NCME Instructional Module on

Design and Development of Performance Assessments

Richard J. Stiggins
Northwest Regional Educational Laboratory

Achievement can be, and often is, measured by means of observation and professional judgment. This form of measurement is called performance assessment. Developers of large-scale assessments of communication skills often rely on performance assessments in which carefully devised exercises elicit performance that is observed and judged by trained raters. Teachers also rely heavily on day-to-day observation and judgment. Like other tests, quality performance assessment must be carefully planned and developed to conform to specific rules of test design. This module presents and illustrates those rules in the form of a step-by-step strategy for designing such assessments, through the specification of (a) reason(s) for assessment, (b) type of performance to be evaluated, (c) exercises that will elicit performance, and (d) systematic rating procedures. General guidelines are presented for maximizing the reliability, validity, and economy of performance assessments.

An Overview of Performance Assessment

When we think of measuring achievement, the natural tendency is to think of published standardized, multiple-choice, true/false, or fill-in tests, all of which rely on written test items that are read and answered correctly or incorrectly by the examinee. In fact, however, recent studies of school assessment suggest that teachers rely at least as much on observation and judgment in evaluating student achievement as they do on paper-and-pencil assessment strategies. Teachers observe and evaluate such behaviors as oral reading fluency or speaking skills and such products as writing samples or art and craft products as they track student development. At some grade levels (especially primary grades) and in some subject-matter areas (e.g., language arts), teachers rely very heavily on measurement that is based on observation and professional judgment. In this module, measurement based on observation and judgment is labeled performance assessment.

Large-scale standardized tests often rely on selection-type paper-and-pencil test items because such test items sample content effectively and can be scored quickly. Efficient assessment is crucial when there are many thousands of tests to be scored. However, even in this context we see increasing use of performance assessment, particularly in instances where observation and judgment represent the most valid way to assess. For example, nearly three quarters of the states are conducting statewide writing assessments based on a teacher's subjective evaluation of student writing samples. Trained raters evaluate overall quality, organization, style, content, and other key factors by applying clearly articulated performance standards in the process of evaluation. When used carefully, performance assessment can produce dependable results.

Whether that use is in the classroom on a daily basis, in an annual statewide testing program, or in evaluation for professional licensing and certification, performance assessments rely on the judgmental rating of achievement. Such rating processes are subject to a variety of measurement errors and must be conducted very carefully. The keys to success—to obtaining valid and reliable results—are (a) to make the judgment-based evaluation process as systematic and objective as it can be while (b) focusing on the most important attributes of performance. Systematic assessments have a clear purpose, are based on explicit criteria, rely on appropriate exercises, and include precise performance rating procedures. Our goal is to be sure that performance ratings reflect the examinee's true capabilities and are *not* a function of the perceptions and biases of the person evaluating performance.

Richard J. Stiggins is Director of the Center for Performance Assessment, Northwest Regional Educational Laboratory, 101 SW Main St., Suite 500, Portland, OR 97204. His specializations are performance assessment methodology and classroom assessment.

Funds for development of this unit were provided in part by the Office of Educational Research and Improvement (OERI), U.S. Department of Education. The opinions expressed in this module do not necessarily reflect the position of OERI, and no official endorsement by OERI should be inferred. The module represents an adaptation of a guide entitled "Evaluating Students by Classroom Observation" published in 1982 by the National Education Association Professional Library, Washington, DC. Published by permission of the National Education Association.

Series Information

This module is written to help the reader reach this goal through the use of practical, economically feasible, efficient performance assessments.

The Components of a Performance Test

Performance assessments call upon the examinee to demonstrate specific skills and competencies, that is, to apply the skills and knowledge they have mastered. The demonstration can take place during the normal course of everyday events (e.g., during normal classroom life) or in response to specific structured exercises provided by the examiners. Regardless, the examinee's task is to construct an original response, which the examiner observes and evaluates.

Performance assessments are valuable tools for measuring communication skills such as reading, writing, speaking, and listening. They serve well in business and industrial education, science lab, visual and performing arts, and physical education. They represent valuable tools in special education contexts where examinees lack paper-and-pencil test-taking skills and/or where paper-and-pencil tests fail to test important skills. They also play key roles in assessment of professional skills.

All performance assessments are composed of four basic components: a *reason* for assessment, a particular *performance* to be evaluated, *exercises* to elicit that performance, and systematic *rating* procedures. The reason for assessment is defined in terms of decisions to be made, decisionmakers, and examinees to be evaluated. Performance to be evaluated is defined in terms of content and/or skills, type of behavior or product to be observed, and performance criteria. Exercises are structured or natural, preannounced or unannounced, and variable in number. Finally, rating procedures are defined in terms of the type of data needed, identity of the rater, and nature of the recording method. Those who design and develop performance assessments must consider all of these factors.

The goal of this training module is to spell out these performance assessment design alternatives and teach the user how to select from among available options to develop performance tests that fit particular purposes.

Performance Assessment in the Larger Scheme

Educators and assessment specialists use a wide variety of tools to evaluate achievement. They use formal and informal performance assessments, objective tests and quizzes, essay tests, and oral questions in class, to mention a few. How do performance assessments compare with the other alternatives in terms of active ingredients and keys to successful use? Consider the differences spelled out in Table 1. Each form of assessment is different. Each is a valuable tool in the hands of a skilled user. Each can contribute to an effective assessment.

Purpose of This Module

Many measurement textbooks deal effectively with objective tests (whether standardized or teacher-developed) and essay tests. Since oral questions are often like short-answer test items in form and evaluation, some measurement texts may help with these also. However, fewer treat performance assessment in a systematic manner. This module is designed to help fill that void.

Performance assessments can vary in their formality. The most formal of these assessments are structured, preplanned events designed in advance to provide a decisionmaker with a specific piece of performance information. For instance, a remedial reading teacher might listen to each student read orally early in the year to help make placement decisions. Informal assessments, on the other hand, are those casual, spontaneous insights that often come to the teacher as a result of an incident noticed during an instructional activity. For example, a third grade teacher might notice a particular student stumbling repeatedly over a particular beginning sound in a reading group. The teacher might make a mental note for later action. The rules of good evidence and sound assessment need to be applied to both formal and informal performance assessment, if each is to produce useful information. However, the evaluator often has little control over the circumstances surrounding informal events. For this reason, our purpose is to concentrate on performance assessment in the formal case through the remainder of this module. Extrapolations from the formal to the informal case will be highlighted as we go.

