

1994

In Service of Service-Learning: Framing an evaluation for effective programs

Steven Schuman

Follow this and additional works at: <https://digitalcommons.unomaha.edu/slceeval>



Part of the [Service Learning Commons](#)

Recommended Citation

Schuman, Steven, "In Service of Service-Learning: Framing an evaluation for effective programs" (1994). *Evaluation/Reflection*. 28.
<https://digitalcommons.unomaha.edu/slceeval/28>

This Report is brought to you for free and open access by the Barbara A. Holland Collection for Service Learning and Community Engagement (SLCE) at DigitalCommons@UNO. It has been accepted for inclusion in Evaluation/Reflection by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.



DRAFT**In Service of Service-Learning: Framing an evaluation for effective programs**

Steven Schuman
 Assessment and Evaluation Center
 Federal Way School District
 Federal Way, Washington

In 1991 at the Wingspread Conference (Giles, Honnet, and Migliore, 1991) two broad questions emerged to help define and frame the evaluation of service-learning programs:

- 1) What was the effect of service-learning on intellectual, moral, and citizenship development of participants?
- 2) What was the effect of service-learning on the advancement of social institutions and democracy?

These two themes are parallel but differ significantly in focus. The first concerns the individual's ability and willingness to participate in a democratic society, while the second concerns the collective process of building a just and effective society. The following five categories of specific questions about service-learning were developed by the conference participants.

- Focus 1: The Participant**
- Focus 2: The Educational System**
- Focus 3: The Community Beyond the School Walls**
- Focus 4: The Theoretical Basis for Service-Learning**
- Focus 5: The Program Approaches to Service-Learning**

NSLC
 c/o ETR Associates
 4 Carbonero Way
 Scotts Valley, CA 95066

Focus 1: The Participant

What are the general effects of the service-learning experience on the service giver and the served?

- What is the effect of service-learning on students as learners?
- What knowledge do students gain as a result of service-learning?
- Does participation in service-learning affect the participants' perception of self and others, prosocial attitudes and behaviors, and view of the world?
- What is the effect of service-learning on participants as citizens?
- Do learner characteristics, such as age, socio-economic status, developmental stage, and family background and support lead to different social developmental outcomes?

In Service of Service-Learning: Framing an evaluation

- Do different models of service-learning lead to different types of world views, value constructions, or skill development in participants?

Focus 2: The Educational System

What is the effect of service-learning on the improvement of the educational system and on specific types of participating agencies?

- What are the outcomes of service-learning which contribute to institutional missions?
- How can service-learning lead to the effective integration of teaching, assessment, and service?
- How can service-learning be used as a vehicle for reform in areas of teaching effectiveness, curriculum design, teacher training, school mission and structure, and practical use of theories of learning and development?
- How can traditional subjects be taught effectively by incorporating a service-learning component?

Focus 3: The Community Beyond the School Walls

What is the effect of service-learning on community improvement?

- To what extent does service-learning promote multicultural understanding within institutions, communities, and society?
- Does service-learning result in the development of long-term habits of participation in the community?
- What are the benefits and costs for communities as a result of service-learning?
- How does service-learning contribute to the collaborative development of democratic community?

Focus 4: The Theoretical Basis for Service-Learning

How can service-learning research contribute to the development of theories that can further undergird and illuminate service-learning?

- How can service-learning contribute to the development of more comprehensive theories of human development?
- How can service-learning contribute to the development of more comprehensive theories of community development?
- How can service-learning contribute to the development of more comprehensive theories of epistemology and learning?
- How can human development, community development, and learning theories be used to increase our understanding of effective service-learning?

Focus 5: The Program Approaches to Service-Learning

What are the components and outcomes of various models of service-learning?

- Is there a difference in impact on students between programs which use systematic reflection and those that don't?
- What program characteristics have enhanced or deterred the institutionalization of service-learning?
- What program characteristics, such as duration, intensity, content, and mandatory or voluntary participation, promote various outcomes?

In Service of Service-Learning: Framing an evaluation

- How can service-learning be incorporated effectively into the curriculum at a variety of grade levels and throughout the disciplines?

Types of Service-Learning Evaluation Targets

There are five targets to aim service-learning program evaluations around. Typically, programs focus on one type of target and provide supporting evidence that may relate to another target. For example, a program may target student attitude change and then complement such evidence with indications of related student achievement or attendance changes. In practice, it is better to include data that are from two or more target areas.

Target 1: Document Program Existence

Target 2: Growth in Student's Attitudes and Preferences

Target 3: Growth in Student's Academic Skills and Knowledge

Target 4: Improved Program Practices and Procedures.

Target 5: Improved Staff Attitudes and Behaviors

Target 1: Document Program Existence

Evidence of improvements in practices and procedures will be in the form of existing records. With careful planning, special recordkeeping procedures may be instituted as a program begins in order to measure change. Types of records include:

- programmed and actual budgets
- records Or expenditures
- records Or staff utilization
- participation, enrollment, and attendance counts
- materials circulation records
- number Or requests received
- number of requests filled
- response time records

Questionnaires, surveys, interviews and structured observations are sometimes used when new or additional data must be collected. Evaluation designs usually rely upon pre/post measurement. A no-treatment expectation is derived from the previously existing conditions, or in some cases from results Or programs having similar goals.

Of all the target models, this one has the most difficulty establishing a suitable comparison standard. Programs making targets of intermediate improvement in practices and procedures may fall into one of two categories:

assessed strictly on their own merits. for there are no data on similar programs with which to make comparisons. Nonetheless, the program may be a successful and highly valuable innovation that merits widespread adoption.

In Service of Service-Learning: Framing an evaluation

In the case of a competitive program it should be possible to compare a program's results not only with preexisting conditions, but also with the results of programs addressing similar goals in similar locations. Thus, appropriate standards of comparison are available and the evaluation task should be similar to that for other models.

In the case of a unique practice, there may be no legitimate standard of comparison. The comparison to preexisting conditions should certainly be made; at least it establishes the existence of change. Program evaluators should be able to answer questions such as the following: Is a demand created? Of what size? How permanent? Who are the users? What are their comments? How do costs and usage compare?

