a wide variety of other useful education databases. The gopher address is: **gopher.nwrel.org**.

Part VII: *Glossary of Useful Assessment Terms*

Age norms—values representing typical or average performance of people of age groups.

Content standards—specify “what students should know and be able to do” in various subjects and domains, such as mathematics or applied learning. Set out the knowledge, skills and other necessary understandings that schools should teach in order for all students to attain high levels of competency in the subject matter.

Criterion-referenced assessment—a measurement of achievement of specific criteria or skills in terms of absolute levels of mastery. The focus is on performance of an individual as measured against a standard or criterion rather than against performance of others who take the same assessment, as with norm-referenced assessments.

Dimensions, traits, or subscales—the subcategories used in evaluating a performance or portfolio product; e.g., in evaluating students’ writing one might rate student performance on subscales such as organization, quality of content, mechanics and style.

Diagnostic assessment—an intensive, in-depth evaluation process with a relatively detailed and narrow coverage of a specific area. The purpose of this assessment is to determine the specific learning needs of individual students and to be able to meet those needs through regular or remedial classroom instruction.

Grade equivalent—the estimated grade level that corresponds to a given score.

Holistic scoring—Scoring based upon an overall impression (as opposed to traditional assessment scoring which counts up specific errors and subtracts points on the basis of them). In holistic scoring the rater matches his or her overall impression to the point scale to see how the portfolio product or performance should be scored. Raters usually are directed to pay attention to particular aspects of a performance in assigning the overall score.

Item—an individual question or exercise in an assessment or evaluative instrument.

Norm-referenced assessment—an objective assessment that is standardized on a group of individuals whose performance is evaluated in relation to the performance of others; contrasted with criterion-referenced assessment.

Opportunity-to-learn standards—suggest criteria for assessing whether schools are giving students the opportunity to learn material reflected in the content standards. Such criteria may include the availability of instructional materials or whether teachers are prepared to teach according to the content standards.

Percentile—the percent of people in the norming sample whose scores were below a given score.

Percent score—the percent of items that are answered correctly.

Performance assessment—performance assessments require students to generate rather than choose a response. Performance assessment by any name requires students to actively accomplish complex and significant tasks, while bringing to bear prior knowledge, recent learning, and relevant skills to solve realistic or authentic problems. Exhibitions, investigations, demonstrations, written or oral responses, journals, and portfolios are examples of the assessment alternatives we think of when we use the term alternative assessment.

Performance criteria—a predetermined list of observable standards used to rate performance assessments. Effective performance criteria include considerations for validity and reliability.

Performance standards—Answers the question “How good is good enough?” Establishes the degree or quality of students’ performance in the subject matter set out by the content standards. Performance standards are often illustrated by exemplars of student work and may contain gradations of performance, i.e., needs improvement, meets standard, above standard.

Portfolio—A collection of representative student work over a period of time. A portfolio often documents a student’s best work, and may include a variety of other kinds of process information, e.g., drafts of student work, students’ self-assessment of their work, parents’ assessments. Portfolios may be used for evaluation of a student’s abilities and improvement.

Prompt—an assignment or directions asking the student to undertake a task or series of tasks. A prompt presents the context of the situation, the problem or problems to be solved, and criteria or standards by which students will be evaluated.

Rating scales—a written list of performance criteria associated with a particular activity or product which an observer or rater uses to assess the pupil’s performance on each criterion in terms of its quality.

Raw score—the number of items that are answered correctly.

Reliability—the extent to which an assessment is dependable, stable, and consistent when administered to the same individuals on different occasions.

Technically, this is a statistical term that defines the extent to which errors of measurement are absent from a measurement instrument.

Rubric—a set of guidelines for giving scores. A typical rubric states all the dimensions being assessed, contains a scale, and helps the rater place the given work properly on the scale.

Standardized assessment—a form of measurement that has been normed against a specific population. Standardization is obtained by administering the assessment to a given population and then calculating means, standard deviations, standardized scores, and percentiles. Equivalent scores are then produced for comparisons of an individual score to the norm group's performance.

Standard scores—a score that is expressed as a deviation from a population mean.

Stanine—one of the steps in a nine-point scale of standard scores.

Task—a goal-directed assessment activity, demanding that students use their background knowledge and skills in a continuous way to solve a complex problem or question.

Validity—the extent to which an assessment measures what it was intended to measure. Validity indicates the degree of accuracy of either predictions or inferences based upon an assessment score.