Structure of This Module

The presentation that follows is designed to take you through the step-by-step process of designing a blueprint for a performance assessment. The process is broken down into four steps, one for each of the major components specified earlier. You will specify a reason for assessment, describe the performance to be evaluated, plan exercises, and outline rating procedures. In all, you will make 13 different design decisions. In each case, you will learn the design alternatives available to you and the factors to consider in making your choices among alternatives. When you have completed all choices, you will have a very detailed picture of a useful assessment tool.

This test design simulation is intended to inform you about each of the active ingredients in a sound performance assessment. We do not advocate the systematic completion of each step every time a new professional judgment is made. However, the more important the purpose for the assessment the more crucial it is that we obtain valid and reliable data from the assessment and, therefore, the more crucial it is that all active ingredients be carefully considered.

Assessment Design

As we begin the process of performance assessment design, please reflect on the educational environment around you and identify a particular situation in which a performance assessment might be (or has been) useful to you. For teachers, this might be a reading, writing, science, or arts context. For administrators, it might be a program evaluation or a personnel evaluation situation. For those involved in professional assessment, it might be a work sample–based performance review. Select a situation requiring a systematic, evaluation-based professional judgment. If possible, make it a real, current need. This will make the simulation most meaningful and useful to you.

When you have identified such a situation, follow the Performance Assessment Blueprint, which appears on page 40, as you proceed through this instructional guide, beginning the design process with Step 1, "Clarify reason(s) for assessment." To illustrate the test design process, we will accom-

pany each step with the detailed development of an example of a performance assessment of writing skills.

Step 1: Clarify Reason(s) for Assessment

No assessment should ever be conducted unless and until the evaluator knows exactly how the results are to be used. As you will see, the manner in which results are to be used influences many of the subsequent design decisions. Therefore, your first step is to state why you are designing this assessment.

A. Specify decision(s) to be made on the basis of assessment results. Some of the alternatives are listed below with brief descriptions. Choose from among these or specify another that fits your context and enter it in the space provided on the Blueprint for Item 1A. You may identify more than one decision if appropriate.

- Individual diagnosis—identifying the strengths and weaknesses in the performance of individuals.
- Group needs assessment—diagnosing the strengths and weaknesses of a group of examinees, such as a particular class.
- Grading—assigning a grade (A, B, C, etc.) to individual

TABLE 1

Comparison of Various Types of Assessment

	Objective test	Essay test	Oral question	Performance assessment
Purpose	Sample knowledge with maximum efficiency and reliability	Assess thinking skills and/or mastery of a structure of knowledge	Assess knowledge during instruction	Assess ability to translate knowledge and understanding into action
Typical exercise	Test items: Multiple-choice True/false Fill-in Matching	Writing task	Open-ended question	Written prompt or natural event framing the kind of performance required
Student's response	Read, evaluate, select	Organize, compose	Oral answer	Plan, construct, and deliver original response
Scoring	Count correct answers	Judge understanding	Determine correctness of answer	Check attributes present, rate proficiency demonstrated, or describe performance via anecdote
Major advantage	Efficiency—can administer many items per unit of testing time	Can measure complex cognitive outcomes	Joins assessment and instruction	Provides rich evidence of performance skills
Potential sources of inaccurate assessment	Poorly written items, overemphasis on recall of facts, poor test-taking skills, failure to sample content representatively	Poorly written exercises, writing skill confounded with knowledge of content, poor scoring procedures	Poor questions, students' lack of willingness to respond, too few questions	Poor exercises, too few samples of performance, vague criteria, poor rating procedures, poor test conditions
Influence on learning	Overemphasis on recall encourages memorization; can encourage thinking skills if properly constructed	Encourages thinking and development of writing skills	Stimulates participation in instruction, provides teacher immediate feedback on effectiveness of teaching	Emphasizes use of available skill and knowledge in relevant problem contexts
Keys to success	Clear test blueprint or specifications that match instruction, skill in item writing, time to write items	Carefully prepared writing exercises, preparation of model answers, time to read and score	Clear questions, representative sample of questions to each student, adequate time provided for student response	Carefully prepared performance exercises; clear performance expectations; careful, thoughtful rating; time to rate performance

students as a reflection of the amount of material learned.
- Grouping—assigning examinees to specific instructional groupings within a class.
- Selection—identifying those who are far ahead of or behind the rest for placement into special advanced or remedial programs.
- Certification—verifying the mastery of specified skills by individuals.
- Evaluation—determining if a particular program of instruction has worked effectively or needs revision.

Reflect on your context and decide which decision(s) best describe(s) your intended use of the results. If you have a purpose in mind that does not appear above, write it in the space provided.

Example: We will conduct a writing assessment to diagnose strengths and weaknesses in individual students' writing skills.

B. Specify decisionmaker(s) by identifying the person(s) who will use the results to make the decision(s) listed under 1A. This may be a teacher, administrator, student, parent, board member, employer, certification board, or some combination of these. It may or may not be the person who actually rates performance. It is the person who acts upon performance ratings to make decisions that influence how examinees and programs interact.

Example: The purpose for our writing assessment is to help the teacher and student identify weaknesses in the student's writing skills.

In the space provided for 1B, specify the decisionmaker(s) who will use your assessment results.

C. Specify use to be made of results. You have two choices. All educational decisions require that we either

- rank examinees—place in order of achievement from lowest to highest; or
- determine mastery—verify that each performer has mastered material regardless of how others perform.

You may also use these in combination as in ranking students and exploring the skills not mastered by those who rank low.

The key factor to consider in making your selection is the nature of the decision to be made. Some decisions, such as grouping and selection, require rank order data, whereas others, like diagnosing strengths and weaknesses, require information about specific skills mastered and not mastered.

Example: In our writing assessment, we will determine mastery of specific writing skills.

Which do you need? Make your choice in space 1C of the Blueprint.

D. Describe students to be assessed—Describe the examinees whose performance is to be evaluated by specifying how many will participate, at what grade level, and any salient characteristics of those examinees. Salient characteristics include those factors that might need special attention in assessment design. Are the students advanced or remedial? Participants in a special new program? If they are unique in some special way, note that fact here. This information will be useful later as we try to design assessments that are feasible in terms of examinees and time required for assessment. Fewer examinees allow more time per assessment; more permit less time per assessment. Our goal is to design assessments that are as economical as they can be and to make them realistic.