The problem still remains that comparisons between a unique practice and a "do nothing" treatment are basically artificial, proving that something is better than nothing does not prove that it is worthwhile. For programs of the unique type, a sound evaluation design is important, but because it cannot offer a realistic standard of comparison, it provides far less support than usual for targets of effectiveness. The evaluator or program developer will have to make up for this deficiency by addressing the question of educational importance thoroughly and persuasively.

In the future, the process of panel review is likely to contribute to the development of standards for consideration of programs which target to improve practices. As such programs come before the panel in growing numbers, a body of comparison data will emerge for certain types of programs. As targets are scrutinized and the body of knowledge grows, standards will be formulated, discussed, challenged, and adjusted, just as they have been for more traditional educational programs in the past.

Target 2: Growth in Student's Attitudes and Preferences

After documenting that the program exists, this target is often the first thought of by program planners. The evaluation focuses on documenting how the program changed students' affective behaviors that, in the long run, will improve both educational and community-based outcomes. Activity of this target typically uses both soft (case study or anecdotal) and hard data to show positive affect change in the server and the served. Typically targets of this type are intended to achieve changes in a specific subgroup rather than in the entire participants in the program.

To substantiate program effects, teachers may assemble a body of evidence from various sources. Thus, several different types of measures are often used in combination, including questionnaires or other self-report inventories; observations or interviews; and unobtrusive measures, like attendance. Examples include school records; case studies of individual student, classes, or schools; focused interviews with students, parents, teachers, community service agencies, or government groups; and post-program follow-up of students' course, college or job choices.

Target 2 assessments are the appropriate evaluation strategy to be used if the program meets both of the following conditions:

In Service of Service-Learning: Framing an evaluation

- ✓ It is aimed at the immediate effect of producing changes in specific attitudes and behaviors of students; and
- ✓ It presumes that the attitude or preference outcomes will contribute to student achievement some time in the future; or
- ✓ It fosters better attitudes by participants toward the schooling process and the educational system in general through immediate attitude-behavioral change.

Examples of Target 2

- *Positive attitude about goals of the program:* With increased involvement in the program, students demonstrated significant increases in tolerance for individual differences, more awareness of the community's social issues, and greater enthusiasm for creating change through personal action as indicated by structured interviews.
- *Greater preference for a subject:* After participating for nine months in a semester Social Science class centered around the activities of the school district's sheltered workshop, junior high students took more Social Science Department classes in high school than was required for graduation.
- *Improved personal or physical change:* Based on performance measures, students in the program improved their sense of self control (i.e., locus of control) and fitness scores significantly more than students in the comparison group not in the program.
- *Increased and sustained attendance:* Students in the program and served persons at the Care Center showed an attendance rate significantly higher than that of the comparison group during the program year. This attendance gain was maintained during the following year.
- *Seeking additional experiences:* Students exhibited more favorable attitudes toward students with disabilities, as measured by an attitudinal inventory. Parent questionnaires and staff attendance records showed program students seeking additional time to work with the Special Olympics Program outside of the school day's program.
- *Rise in specific course selection:* During the three subsequent years after the service-learning program, participants were more likely than contrast group students to enroll in courses designated academic.
- *Completion rates are higher:* Students were more likely to complete a course and received better grades if it had service-learning components in its learning units. This trend was maintained for five subsequent years after out-of-the-district funding for the program had lapsed.
- *Lower dropout rates:* This "Care for One; Share with All" program which was aimed at students who had previously dropped out of high school resulted in an increase in the number of students successfully graduating. Two years after implementation of the

In Service of Service-Learning: Framing an evaluation

service component in the Senior Project, the school dropout rate fell below the average of ten years previously. No change was observed in similar schools nearby.

• *Decreased vandalism and detention rates:* One year after the Community in Schools program began in all classrooms, school records indicated a decline in the monthly average of incidents of vandalism and disciplinary actions and a reduced cost to repair what damage did occur.

Target 3: Growth In Student's Academic Skills and Knowledge

This is the conventional target. It requires measurement of learning in specific school subjects or interdisciplinary skills that span across classrooms or school walls. Typically, comparisons are made between the server's post service skills and knowledge and other appropriate contrast groups, standards, or pre service levels.

Projects using this target present program evidence based on familiar educational measures. Standardized norm-referenced tests, locally or regionally developed performance-based assessments, and criterion-referenced tests are typical measures. These measures usually have well established and tolerable reliability (both inter-rater agreement and traditional reliability estimates). More importantly they should have content that accurately reflects the outcome(s) under study, i.e., content and construct validity.

Target 3 evaluations are the appropriate evaluation strategy to be used if the program meets both of the following conditions:

- ✓ It is aimed at the immediate effect of producing changes in specific academic knowledges, skills and behaviors of students; and
- ✓ It presumes that service-learning activities will significantly lead to improved general learning of skills and will contribute to student betterment in the schooling process and the educational system in general.

Examples of Target 3

• *Gains of commonly shared or valued knowledge/skills:* Students in the service-learning program made greater advancement than (1) themselves before they were in the program, (2) other local students not involved in the program, (3) a nationally normed group on a standardized test or a comparable group of students on locally accepted standards.

• *Generalizing knowledge and skills:* Even though matched on preprogram skills, program students achieved significantly better ratings on both holistic and analytically scored writing samples than did non involved students in the regular language arts program.

• *Acquiring new skills or forms of knowledge (i.e., skills or knowledge not normally presented in the typical curriculum):* When compared with a contrast group, service-

In Service of Service-Learning: Framing an evaluation

learning students achieved significantly better on a valid and reliable measure of law-related education about the civil rights of the homeless.

• *Applied use of a skill or knowledge:* Program students required significantly less assistance in finding community volunteer resources than comparison students, as measured by structured checklists used at the Volunteer Bureau.

• *Efficiency of learning:* Students completing a six month period in an unique service-learning program performed as well or better on a common assessment measure as did a matched comparison group of students serving in the traditional twelve month program.