Example: We will assess the skills of 100 11th grade language arts students.

Describe your performers under Blueprint entry 1D.

Step 2: Clarify Performance to Be Evaluated

An obvious key to quality assessment is the clear definition of the performance to be evaluated. We need to specify the general content area, the type of performance to be observed, and the specific dimensions of performance to be considered in evaluation.

A. Specify the content or skill focus of the assessment, in a very general way. What subject matter area and class of proficiencies are to be demonstrated?

Example: We will assess the *English composition* skills of high school students.

Place your answer in Item 2A of your Blueprint.

B. Select the type of performance to be evaluated. You can observe and evaluate a process of behavior as it occurs, as when the teacher listens to and evaluates oral reading fluency. Or you can observe and evaluate a product developed by the performer, as when he or she creates a woodworking project. You also can use both, as when a typing teacher evaluates both typing technique and the finished product.

The key question is: Where can you find the best evidence of proficiency? Is the process important, as in speaking assessment? Must steps be carried out in a specified order, as in carrying out a science experiment? Or does a tangible product result which is to have certain characteristics?

Example: Our writing assessment will be based on a product—actual samples of student writing.

On what will you base your judgment? Check the appropriate indicator on Blueprint Item 2B.

C. List performance criteria. **No other single specification will contribute more to the quality of your performance assessment than this one.** Before the assessment is conducted, you must state the performance criteria, in other words, the dimensions of examinee performance (observable behaviors or attributes of products) you will consider in rating.

Each dimension of performance is to be specified in two parts: a definition and a performance continuum. Performance criteria should reflect those important skills that are the focus of instruction. Definitions spell out what we, as the evaluators, mean by each particular criterion. The continuum specifies the range of possible ratings from high to low for each criterion—*in observable terms* (see the example below).

The key to identifying sound performance criteria is to place yourself in the hypothetical situation of having to give feedback to someone who has performed poorly on the task. What factors are you likely to mention? The answer to this question requires and deserves a great deal of thought on

the part of the evaluator. **If you do not have a clear sense of the key dimensions of sound performance—a vision of poor and outstanding performance—you can neither teach students to perform nor evaluate their performance.**

Example: Key dimensions of writing performance to be evaluated in our writing assessment

Factor	Meaning	Continuum
Organization	Arrangements follow in an effective sequence	Beginning, middle, and end are clear and effective, *to* very haphazard arrangement, hard to follow
Voice	Extent to which the writer speaks to the reader	Writer speaks directly to the reader and seems sincere, *to* writing is flat, lifeless, with no sense of the individual writer
Writing mechanics	Grammar, usage, punctuation, etc.	Good grasp, no glaring errors, not a factor in reading paper, *to* too many errors make paper hard to read and understand

In the table provided in the Blueprint under 2C, space is provided for up to three performance criteria. In fact, you may specify more or fewer, depending on your assessment. But be sure to specify all relevant criteria and expand your initial sketch as needed to ensure clear understanding of performance expectations by both raters and performers (see the Appendix for more detail on criteria listed above).

Step 3: Design Exercises

Performance assessment exercises present those instructions or describe those events that give rise to examinee performance. Once the performance to be evaluated is defined, we must decide how to elicit or sample performance so as to observe and evaluate it. We need to specify the form, obtrusiveness, and number of exercises to be used.

A. Select form of exercises. Specify if you plan to (a) observe everyday events as they occur, such as by watching examinees during everyday discussion to evaluate speaking skills, or (b) design specific exercises to cause them to speak on a standard topic, for example, under standard, controlled circumstances so you can evaluate performance. Please note that you can use both in gathering evidence of proficiency.

Consider two factors in making this selection: the natural availability of dependable evidence and the seriousness of the decision to be made. Evidence available through natural observation can be relatively inexpensive, as these events will unfold whether or not you use them as a source of performance ratings. So, in effect, the design of the exercise has no cost. When natural circumstances provide sufficient samples of the kind of behavior you wish to observe, those circumstances provide the basis for a very efficient, low-cost performance assessment. Just remember, however, that frequency of occurrence of desired behavior is a factor in selecting this option. If the desired performance is likely to be infrequent, you can waste a great deal of time in unproductive observation. In this case, structured exercises may be needed to elicit evidence of proficiency.

Structured exercises also may be needed when very important decisions hang in the balance, such as grade promotions, high school graduation, college scholarships, or professional certification. In these cases, it is crucial that every examinee have a fair and equal opportunity to demonstrate his or her skill. Under such circumstances, a carefully planned, standardized, and tightly controlled assessment is essential.

Example: Sample writing exercise: Assume you have a friend who is moving to another city to find a job. You know someone there who may be able to hire your friend. Write a letter of introduction for your friend to take to the new city. Describe your friend in a way that will make the reader feel he or she would like to meet and interview this person.

Make your choice under Blueprint 3A now, providing a sample exercise, if you select that option, and/or describing the relevant natural events to be observed. In your sample exercise, provide a complete set of instructions to the examinees letting them know what they are to do and under what conditions. If you choose naturally occurring events, describe what the performer will be doing and under what circumstances.

B. Determine the obtrusiveness of assessment by specifying if it will be preannounced, open, and public assessment or if you plan to conduct your assessment without advising performers that an assessment is taking place. Again, you may wish to use a combination, gathering evidence of proficiency under both obtrusive and unobtrusive circumstances.

The key difference between the two is seen in examinee motivation and level of test anxiety. Under standard testing circumstances, performers will operate at levels of maximum motivation—demonstrating their best possible work. However, in a few situations (such as with safety rules in a science lab) it may be more helpful to know how they perform under conditions of typical everyday motivation. Unobtrusive—but systematic—observations can provide that information.

Under normal testing conditions, examinees often experience varying degrees of test anxiety. In moderation, this anxiety can create a keen performance edge that maximizes performance. On rare occasions, however, some people experience debilitating test anxiety. At test time, capable but extremely anxious performers fail to demonstrate skills you know they have. When this occurs, unobtrusive observations can disarm the test situation, eliminate the anxiety, and provide needed evidence of proficiency.

Example: Our writing assessment will be announced in advance.

Specify your choice under Blueprint 3B now.

C. Determine the amount of evidence you plan to gather. You have three choices: one sample of performance gathered at one time, multiple samples at one time, or multiple samples collected over several occasions.

In evaluating these design options, consider three factors:

- The importance of the decision—the more important the decision, the surer you have to be and therefore the more evidence you need to gather.
- Representativeness of samples—any assessment reflects a sample of all of the possible examples of performance that could have been gathered. Gather enough samples of performance to be sure you have sampled the range of possible applications of the skills to be evaluated.
- Time—consider the amount of time you have available to rate each sample, the number of performers to be evalu-

ated, and the time until the decision must be made. This will indicate how many samples you can realistically gather and evaluate.

There are no simple rules as to how many samples are enough. The number varies greatly from context to context. Rely on your own judgment. How many will you need to be relatively sure you can accurately judge an examinee's true capabilities?

Example: Students in our writing assessment will provide three 30-minute samples of writing. They will produce one per day for 3 days.

Specify your sampling plan under Blueprint 3C.

Step 4: Design Performance Rating Plan

To complete the performance assessment blueprint, you must plan how performance will be "scored" and how results can be put in a communicable form. The performance rating plan must be designed to give the decisionmaker(s) the assessment results they need in a timely fashion and in a form they can use. Therefore, the key factor to consider in designing this component of the assessment is the nature of the decision to be made. To complete the performance rating plan, answer these questions: What type of score is needed? Who is to rate performance? How are the data to be recorded?

A. Determine the type of score needed by reflecting on the information needs of the decisionmaker. Some decisions, such as grouping and selection, can be based on a ranking of examinees (see Blueprint 1C) and therefore require only a general overall index of performance—a holistic rating, if you will. Other decisions, such as diagnosing individual or group needs and certifying minimum competencies, require a more detailed breakdown of dimensions present or absent in student performance. These require analytic scoring.

Please note that both holistic and analytic scores are based on the performance criteria spelled out under Blueprint 2C. But the criteria are used differently in each case. In analytic scoring, each criterion is evaluated individually and is assigned a score. In holistic scoring, all criteria are considered simultaneously when assigning an overall score. Some criteria may be given more weight than others in this process, depending on the contribution of each to the quality of performance in the evaluator's opinion.

The advantage of holistic scoring is that it is less time-consuming and therefore less expensive than analytic scoring. It takes less time to make one overall judgment than it does to make several individual judgments. When holistic will suffice in terms of information needed, it is definitely the score of choice. Analytic, on the other hand, offers more focused information. It offers a profile of performance characteristics. The price we pay, of course, is the greater time required to make several evaluations—one for each criterion. But when we need the diagnostic information, analytic scoring is worth the price.

Example: In our writing assessment, we will score all papers analytically, because we wish to determine the specific weaknesses of the students.

Specify the type of score you need for your assessment in Blueprint Item 4A now. Remember, you can use them in combination, as when you score all performance samples holistically and then analyze the specific weaknesses of those scoring low in an overall sense.

B. Determine who is to rate performance. In this case, as in others, you have several choices. You can rely on the expert professional judgment of a teacher or other qualified specialist. Or you can have the performers evaluate their own and/or each other's performance. Which you choose depends on your assessment context.

Use teacher or other professional judgment if

- rating requires the specialized, technical knowledge and experience that only a professional possesses;
- important competitive decisions require standard, uniform test conditions and unbiased ratings for all examinees; and/or
- examinees have a vested interest in results and may be perceived as having a chance to benefit unfairly from self-evaluation (e.g., a grade hangs in the balance).

Consider examinee peer or self-rating if

- examinees are capable of learning and applying the performance criteria;
- slight variation from rater to rater is acceptable in your context;
- examinees have nothing to gain from artificially inflating or deflating their ratings; and/or
- the evaluator's scoring task exceeds time available to complete it—examinees sometimes represent a low-cost, efficient scoring resource.

Using more than one rater is often an excellent idea. By using two evaluators, we can gather evidence on the extent to which independent ratings of the same examinee's performance agree. Such agreement argues that criteria have been dependably applied and rater bias has been controlled. On the other hand, if independent raters disagree, we may need to reconsider criteria and/or train raters more thoroughly.

Remember that *all* raters must be trained to rate examinee performance—whether raters are experts or students. Just because experts are selected from the same field, there is no guarantee that they view key dimensions of achievement alike. They must be trained to agree by practicing with the criteria on samples of performance.

Remember also that training students to be raters represents an excellent instructional (and assessment) strategy. Once they internalize performance criteria and see how those criteria come into play in their own and each other's performance, students often become better performers.

Example: In our writing assessment, students and the teacher will rate performance. Each student will rate his or her own paper plus the writing samples of one other student. The teacher will independently rate all papers.

Specify the raters you will use in your assessment under Blueprint entry 4B now.

C. Clarify score recording method to be used. This is the final design decision and represents the place where we merge the performance criteria into the scoring process. Recording alternatives include:

- a *checklist* of specific attributes present or absent in the performance, such as characteristics of an art object;

- a *rating* scale in which dimensions of performance are evaluated along a continuum from adequate to inadequate, such as rating key aspects of a speech;
- an *anecdotal* record in which important behaviors are described in a verbal form that communicates evaluative judgments, as in anecdotal records of social/emotional problems; and,
- a *portfolio* of examples of examinee performance selected to illustrate the level of skill and/or development over time.

Please note that these can be used in combination to create a very thorough record of performance, such as by developing a portfolio of work samples, each of which has a rating scale attached.

Rating scales and checklists have distinct advantages. For instance, they

- combine the observation and judgment into an easy-to-interpret record;
- apply equally well to process and product evaluations;
- can be tailored to many types of performance criteria;
- provide a convenient frame of reference for summarizing and comparing assessment results; and
- permit a quick, efficient means of recording data.

Rating scales and checklists are most efficient, and therefore most useful, when many examinees and/or many performance criteria are involved.

Anecdotal records and portfolios are more cumbersome but provide a level of detail and richness of information not achievable with rating scales. They also allow the quality of behavior or products to be described with examples, for instance, as performance changes over time. These are most useful when the assessment involves a few examinees and a limited range of traits. Please note that video and audio tape can also serve as a basis for recording examples of achievement for inclusion in an anecdotal record or portfolio.

One final recordkeeping system that is often used, and almost always is ill-advised, is mental recordkeeping. When we attempt to remember records of achievement, five different things can happen, and four of them are bad. The one good (but unlikely) outcome can be that we maintain a vivid, accurate recollection of the performance. On the downside, we can

- forget and lose valuable information;
- ascribe the recollection to the wrong examinee;
- change the mental record as a result of observing subsequent performance; and,
- most dangerous of all, filter all subsequent observations through the recollection of that initial event, biasing results.