Target 4: Improved Program Practices and Procedures.

There is another class of programs in which the goals have to do with changes in the education system—its efficiency, the types of services it provides, or coordination among its different elements. Examples might be programs that reduce costs, save labor, promote interdepartmental cooperation, provide new types of services, or improve services to particular client groups. Such programs may operate in schools or in other institutions with education-related missions and links to schools, such as libraries and museums.

Targets that are related to student learning either directly or in intermediate fashion and not simply in improved functioning of educational institutions. For example, efficiency in plant maintenance or cafeteria food savings are important objectives for schools but such types of changes are outside the scope of most teacher's missions.

Much less the focus of service-learning programs is the attempt to document its effect on the environment of learning itself. Targets of this type are outcomes that have systemic changes to the culture of learning, itself. Much like the movement toward performance-based or authentic-based assessments which is attempting to improve the function of assessment in the context of the instruction, service-learning targets of this type document change to the selection of educationally relevant contexts and outcomes; like life-long learning, serving, and civic responsibility or the cost savings in educational environs as a result of the linkage of service to the organization of the school.

If the goals of the program relate to intermediate changes in a specified participant group (i.e., teachers or students), then Target 2 or 3 should be used. Target 4 is appropriate when the program meets the following conditions:

- ✓ It is aimed at the immediate effect of producing changes in the school, system, or institution; and/or changes in a general population or service area;
- ✓ It consists of a coherent set of procedures that can be transferred to similar institutions; and
- ✓ It postulates that the outcomes will contribute to student achievement some time in the future.

In Service of Service-Learning: Framing an evaluation

Examples of Target 4

- *Improvements in service to particular client groups:* Through the satellite programs for the homeless, program children at five sites participated in the series for a cost of approximately \$1.00 per student.
- *Reduction in costs and improvements in efficiency of service delivery:* By instituting a cooperative program among three school districts, duplicate management service-learning efforts were eliminated and expanded opportunities were provided at significantly reduced costs.
- *Increase in use of information:* As a result of the program, average monthly circulation of meals-on-wheels to home-bound seniors increased by one-half standard deviation over monthly figures for the pre-program year.
- *Increase in use of resources and facilities:* One year after conversion of a surplus school building into a homework library staffed by student-librarians and stocked with young-adult level materials, monthly figures for library visits quadrupled, the number of hours of tutoring doubled, and the local schools' average daily attendance rate was significantly three percentage points better.
- *The competitive practice.* Certain service-learning programs provided by education-related institutions are fairly standard; they have traditionally been provided, and probably will always be needed. In schools these include basic instructional and support functions. In local agencies, they include things like the circulation of food and the provision of social services. In museums, planetariums, zoos, or aquariums, they include the dissemination of knowledge about natural phenomena and cultural or historical artifacts. A given program may have better methods for these standard activities than do most programs in similar institutions. Compared to others, the program may result in greater efficiency, increased use, or lower costs.
- *The unique practice.* This may be a program or practice that is being reported for the first time, that addresses different goals and targets from any seen previously. It may result from a new technology, an attempt to serve an unserved population, or the introduction of new knowledge. The changes produced by such a program show a cutting edge solution to problems shared by others.

Target 5: Improved Staff Attitudes and Behaviors

Many programs seek to improve teaching and learning by influencing teachers' attitudes and changing their teaching behaviors, or both. Targets of this type focus on programs that change teachers' attitudes and behaviors in order to improve the teaching process. They require demonstration of changes in attitudes or behaviors, and presentation of a reasonable link between these results and an educationally important goal.

In Service of Service-Learning: Framing an evaluation

Target 5 should be used if the program meets both of the following conditions:

- ✓ It is aimed at the intermediate effect Or producing changes in the attitudes and behaviors of teachers; and
- ✓ It postulates that these changes will contribute to student achievement some time in the future.
- ✓ Targets of this type are intended to achieve changes in a focus group rather than in an entire institutional population.

• *Increase in the amount of instruction devoted to a subject:* After implementation of the new hands-on science program, participating teachers reported an increase or at least 20 minutes per week in the time devoted to science instruction, while non-program teachers showed no increase ($p < .01$). Pre- and post-classroom observation figures confirmed this finding.

• *Increase in total instructional time:* Teachers who participated in the computer management program reduced the time spent on recording attendance, tardiness, homework completion, lesson assignment, and test scores. Classroom observations showed that they increased their time spent on direct instructional contact by one-third standard deviation over a 1-year period, and maintained this gain throughout the following year.

• *Change in instructional methods:* This program produced changes in teachers' instructional strategies for teaching Shakespeare, including greater interest and enthusiasm for the subject, and greater use of methods emphasizing student participation in actual dramatic performance. These effects were documented through questionnaire response and through voluntary teacher participation in the program, which over 7 years increased from 30 to 100 teachers at the elementary level and from 30 to 150 teachers at the secondary level

• *Change in emphasis within a discipline:* Social studies teachers who participated in the research and problem-solving workshops modeled problem-solving approaches more frequently in the classroom and gave more assignments requiring use of research skills than did a comparison group Or teachers drawn from the same schools. After 1-year, teacher questionnaires, student questionnaires, and pre- and post-classroom observations all showed statistically significant differences.

Assessment Issues 1: Design of the Evaluation

1. Is the program and the contrast groups similar enough in educationally relevant ways, e.g., age; gender; years in courses X, Y or Z; parent involvement in school, etc.?
2. If only a sample of the program participants is used, will it be a representative sample, i.e., was the sample selected non biasedly? Is the same true for the contrast group?

In Service of Service-Learning: Framing an evaluation

3. Will the sample size be large enough to generalize to the program group as a whole? The fewer the number of students that represent the program at large the harder it is to show effects of the program.
4. Have enough learners remained in the program so that attrition and its effects won't jeopardize the findings?
5. What might happen if the program hadn't taken place, e.g., projected trends in vandalism?
6. Is the skill, knowledge, behavior, or attitude standard (e.g., norms, measure's criterion) an appropriate comparison?
7. Has the design used the most realistic no-treatment expectation available?
8. If comparison was to other treatments, how similar were the situations?
9. Does the design allow for pre-post assessment or time-series assessment over a time period sufficiently long to assure stability? Has attention been paid to assessment of implementation? Has the population in question been exposed to the treatment in a uniform way?
10. Is a variety of measures and comparisons used?
11. If sampling was used, how representative is each sample of the larger group?
12. Does the design allow assessment of progress toward long-term academic goals?
13. If comparison was between groups of teachers, how was their pre-treatment equivalence documented?
14. If the evaluation used a sample of teacher participants, how was the sample selected? How representative is the sample of the participant group? How valid is the comparison sample?
15. To what extent are the teachers representative of the general teacher population in terms of background, training, and experience?
16. Are the samples large enough to generalize with confidence to the population as a whole?
17. Is the sample large enough to have confidence in the reliability of the observed effect?
18. Are selection methods unbiased, as opposed to having the treatment group formed of teachers who volunteered and the comparison group of those who did not volunteer?
19. What was the response rate of self-reporting measures? How was non response bias addressed?

Assessment Issues 2: Measures, Data Collection and Procedures

In assembling the battery of measures to be used, the evaluator should carefully consider a number of common sense questions: If the hypothesized change is taking place in teachers, how would we see it? How many different kinds of indicators can we identify? Which of these can be measured by unobtrusive means (the most objective)? Which by systematic observation (some possible subject or observer bias)? Which by self-reporting instruments (the most subjective)? After making a list of possibilities under each of these categories, appropriate choices can be made, taking into account the combinations of indicators which would be most scientifically sound, most persuasive to an outside review panel, and most feasible to implement given the resources available for evaluation.

In Service of Service-Learning: Framing an evaluation

Evaluators should not only present information regarding the appropriateness, reliability, and validity of the instruments, but also document the adequacy of data collection procedures. For example, if observations or interviews are used, the training of data collectors and the means used to determine their reliability should be described.

20. Are the measures (and procedures) valid and reliable for the outcomes of service-learning being evaluated?
21. Are the attitudes or behaviors being measured representative of the desired outcome(s)? How is this validity demonstrated?
22. Are the measures acceptable to others in the field of service-learning?
23. Will impartial procedures be used to insure reliable scoring if non objectively scored measures are used?
24. What will be done to ensure objectivity of and consistency between the interviewers or observers? Will observers be trained to ensure that the same attribute was seen across the observers?
25. Will instruments or procedures be administered uniformly across students?
26. Will the data be collected at a time appropriate for the measures (e.g., norming periods) and at the time the trait(s) under study is at its maximum strength (e.g., just after learning or after being applied to real world, sustaining environs) ?
27. Will alternate forms and standardized administrations to be used?
28. Will the range of possible scores, ratings, etc. be large enough to avoid negative floor and ceiling effects?
29. What steps will be taken to promote objectivity on self-reports measures?
30. What kind of response rate was obtained? How was non response bias addressed?
31. performance ratings were assigned by raters (e.g., grades), will care be taken to ensure systematic rating across teachers?
32. Will the measurements be taken often enough to estimate stability of the attitude/behavior over time?
33. Is the information collected valid for the targets made?
34. If existing records are used, what safeguards exist to ensure accurate completion and maintenance of recorded information?
35. Is information collected from records complete?
36. Does the information reported reflect all aspects of the treatment?
37. Are the reporting units reasonable for purposes of comparison with existing standards?
38. If self-reporting measures were used, what was done to promote objectivity? What was done to deal with non response?
39. How well does the attitude or behavior measured by the instruments correspond to the underlying treatment construct?
40. How reliable are the instruments and data collection procedures?
41. If self-report, what cautions have been taken to ensure objectivity?
42. If observations are used, how is the observation schedule related to implementation Or the treatment? Have multiple observations been used to measure stability of results?
43. Has there been attrition from the sample(s)? What are the reasons for attrition? What effect might this have had on results?

In Service of Service-Learning: Framing an evaluation

44. If the major measures rely on self-report, what other evidence corroborates this data?

Assessment Issues 3: Analysis and Conclusions

In addition to questions raised earlier regarding correctness, clarity, and plausibility, do the following apply:

45. Are the samples large enough to generalize with confidence to the population as a whole?
46. Are results differentiated by student characteristics such as age, sex, ethnic identity, ability?
47. Is there a plausible relationship between the nature of the treatment and the effects targeted?
48. Is there evidence that the effects are sustained or do they diminish after completion of the treatment?
49. Has the program examined unintended outcomes as well as intended results?
50. What is the hypothesized link between what is measured and student achievement?
51. Do experts agree that the change suggested by outcomes is an important one?
52. Are there harmful effects?
53. Are there rival hypotheses that could account for the observed change in student attitudes or behaviors? For example, might the change be attributed to
 - other programs sanctions, or incentives?
 - outside social influences or larger societal trends?
 - specific local events, such as student deaths related to behaviors later targeted by the program?
54. What is the range of situations in which the results have been observed? Different departments? Different school organizations? Different types of communities?
55. Are there any unintended benefits from the program? For example, are there positive carryovers into other teaching areas?
56. Are there any unintended negative effects from the program? For example, have time or resources been taken away from other disciplines?
57. Are the observed effects accounted for by rival hypotheses, such as:
 - other training, incentives, or requirements that affect teachers?
 - unique or unusual characteristics of the program school(s)?
 - Hawthorne or halo effects from participation in the program?
58. How long-lasting are the observed changes?
59. Will analysis procedures be appropriate for the evaluation design and goal, i.e., appropriate statistics for correlational or causative conclusions?
60. Will the implementation of the program be measurable?
61. Were appropriate scores used and did the results show statistical significance?
62. Was the program implemented enough so that the gains or lack of same can be attributed to the program's key features and not to low levels of implementation?
63. Are the obtained differences between the program and contrast group (i.e., effect sizes) large enough to indicate real world differences? Are the findings plausible and consistent with personnel's observations?
64. What expense, if any, was made on other subjects or procedures to strengthen the program area?