Written, audio tape, or even video tape recorded recordkeeping is advised.

Select your recording strategy on Blueprint entry 4C, then develop your checklist or rating scale, if selected. The checklist should include all key attributes of performance. The rating scale should include all performance criteria and a descriptive scale or continuum with each point described in specific observable terms. Scales relying on general labels such as "poor" to "excellent"will not suffice. Poor performance needs to be defined in concrete performance terms, if we are to provide valid, dependable, and meaningful feedback to performers. See examples of a rating scale for writing assessment in the Appendix.

Ensuring High Quality

Professional observations and carefully considered judgments represent a viable means of measuring achievement. They can and do serve us very well in the day-to-day management of instruction. They also play a key role in large-scale assessment programs—especially in writing assessment. If they are to fulfill their potential in these contexts, the performance rating results must be dependable. They can be dependable only if performance assessment developers and users adhere to basic rules of test design. They must

1. be clear on the *purpose* of the assessment. Without knowledge of the decision(s) to be made and decisionmaker(s), we cannot plan useful, economical assessments.
2. *communicate effectively,* with (a) performance criteria conveyed in an understandable way to students *prior to the assessment* and (b) challenging performance exercises.
3. maximize the *validity* of the assessment by (a) being sure about the purpose; (b) defining the student characteristics to be evaluated; (c) specifying levels of performance along appropriate continuums; (d) using exercises that sample the range of performance contexts; and (e) comparing ratings with other achievement data when possible.
4. maximize the *reliability* of assessment by (a) using clear criteria; (b) training raters thoroughly; (c) planning and implementing appropriate scoring procedures; (d) gathering enough samples of performance; (e) minimizing rater bias through cultural awareness, clear criteria, and thorough training; and (f) providing for a standard, uniform assessment condition when needed.
5. attend to the *economy* of assessment by (a) adapting the form of the assessment to the purpose; (b) gathering only as many samples of performance as the decision requires; (c) reusing good exercises when possible; (d) rating in terms of the relevant criteria only; (e) training raters to be efficient and to reduce the need to reevaluate performances; (f) using rating scales, checklists, and work samples (portfolios) when appropriate; and (g) relying on students as raters when possible and appropriate.

Adherence to these guidelines will result in high-quality performance assessment results.

Teaching Aids Are Available

A set of teaching aids, designed by Richard J. Stiggins to complement his ITEMS module of "Design and Development of Performance Assessment," is available at cost from NCME. These teaching aids consist of brief additional text and 19 overhead masters. As long as they are available, they can be obtained by sending $1.00 to: **Teaching Aids, ITEMS Module #1, NCME, 1230 17th St., NW, Washington, DC 20036.**

PERFORMANCE ASSESSMENT BLUEPRINT

Step 1: Clarify reason(s) for assessment

A. Specify decision(s) to be made _____

B. Specify decisionmaker(s) _____

C. Specify use to be made of results

____ Mastery ____ Rank order ____ Combination

D. Students to be assessed

How many? ____ What grade level? ____

Unique characteristics? _____

Step 2: Clarify performance to be evaluated

A. Specify the content or skill focus of the assessment _____

B. Select type of performance to be evaluated

____ Process; identify process _____

____ Product; identify product _____

____ Combination; identify each in spaces provided above

C. List performance criteria

Factor	Meaning	Continuum

Step 3: Design exercises

A. Select form of exercises

____ Structured exercises; devise sample exercise _____

____ Natural events; describe events to be observed _____

____ Combination; devise sample exercise and describe events in spaces provided above.

B. Determine obtrusiveness of assessment

____ Examinees aware of assessment ____ Examinees unaware of assessment ____ Combination

C. Determine amount of evidence needed

____ 1 sample, 1 time ____ Multiple samples, 1 time ____ Multiple samples over time

Step 4: Design performance rating plan

A. Determine type of score needed

____ Holistic ____ Analytic ____ Combination

B. Determine who is to rate performance

____ Teacher ____ Self-rating ____ Combination

____ Other expert ____ Peer-rating

C. Clarify score recording method

____ Checklist ____ Anecdotal record ____ Combination (specify)

____ Rating scale ____ Portfolio _____

Self-Test

Instructions: Use these exercises for additional practice or to assess your proficiency in performance assessment design. As you do so, please realize that there is more than one appropriate design. Any of a variety of assessments might serve the purpose. In addition, as you design the assessments, address issues of performance criteria within the limits of your background knowledge.

1. For the sake of this exercise, assume you have been selected to design and carry out a science fair in your school. Students are to develop and present science projects. Developers of the best projects are to be awarded scholarships. Complete a performance assessment blueprint, filling in details where needed, to reflect how you would conduct the assessment.

2. If you were to design a performance assessment to evaluate whether someone could drive an automobile on public streets and highways, what would the blueprint look like?

3. Primary level teachers (grades 1, 2, and 3) often need to identify students who may require special help in the development of their reading skills. Since these students have not yet developed traditional test-taking skills, paper-and-pencil tests cannot be used. Using performance assessment, how might primary teachers identify these students?

4. Assume that you are a high school chemistry teacher whose objective is to determine if safety rules are being adhered to in the lab. Because the school has had a history of accidents, your supervisor has asked for proof that your classes are being conducted safely. Design a performance assessment to provide needed data.

5. Teacher evaluation is often based, at least in part, on the observation and judgment of teacher behavior in the classroom. Assume you are a principal whose assignment is to evaluate a first-year teacher to be sure that that teacher is competent in terms of state-specified minimum competencies. Your job will be to decide to retain or release this teacher at the end of the year. How would you design and conduct this assessment?

Additional Readings

Berk, R. A. (Ed.). (1986). *Performance assessment: Methods and applications.* Baltimore, MD: The Johns Hopkins University Press.

This is the most comprehensive treatment of performance assessment methodology completed to date. Nineteen chapters address methods and methodological issues, as well as applications in business and education. Of particular interest to educators are chapters on teacher evaluation, student evaluation, writing assessment, and listening and speaking assessment.

Fitzpatrick, R., & Morrison, E. J. (1971). Performance and product evaluation. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 237–270). Washington, DC: American Council on Education.