In Service of Service-Learning: Framing an evaluation

65. Could staff ability, attitude, experience, or charisma account for the program results, e.g., teaching to the test measure, Halo or Hawthorne effects, years of teaching, maturation?
66. Do the changes or improvements compare favorably to standard practices in similar institutions?
67. Is there a clear link between potential student achievement and the attainment of immediate goals?
68. Are the ultimate goals worthy ones?
69. Is the demonstrated change worthwhile in terms of cost? Are the time savings worthwhile?
70. Are there any unintended negative effects from the program? For example, are time or resources taken away from other areas?
71. Are there rival hypotheses that could account for the observed effects?
72. Is the scope of the change significant? Are other institutions likely to be interested in adopting the treatment?

In Service of Service-Learning: Framing an evaluation

FRAMEWORK

- Focus 1: The Participant**
- Focus 2: The Educational System**
- Focus 3: The Community Beyond the School Walls**
- Focus 4: The Theoretical Basis for Service-Learning**
- Focus 5: The Program Approaches to Service-Learning**

- Target 1: Document Program Existence**
- Target 2: Growth In Student's Attitudes and Preferences**
- Target 3: Growth In Student's Academic Skills and Knowledge**
- Target 4: Improved Program Practices and Procedures.**
- Target 5: Improved Staff Attitudes and Behaviors**

- Assessment Issues 1: Design of the Evaluation**
- Assessment Issues 2: Measures, Data Collection and Procedures**
- Assessment Issues 3: Analysis and Drawing Conclusions**

Educational Assessment References and Resources

Articles

- Airasian, P. W. (1988). Measurement driven instruction: A closer look. *Educational Measurement: Issues and Practices*, 7, 6-11.
- Baker, E. L., O'Neil, H. F. & Linn, R. L. (1991). *Policy and Validity Prospects for performance-based Assessment*. Annual Conference of the American Psychological Association, August.
- Ericson, D. P. & Ellett, F. S. (1990). Taking student responsibility seriously. *Educational Researcher*, 19 (9), 3-10.
- Fenstermacher, G. D. (1986). Philosophy of research on teaching: Three aspects. In M. C. Wittrock (Ed.). *Handbook of Research on Teaching*. (3rd ed., pp. 87-99). NY: MacMillan.
- Frederiksen, J. R. & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18 (9), 27-32.
- Goslin, D. (1963). *The Search for Ability*. In H. Black, *They Shall Not Pass*. NY: Morrow and Company.
- Isham, M. & Kite, R. (1991). *Developing & Conducting Performance Assessments*. Annual Conference of the Association of Supervision & Curriculum Development, March.
- Jackson P.W., (19xx). *Life in Classrooms*, New York, NY: Teachers College Press.
- Kulieke, M. (1990). Assessing outcomes of significance. *The Quarterly Journal of the Network for Outcome-Based Schools*, 9 (4).
- Maeroff, G. I. (1991). Assessing alternative assessment. *Phi Delta Kappan*, Dec, 272-281.
- Mitchell, R. (1990). Performance Assessment: An emphasis on 'Activity.' *Education Week*, January 24, 36.
- Paris, S. G., Lawton, T. A., Turner, J. C., & Roth, J. L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher*, 20 (5), 12-20.
- Paulson, Paulson & Meyer (1991). What makes a portfolio a portfolio? *Educational Leadership*, February, 48(5), 60-64.
- Resnick, L. B. & Klopfer, L. E. (1989). Toward the thinking curriculum: Concluding Remarks, In L. B. Resnick & L. E. Klopfer (Eds.) *Toward the Thinking Curriculum: Current cognitive research*. Alexandria, VA: Association for Supervision & Curriculum Development, 206-211.
- Resnick, L. B., & Resnick, D. P. (1985). Standards, Curriculum and Performance: A historical and comparative perspective. *Educational Researcher*, 14, 5-21.
- Resnick, L. B., & Resnick, D. P. (1989). Assessing the thinking curriculum: New tools for educational reform, In B. R. Gifford and M. C. O'Connor (Eds.) *Changing Assessments: Alternative views of aptitude, achievement and instruction*. Boston: Kluwer Academic Publishers.
- Rief, Linda (1990). *Educational Leadership*, , March, 24-29.
- SCANS (Secretary's Commission on Achieving Necessary Skills) (1991). *What Work Requires of Schools: A SCANS report for America 2000*, U.S. Dept of Labor, Washington, DC.
- Schuman, S. (1992) Alternative Assessment: An unfolding story. In J. Bamburg (Ed.) *Assessment: How Do They Know What They Know?* Union, WA: Washington State Association for Supervision and Curriculum Development.

- Shepard, L. A. (1991). Psychometrician's beliefs about learning. *Educational Researcher*, 20 (7), 2-18.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations for a new reform. *Harvard Educational Review*, 57(1), 1-22.
- Silverman, H. (1985). Dynamic cognitive assessment: An alternative to intelligence testing. *Canadian Journal of Special Education*, 1 (1), 63-72.
- Smith, M. C. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20 (5), 8-11.
- Stiggins, R. J. & Bridgeford, N. J. (1985). The ecology of the classroom assessment. *Journal of Educational Measurement*, 22 (4), 271-86.
- Stiggins, R. J. (1985). Revitalizing classroom assessment: The highest instructional priority. *Phi Delta Kappan*, January, 363-368.
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practices*, 8 (2), 5-14.
- Taba, H. (1963). Learning by Discovery. *Elementary School Journal*, 63 (6), 308-316.
- Terwilliger, J. S. (1989). Classroom standard setting and grading practices. *Educational Measurement: Issues and Practices*, 8 (2), 15-19.
- Valencia, S. W. (1990). A portfolio approach to classroom reading assessment: The whys, whats and hows. *The Reading Teacher*, January, 338-340.
- Wang, M.C., Haertel, G.D., & Walberg, H.J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, Fall, 63(3), 249-294.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, May, 703-712.
- Wiggins, G. (1991). Standards, not standardization: Evoking quality student work. *Educational Leadership*, Feb, 48(5), 18-25.

Books

- Airasian, P. (1991). *Classroom Assessment*, New York: McGraw-Hill.
- Aiken, L.R. Jr. (1976) *Psychological Testing and Assessment*. Boston, MA: Allyn & Bacon, Inc.
- Alternative Assessment: Evaluating student performance in elementary mathematics* (1993), Palo Alto, CA: Dale Seymour Publications.
- Anderson, Lorin W. (1981). *Assessing Affective Characteristics in the Schools*. Boston, MA: Allyn and Bacon, Inc.
- American Psychological Assoc., *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Assoc., American Psychological Assoc., National Council on Measurement in Education.
- Angelo, T.A., Cross, K. P. (1992) *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco, CA: Jossey/Bass Publishers.
- Angoff, W.H. (1989) *Scales, Norms, and Equivalent Scores*. Princeton, NJ: Educational Testing Service.
- Anthony, R.J., et al. (1991). *Evaluating Literacy: A Perspective for Change*. Portsmouth, NJ: Heinemann.
- Archbald, D. A. & Newmann, F. M. (1988). *Beyond Standardized Testing: Assessing authentic academic achievement in the secondary school*. Reston, VA: National Association of Secondary School Principals.

- Assessment Standards Working Groups for the National Council of Teachers of Mathematics (1993). *Assessment Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Bamburg, J. (Ed.) (1992). *Assessment: How do we know what they know?* Union, WA: Washington State Association for Supervision and Curriculum Development.
- Batzle, J. (1992). *Portfolio Assessment and Evaluation: Developing and using portfolios in the K-6 classroom*, Cypress, CA: Creative Teaching Press.
- Belanoff, P. & Dickson M. (Eds.) (1991). *Portfolios: Process and Product*. Portsmouth, NH: Boynton/Cook Publishers..
- Bennett, R.E., Ward, W.C. (Ed.) 1993). *Construction Versus Choice in Cognitive Measurement, Issues in Constructed Response, Performance Testing, and Portfolio Assessment*, Hillsdale, NJ: Lawrence Erlbaum Assoc., Publishers.
- Berlak H., Newmann, F.M., Adams, E., Archbald, D.A., Burgess, T., Raven, J., & Romberg, T.A. (1992). *Toward a New Science of Educational Testing and Assessment*, Albany, NY: State University of New York Press.
- Botel, M., Botel-Sheppard, B., Botel-Sheppard, R., Stokes, R., (1986). *Comprehension Test Strategies, Book D*, (middle grades), New York, NY. Scholastic Inc.
- Botel, M., Botel-Sheppard, B., Botel-Sheppard, R., Stokes, R., (1986). *Comprehension Test Strategies, Book E*, (middle grades), New York, NY. Scholastic Inc.
- Bouffler, C. (Ed.). (1992) *Literacy Evaluation, Issues & Practicalities*. Portsmouth, NH: Reed Publishing.
- Bowman, R.P. (1987). *Test Buster Per Rally*. Minneapolis, MN. Educational Media Corporation.
- Brandt, R.S. (Ed.). *Performance Assessment: Readings from Educational Leadership*, Alexandria, VA: Association for Supervision & Curriculum Development.
- Brooks, J. G., Brooks, M. G., (1993), *In Search of Understanding, The Case for Constructivist Classrooms*, Alexandria, VA: Assoc. for Supervision and Curriculum Development.
- Brown, H. & Cambourne, B., (1987), *Read and Retell*. Portsmouth, NH: Heinemann.
- Buros, O. K. (1961). *Tests in Print*. Highland Parks, NJ: The Oryphus Press, xxiii.
- Camp, R., Seidel, S. Wolf, D. Zessoules, R. (1992). *Arts Propel: A Handbook for Imaginative Writing*. Princeton, NJ: Educational Testing Service.
- Cawley, J.F. (Ed.) (1985). *Cognitive Strategies and Mathematics for the Learning Disabled*. Rockville, Maryland: Aspen Systems Publications.
- Cawelti, G. (Ed.) (1993). *Challenges and Achievements of American Education: 1993 yearbook of the Association for Supervision and Curriculum Development*, Alexandria, VA: Association for Supervision and Curriculum Development.
- Centre for Language in Primary Education (1991) *The Primary Language Record: Handbook for Teachers*, Portsmouth, NH: Heinemann Educational Books.
- Chandler, M. & Chapman, M. (Eds.) (1991). *Criteria for Competence*, Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Cognetta, R.A., Malvetti, A.L., Plakos, J., Wells, J., (1986). *Testwiseness, Improving Student Test-Taking Skills*, California Evaluation Improvement Program.
- Cohen, M. (1980), *First Grade Takes a Test*, New York, NY: Dell Publishing.
- Cooper, C. R. & Odell, L. (Eds.) (1977) *Evaluating Writing: Describing, measuring and judging*, State University of

New York at Buffalo, NY: National Council of Teachers of English.

Curriculum and Evaluation Standards for School Mathematics, (1989). Reston, VA: National Council of Teachers of Mathematics.

Davidson, L., Myford, C., Plasket, D., Schipp, L., Swinton, S., Torff, B., Waanders, J. (1992). *Arts Propel: A Handbook for Music*. Princeton, NJ: Educational Testing Service.

DeFina, A.A. (1992). *Portfolio Assessment*. NY, NY: Scholastic Inc.

Dillman, D.A., (1978) *Mail and Telephone Surveys: The Total Design Method*. NY, NY: John Wiley & Sons.

Educational Testing Service. (1992). *Information for National Standards for Education: What They Might Look Like. A Workbook*. Princeton, NJ: Policy Information Center, Educational Testing Service.

English, F.W.D (1992). *Deciding What to Teach and Test: Developing, aligning, and auditing the curriculum*, Newbury Park, CA: Sage Publications.

Finch, F. L. (Ed.) (1991). *Educational Performance Assessment*. Chicago, Ill. Riverside Publishing Company.

Foote, A., Gitomer, D., Melamed, L., Rosenblatt, E., Simmons, S. Sims-Gunzenhauser, A., Winner, E. (1992). *Arts Propel: A Handbook for Visual Arts*. Princeton, NJ. Educational Testing Service.