This discussion of performance assessment addresses the topic from some unique perspectives. For example, the authors give a great deal of attention to the concept of simulation, the application of high technology in performance assessment, the basic concepts of fidelity, cost, and the essentials of good simulations. In addition, they emphasize the use of situation tests such as in-basket tests, games, contests, and diagnostic problem-solving tests in which examinees engage in some real-life tasks. Finally, the authors provide guidelines for developing performance tests, covering test specifications, stimulus conditions, response modes, conditions for appropriate test administration and scoring, and guidelines for test use.

Haynes, S. N., & Wilson, C. C. (1979). *Behavioral assessment.* San Francisco, CA: Jossey-Bass, 526 pp.

Behavioral assessment—based on the observation of actual behavior—is as important to the clinical psychologist as performance assessment is to the classroom teacher. Haynes and Wilson illustrate this by referencing more than 70 journal articles addressing behavioral assessment in educational contexts. Educators will find the description of developments in behavioral assessment and their educational applications both interesting and useful. The publication also includes reference to current research literature on such recent methodological advances as behavioral coding systems, strategies for assessing interobserver agreement, and advances in participant observation; observation in natural environments, including schools; observations of child behavior in structured learning environments; and use of self-monitoring, behavioral questionnaires and behavioral interviews.

Priestley, M. (1982). *Performance assessment in education and training: Alternative techniques.* Englewood Cliffs, NJ: Educational Technology Publications, 263 pp.

This is a basic introductory textbook on performance assessment. In it, Priestley describes 25 different types of performance assessments in terms of their form, uses, advantages, disadvantages, and, most important, steps in test development. He includes many concrete illustrations of assessments and addresses keys to successful administration and coding. Among the illustrations are actual performance tests including work samples, identification tests, and supervisor, peer, and self-ratings. Simulations including job simulations, written simulations, and management simulations are described. Observational assessments, such as checklists, rating scales, and anecdotal records are summarized as are oral assessments, paper-and-pencil assessments, and personnel records.

APPENDIX

EXAMPLES OF ANALYTICAL RATING GUIDES

These scoring guides are adapted from scales developed by Carol Meyer and Vicki Spandel, Beaverton Public Schools, Beaverton, OR.

Organization

5 The writer organizes material in a way that enhances the reader's understanding or that helps to develop a central idea or theme. The order may be conventional or not, but the sequence is effective and moves the reader through the paper.

- Details seem to fit where they're placed, and the reader is not left with the sense that "something is missing."
- The writer provides a clear sense of beginning and ending, with an inviting introduction and a satisfying conclusion ("satisfying"

in the sense that the reader feels the paper has ended at the right spot).
- Transitions work well; the writing shows unity and cohesion, both within paragraphs and as a whole.
- Organization flows so smoothly that the reader doesn't have to think about it.

3 The writer attempts to organize ideas and details cohesively, but the resulting pattern may be somewhat unclear, ineffective, or awkward. Although the reader can generally follow what's being said, the organizational structure may seem at times to be forced, obvious, incomplete, or ineffective.
- The writer seems to have a sense of beginning and ending, but the introduction and/or conclusion tend to be less effective than desired.
- The order may not be a graceful fit with the topic (e.g., a forced conventional pattern, or lack of structure).
- The writer may miss some opportunities for transitions, requiring the reader to make assumptions or inferences.
- Placement or relevance of some details may be questionable (e.g., interruptive information; writer gets to the point in roundabout fashion).
- While some portion of the paper may seem unified (e.g., organization within a given paragraph may be acceptable), cohesion of the whole may be weak.

1 Organization is haphazard and disjointed. The writing shows little or no sense of progression or direction. Examples, details, or events seem unrelated to any central idea, or may be strung together helter-skelter with no apparent pattern.
- There is no clear sense of a beginning or ending.
- Transitions are very weak or absent altogether.
- Arrangement of details is confusing or illogical.
- There are noticeable information "gaps"; the reader is left dangling or cannot readily see how the writer got from one point to another.
- The paper lacks unity and solidarity.

Voice

5 The paper bears the unmistakable stamp of the individual writer. The writer speaks directly to the reader, and seems sincere, candid, and committed to the topic. The overall effect is individualistic, expressive, and engaging.
- The paper is honest. There is a real effort to communicate, even when it means taking a risk (e.g., an unexpected approach or revealing of self).
- The writing is natural and compelling.
- Tone is appropriate and consistently controlled.
- The writer's own enthusiasm or interest comes through and brings the topic to life.
- The reader feels an interaction with the writer and, through the writing, gains a sense of what the writer is like.

3 The writer makes an honest effort to deal with the topic, but without a strong sense of personal commitment or involvement. The result is often pleasant or acceptable, yet not striking or compelling in a way that draws the reader in.
- Writer may seem self-conscious or unwilling to take a risk—may seem to be writing what he/she thinks the reader wants.
- Paper lacks individuality or the ring of conviction.
- The writing communicates, but only in a routine, predictable fashion that tends to make it blend in with the efforts of others.
- Voice may be inconsistent; it may emerge strongly on occasion, only to shift or even disappear altogether.
- The reader has only an occasional or limited sense of interaction with the writer.

1 The writer may not have understood the assignment or may simply have felt indifferent toward the topic. As a result, no clear voice emerges. The result is flat, lifeless, very mechanical, and stilted, or possibly inappropriate.
- The writing has virtually no individual personality or character; there is no identifiable voice behind the words.
- There is little or no evidence of the writer's involvement in the topic.
- The reader has no sense that this writer was "writing to be read" and experiences virtually no writer-reader interaction.

Writing Conventions

5 The writer has a good grasp of standard writing conventions (grammar, capitalization, punctuation, usage, spelling, paragraphing). There are no glaring errors. In fact, errors tend to be *so* minor that reader can easily overlook them unless searching for them specifically.
- Sentence structure and paragraphing tend to be sound.
- Agreement of subject and verb is correct.
- Punctuation is smooth and enhances meaning. (Informalities—use of dashes, contractions—are allowed.)
- Spelling is generally correct.

3 Errors in writing conventions begin to impair readability. Sentence structure is generally correct on simple sentences, though more complicated patterns may contain such problems as faulty parallelism, inconsistent tense, voice shift (e.g., first to second person), dangling modifiers, or vague pronoun reference.
- Errors may reflect hasty writing or lack of careful attention to detail in editing.
- The reader can follow what's being said overall, but errors in conventions may require the reader to pause or reread on occasion.