Gifford, B. R. & O'Connor, M. C. (Eds.) (1993) *Changing Assessments: Alternative views of aptitude, achievement and instruction*. Boston: MA: Kluwer Academic Publication.

Glazer, S.M. & Brown, C.S. (1993). *Portfolios and Beyond: Collaborative assessment in reading and writing*, Norwood, MA: Christopher-Gordon Pub.

Goodman, K. S., Goodman, Y. M., & Hood, W. J. (1989). *The Whole Language Evaluation Book*, Portsmouth, NH: Heinemann Educational Books, Inc.

Goslin, D. A. (1967) *Teachers & Testing*. NY, NY: Russell Sage Foundation.

Grady, E. (1992). *The Portfolio Approach to Assessment (Fastback #341)*, Bloomington, IN: Phi Delta Kappan.

Graves, D.H. & Sunstein, B. S. (Eds.) (1992). *Portfolio Portraits*, Portsmouth, NH: Heinemann Educational Books, Inc.

Green, J.A. (1963). *Teacher-made Tests*. New York, NY: Harper & Row, Publishers.

Greenberg, K.L., et al. (1986). *Writing Assessment: Issues and Strategies*. New York: Longman.

Gruber, Gary (1986). *Dr. Gary Gruber's Essential Guide to Test Taking for Kids Grades 3 - 4 - 5*. New York, NY. William Morrow and Company, Inc.

Gruber, Gary (1986). *Dr. Gary Gruber's Essential Guide to Test Taking for Kids Grades 6 - 7 - 8 - 9*. New York, NY: William Morrow and Company, Inc.

Harp, B. (Ed.) (1993). *Assessment and Evaluation in Whole Language Programs*. Norwood, MA: Christopher-Gordon Publishers.

Hart, D. (1994). *Authentic Assessment: A handbook for educators*, Menlo Park, CA: Addison-Wesley Publishers.

Henning-Stout, M. (1994) *Responsive Assessment: A new way of thinking about learning*, San Francisco, CA: Jossey-Bass Publishers.

Herman, J. (Ed.) (1987). *Program Evaluation Kit*. Newbury Park, CA: Sage Publications.

Herman, J., Aschbacher, P. & Winters, L. (1992). *A Practical Guide to Alternative Assessment*. Alexandria, VA:

Association for Supervision & Curriculum Development.

Herman, J. L. & Winters, L. (1992) *Tracking Your School's Success: A guide to sensible evaluation*. Newbury Park, CA: Corwin Press.

Hoffman, B. (1962). *The Tyranny of Testing*. NY: The Crowell-Collier Press.

Howell, K.W., Fox, S.L., Morehead, M.K., (1993). *Curriculum-Based Evaluation, Teaching and Evaluation, 2nd Ed.*, Pacific Grove, CA: Brooks/Cole Publishing Company.

Hynes, D.L., Chafin, A. E., & Gonder, P. (1991). *The Changing Face of Testing and Assessment: Problems and solutions*. Arlington, VA: American Association of School Administrators.

Jacobs, L.C. & Chase, C.I. (1992). *Developing and Using Tests Effectively*, San Francisco, CA: Jossey-Bass Publishers.

Jasmine, J. (1992), *Portfolio Assessment for your Whole Language Classroom*, Huntington Beach, CA: Teacher Created Materials, Inc.

Kellaghan, T., Madaus, G.F., Airasian, P.W. (1982). *The Effects of Standardized Testing*. Hingham, MA: Kluwer-Nijhoff Publishing.

Kubiszyn, T., Borich, G. (1984). *Educational Testing and Measurement, Classroom Application and Practice*, Glenview, IL: Scott, Foresman and Company.

Kulm, G., Malcolm, S. M. Editors (1991). *Science Assessment in the Service of Reform*, Washington, D.C.: American Assoc. for the Advancement of Science.

Kulm, G. (Ed.) (1991) *Assessing Higher Order Thinking in Mathematics*. Washington, D.C.: American Assoc. for the Advancement of Science.

Lesh, R., Lamon, S.J. (Ed.) (1992) *Assessment of Authentic Performance in School Mathematics*. Washington, D.C.: American Association for the Advancement of Science.

Livingston, S.A., Zieky, M.J. (1982). *Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*. Princeton, NJ: Educational Testing Service.

Madaus, G.F., Haney, W. & Kreitzer, A. (1992). *Testing and Evaluation: Learning from the projects we fund*, NY, NY: Council for Aid to Education.

Marzano, R.J., Pickering, D. & McTighe, J. (1993) *Assessing Student Outcomes: Performance assessment using the Dimensions of Learning Model*, Alexandria, VA: Association for Supervision & Curriculum Development.

Mathematical Sciences Education Board (National Research Council) (1993). *Measuring Up: Prototypes for Mathematics Assessment*. Washington, DC: National Academy Press.

Mathematical Sciences Education Board (National Research Council) (1993). *Measuring What Counts: A Conceptual Guide for Mathematics Assessment*. Washington, D.C.: National Academy Press.

McDonald, M. (1991)., *Teachers' Messages for Report Cards, Revised Edition*, Carthage, IL: Fearon Teacher Aids.

McDonald, J.P., Smith, S., Turner, D., Finney, M., Barton, E. (1993). *Graduation by Exhibition, Assessing Genuine Achievement*. Alexandria, VA: Assoc. for Supervision and Curriculum Development.

Medina, N. & Neill, D. M. (1990). *Fallout From the Testing Explosion: How 100 million standardized exams undermine equity and excellence in America's public schools*. Cambridge, MA: National Center for Fair and Open Testing (FairTest).

Mislevy, Robert J. (1992). *Linking Educational Assessments: Concepts, Issues, Methods, and Prospects*. Princeton, NJ: Educational Testing Service.