1 There are so many errors in usage, sentence structure, spelling, and/or punctuation that the paper is hard to understand.
- The student shows very limited understanding of, or ability to apply, conventions.
- Basic punctuation tends to be omitted, haphazard, or just plain wrong.
- Spelling errors are typically frequent, even in common words.
- Fragments, run-ons, and awkward constructions abound.

Instructor's Guide

for

DESIGN AND DEVELOPMENT OF PERFORMANCE ASSESSMENTS

Richard J. Stiggins
Center for Performance Assessment
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue
Portland, Oregon 97204

January 1987

Funds for development of this unit provided in part by the Office of Educational Research and Improvement (OERI), U.S. Department of Education. The opinions expressed in this publication do not necessarily reflect the position of OERI, and no official endorsement by OERI should be inferred. The material presented herein represents an adaptation of a guide entitled "Evaluating Students by Classroom Observation" published in 1982 by the National Education Association Professional Library, Washington, D.C. Published by permission of the National Education Association.

INSTRUCTOR'S GUIDE

Rationale for the Module

This instructional module deals with a form of measurement that is central to our evaluation of achievement--measurement based on observation and protessional judgment, or performance assessment. One of the important paradoxes in the field of educational measurement, in my opinion, is the fact that performance assessment is so crucial to the effective management of instruction, and yet is so completely disregarded in the measurement training offered to teachers, administrators and assessment specialists. Until recently, professional judgments regarding student achievement have been all but ignored outside the classroom and are often ignored in measurement textbooks used in educator training.

The one skill arena where this pattern has been broken is writing assessment. Over the past decade, multiple-choice tests of language usage, grammar and mechanics have been replaced by writing samples as the primary means of measuring writing skill. This change took place because English teachers across the land felt writing samples represented the most valid way to measure composing skills and they no longer would settle for language usage tests just because someone felt they were less expensive to administer and score. Speaking in a unified voice, the National Council of Teachers of English demanded the most valid assessment and now they have it. Virtually every state-level writing assessment program now bases at least part of its score on a teacher-rated writing sample, and a majority of standardized achievement battery published in the last decade includes a writing sample as a local testing option.

How did this occur? Why is it that teacher judgments gained such credibility in the writing domain? What lessons can we learn from this example about the ingredients that have to be present for a performance assessment to be of high quality and to be perceived as sound? This module explores answers to these questions.

In my opinion, professional judgments are regarded as second-class measurement citizens because those judgments appear to be "intuitions" or unsubstantiated guesses. In many cases, they may be far more. They may be high quality professional judgments based on sound underlying measurement methodology. In other cases, however, they may be based on unsound measurement. In cases where judgments are based on sound methodology, we need to provide evaluators with an assessment framework and means of communication that will make the underlying rigor apparent. In cases where assessments are based on unsound methodology, we need to teach users a framework and means of communication that will raise the quality of the measurement. This training module is designed to serve both of these purposes.

Performance assessments in the area of writing have gained credibility because many feel that writing samples offer the highest fidelity and most valid means of measuring the important skills in that domain. In addition, measurement specialists and teachers have worked hand-in-hand to (1) develop clearly articulated performance criteria and strategies for training teachers to apply them in a rigorous manner, (2) devise writing exercises capable of eliciting useful samples of writing, and (3) plan and conduct scoring procedures that ensure reliable ratings of student performance. This approach can be followed in any subject matter area to devise performance assessments yielding valid, reliable and credible data. This training module shows how this goal can be reached.

Training Options

These training materials are equally effective presented as a group workshop or as an individual study guide. The study guide format takes the user through the step-by-step process of designing a performance assessment. The four major steps are (1) clarifying the reason for assessment, (2) defining the performance to be evaluated, (3) designing exercises, and (4) planning scoring procedures. Within these four steps are 13 test design decisions. Each design decision is spelled out in terms of design alternatives and factors to consider in selecting from among those options. The user who makes all 13 decisions and fills in detail around each has completed a detailed blueprint of a performance test.

The same material also can be presented in a 90 to 120 minute workshop. Transparencies are included here to allow the instructor to take a student group through the same step-by-step design process. Workshop participants can work individually or in small groups to design a performance assessment relevant to them. The instructor begins by establishing the importance of observation and professional judgment as a means of measuring achievement (via discussion or lecture), then turns to the 13 design decisions, spelling each out and allowing participants time to make their choices. After reviewing key quality control guidelines, the instructor can have participants share their blueprints with the class. As an assignment, the instructor can have developers polish and refine their blueprints to hand in for evaluation. Before presenting such a workshop, carefully review the instructional module along with the transparencies to ensure coordination.

Please bear in mind the fact that the design of a performance assessment is not a simple process. There are some complex decisions to be made, such as

that of spelling out clear and relevant performance criteria. The entire design process is likely to look long and complicated to workshop participants the first time through. Be encouraging and let them know it becomes easier with practice.

In addition, be careful not to advocate too much. The idea of this unit is not to teach a rigid framework that must be used in its entirety every time a professional judgment is to be made. Rather, the idea is to make participants aware of the key dimensions of sound performance assessment. Advise them to account for as many as possible each time they devise a performance assessment and to follow the entire blueprint process step by step for the most important assessment.

Intended Audience

This training is relevant for a wide variety of educators and personnel assessment specialists. For example, teachers at all levels and in all subjects use performance assessments. Therefore, this training module fits well in both preservice and inservice teacher training programs in assessment. Principals and district administrators are urged to be instructional leaders. Assessment is a key ingredient of effective instruction. If school administrators are to provide leadership in assessment, they must do so with a clear understanding of performance assessment methodology. Therefore, this training module is relevant for them.

Those who are concerned with professional certification often rely on performance assessment methodology. For instance, personnel evaluation is a performance assessment task. Therefore, the design decisions outlined in this module must be carefully considered in that context. Those who design and

develop professional licensing examinations or who supervise trainees in field internship experiences are also concerned with the valid, reliable and efficient measurement of work-related performance. Training along the lines spelled out in this module will help maximize the quality of these assessments also.

However, for all of these audiences, this module represents only the basic course in performance assessment methodology. Depending on the context and the seriousness of the decisions to be made on the basis of assessment results, the test developer may need to know a great deal more than we can present in this one module (see references below). But regardless of the context, the module represents an appropriate starting point.

Additional Resources

Those who are interested in developing a stronger background in performance assessment, so they are better equipped to teach performance assessment methodology and/or to develop sound assessments, are urged to select readings from the bibliography provided below.

Those particularly interested in historical perspective and the evolution of thought on performance assessment might focus particularly on references by Ryans and Fredericksen (1951), Glaser and Klaus (1962), Fitzpatrick and Morrison (1971) and Priestly (1982). These present detailed frameworks as conceptualized at the beginning of each decade.

Others may be interested in information on educational applications of performance assessment. In this regard, Gronlund (1981), Priestly (1982) and Stiggins (1984) are likely to be helpful. In addition, Stiggins and Bridgeford (1986) explore the role of performance assessments in day-to-day classroom assessment.

Advanced topics of interest may include the use of assessment center methodology (Moses and Byham, 1977), estimation of dependability of performance ratings via generalizability analysis (Brennan and Kane, 1979) and application of performance assessment methodology in the context of behavior therapy (Haynes and Wilson, 1979; and Cimmero, Calhoun and Adams, 1986).

For a general desk reference on the topic of performance assessment as developed and used across educational and work-related contexts, instructors and scholars are urged to refer to the volume edited by Berk (1986).

REFERENCES

Berk, R.A. (Ed.) (1986) Performance Assessment. Baltimore, MD: The Johns Hopkins University Press, 530 pp.

Brennan, R.L. & Kane, M.T. (1979) Generalizability theory: A review. New Directions for Testing and Measurement. San Francisco, CA: Jossey-Bass Publishers, 33-51.

Cimmero, A.R., Calhoun, K.S. & Adams, H.E. (Eds.) (1986) Handbook of Behavioral Assessment (2nd ed.). New York, NY: John Wiley & Sons, 789 pp.

Fitzpatrick, R. & Morrison, E.J. (1971) "Performance and production evaluation." In R.L. Thorndike (Ed.) Educational Measurement. Washington, DC: American Council on Education, 237-270.

Gronlund, N.E. (1981) Evaluating learning and development: Observational techniques. Chapter 16 in N.E. Gronlund Measurement and Evaluation in Teaching (4th ed.). New York, NY: MacMillan, 433-455.

Haynes, N. & Wilson, C. (1979) Behavioral Assessment. San Francisco, CA: Jossey-Bass Publishers, 526 pp.

Moses, J.L. & Byham, W.C. (1977) Applying the Assessment Center Method. New York, NY: Pergamon Press, 310 pp.

Priestly, M. (1982) Performance Assessment in Education and Training: Alternative Techniques. Englewood Cliffs, NJ: Educational Technology Publications, 263 pp.

Ryans, D.G. & Fredericksen, N. (1951) Performance tests of educational achievement. In E.F. Lindquist (Ed.) Educational Measurement. Washington, DC: American Council on Education, 455-493.

Stiggins, R.J. (1984) Evaluating students by classroom observation: Watching students grow. Washington, DC: National Education Association Professional Library, 32 pp.

Stiggins, R.J. & Bridgeford, N.J. (1986) The ecology of classroom assessment. Journal of Educational Measurement, 22(4) 271-286.

WORKSHOP VISUALS

PERFORMANCE ASSESSMENT

An Alternative for the 80s

PERFORMANCE ASSESSMENT

* Apply knowledge + skills
* Naturally occurring or structured exercises
* Original responses
* Observe + rate

❖EXAMPLES❖

Oral reading

Writing

Speaking

Foreign language dialogue

Product evaluation

SUBJECT MATTER AREAS

- Communication Skills – Reading
  Writing
  Speaking
  Listening
- Business and Industrial Education
- Psychomotor Development
- Visual and Performing Arts
- Social-Emotional Development
- Affective Assessment

Why Use Performance Tests?

- Demonstrate Skills
- Real-life Situations and Competencies
- Effective Diagnosis
- Observe Natural Events
- Continuous Feedback

PERFORMANCE ASSESSMENTS

Formal: Preplanned, structured

Informal: Spontaneous observations

PERFORMANCE ASSESSMENT DEVELOPMENT

* Clarify Reason(s) for Assessment

* Clarify Performance

* Design Exercises

* Design Rating Plan

1. Clarifying Reason(s)

A. Specify Decisions to be Made

___Individual Diagnosis

___Group Needs Assessment

___Grading

___Grouping

___Selection

___Certification

___Evaluation

___Other _____

B. Specify Decision Maker(s)

___Teacher

___Student

___Administrator

___Parent

___Other_____

C. Use of Results

___Mastery

___Rank

___Combination

D. Describe Students

___Number

___Grade

FACTOR TO CONSIDER

* NATURE OF THE DECISION

TO BE MADE

2. Specify Performance

A. Specify Focus of Assessment

___Content Domain

___Skill Focus

B. Select Type of Performance

___Process

___Product

___Combination

C. Identify Performance Criteria

List All Key Ingredients

FACTORS TO CONSIDER

* IN OBSERVABLE TERMS

* FACTOR DEFINITIONS

* CONTINUUM (POOR–GOOD)

* FEEDBACK TO PERFORMER

3. Design Exercises

A. Select Form

__Structured Exercises
__Natural Events
__Combination

B. Obtrusiveness

__Aware of Assessment
__Unaware of Assessment
__Combination

C. Amount of Evidence

__1 Sample, 1 Time
__Mult. Samples, 1 Time
__Mult. Samples Over Time

FACTORS TO CONSIDER

* SERIOUSNESS OF DECISION

* EQUITY ISSUE

* NATURAL AVAILABILITY

  OF EVIDENCE

FACTORS TO CONSIDER

* ALL NECESSARY INSTRUCTIONS

* CLARITY OF COMMUNICATION

FACTORS TO CONSIDER

* IMPORTANCE OF DECISION

* NATURAL AVAILABILITY OF EVIDENCE

* MOTIVATION ISSUE

* STUDENT PRIVACY RIGHTS

FACTORS TO CONSIDER

* REPRESENTATIVENESS

* IMPORTANCE OF DECISION

* TIME PER SAMPLE

* TIME UNTIL DECISION

4. Design Rating Plan

A. Type of Score Needed

___Holistic

___Analytical

___Combination

B. Specify Rater

___Teacher

___Peer

___Self

___Combination

C. Design Recording Method

___Checklist

___Rating Scale

___Anecdotal Record

___Portfolio

___Combination