- Mitchell, R. (1992). *Testing for Learning: How new approaches to evaluation can improve American schools*. New York, NY: The Free Press.
- Morris, L. L., Fitz-Gibbon, C. T. & Lindheim, E. (1987). *How to Measure Performance and Use Tests*. Newburg Park, CA: SAGE Publications.
- Mueller, D.J. (1986). *Measuring Social Attitudes: A Handbook for Researchers and Practitioners*. New York, NY: Teachers College Press.
- National Comm. on Testing and Public Policy, (1990), *From Gatekeeper to Gateway: Transforming Testing in America, Executive Summary*, Chestnut Hill, MA: Nat'l. Comm. on Testing & Public Policy.
- National Education Assoc. Library Publication, (1993). *Student Portfolios*, National Education Assoc. of the U.S.
- Neimeyer, G. J. (Ed.) (1993). *Constructivist Assessment: A casebook*, Hillsdale, NJ, Lawrence Erlbaum Associates.
- Office of Technology Assessment (1992). *Testing in American Schools, Asking the Right Questions*. Congress of the United States.
- Ferrone V. (Ed.) (1991). *Expanding Student Assessment*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Rhodes, L.K. (Ed.). (1993). *Literacy Assessment: A Handbook of Instruments*. Portsmouth, NJ: Heinemann.
- Rhodes, L.K. and Shanklin, N. *Sideways into Literacy Assessing Learners K-8*. Portsmouth, NJ: Heinemann.
- Roid, G. H. & Haladyna, T. M. (1982). *A Technology for Test-Item Writing*. New York, NY: Academic Press.
- Salvia, J. & Ysseldyke, J.E. (1991). *Assessment*. Boston: Houghton Mifflin.
- Savage, C.W. & Ehrlich, P. (Eds.) (1992). *Philosophical and Foundational Issues in Measurement Theory*, Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Sharp, Q. Q. (1989). *Evaluation: Whole language checklists for evaluating your children for grades K to 6*, New York, NY: Scholastic, Inc.
- Shaw, M.E., & Wright, J.M. (1967). *Scales for the Measurement of Attitudes*. New York, NY. McGraw-Hill Book Company.
- Sizer, T.R. (1992). *Horace's Compromise: The dilemma of the American high school*. NY, NY: Houghton Mifflin Co.
- Sizer, T.R. (1992). *Horace's School: Redesigning the American high school*. NY, NY: Houghton Mifflin Co.
- Smith, N. L. Ed., (1981), *Metaphors for Evaluation, Sources of New Methods, New Perspectives in Evaluation, Vol. 1*, Beverly Hills, CA: Sage Publications.
- Spandel, V. & Stiggins, R. J. (1990) *Creating Writers: Linking assessment and writing instruction*, White Plains, NY: Longman Publishers.
- Stenmark, J. K., (Ed.) (1991). *Mathematics Assessment: Myths, models, good questions, and practical suggestions*, Reston, VA: National Council of Teachers of Mathematics.
- Stiggins, R. J. & Conklin, N.F. (1992). *In Teachers' Hands: Investigating practices of classroom assessment*, Albany, NY: State University of New York Press.
- Stiggins, R. J. (1994). *Student-Centered Classroom Assessment*. New York: Merrill.
- Suen, H.K. (1990). *Principles of Test Theories*. Hillsdale, NJ. Lawrence Erlbaum Associates, Publishers.

- Test Alert, (1990). *Test Alert Level B, Teacher's Guide*, Chicago, IL. The Riverside Publishing Co.
- Test Alert, (1990). *Test Alert Level B, Teacher's Guide*, Chicago, IL. The Riverside Publishing Co.
- Thelen, J.N., Itzkowitz, L., Krulik, S., (1989). *How to Be a Better Test Taker, Book A (Grade 1), and Teacher's Edition*, New York, NY. Scholastic Inc.
- Thelen, J.N., Itzkowitz, L., Krulik, S., (1989). *How To Be a Better Test Taker, Book B (Grade 2), and Teacher's Edition*, New York, NY. Scholastic Inc.
- Thelen, J.N., Itzkowitz, L., Krulik, S., (1984). *How To Be A Better Test Taker, Book C, (grade 3)* New York, NY. Scholastic Inc.
- Thelen, J.N., Itzkowitz, L., Krulik, S., (1984). *How To Be A Better Test Taker, Book E. (grade 5)* New York, NY. Scholastic Inc.
- Thelen, J.N., Itzkowitz, L., Krulik, S., (1984). *How To Be A Better Test Taker, Book F, New York, NY. Scholastic Inc.*
- Tierney, R. J., Carter, M. A. & Desai, L. E. (1991). *Portfolio Assessment in the Reading-Writing Classroom*. Norwood, MA. Christopher-Gordon Publishers.
- Thorndike, R.M., Lohman, D.F. (1990) *A Century of Ability Testing*. Chicago, IL: Riverside Publishing Co.
- Valette, R.M. (1967). *Modern Language Testing. A Handbook*. New York, NY: Harcourt, Brace & World, Inc.
- Walberg, H. J. (1982). *Improving Educational Standards and Productivity: The research bases for policy*. Berkeley, CA: McCutchan.
- Webb, N.L., (Ed.) (1993) *Assessment in the Mathematics Classroom*. Reston, VA: National Council of Teachers of Mathematics.
- Wernick, R. (1956). *They've Got Your Number*. NY, NY: Norton Press.
- White, E.M. (1985). *Teaching and Assessing Writing*. San Francisco: Jossey-Bass.
- Wiggins, G. P. (1993). *Assessing Student Performance*, San Francisco, CA: Jossey-Bass Publishers.
- Winner, E. (1992). *Arts Propel: An Introductory Handbook*. Princeton, NJ: Educational Testing Service.
- Yancy, K.B., (Ed.) (1992). *Portfolios in the Writing Classroom: An Introduction*. Urbana Illinois: National Council of Teachers of English.

Journals

- Applied Measurement in Education*, 365 Broadway, Hillsdale, NJ, 07642, Lawrence Erlbaum Associates.
- Educational Assessment*, 365 Broadway, Hillsdale, NJ, 07642, Lawrence Erlbaum Associates.
- Educational Leadership*, Alexandria, VA: Association for Supervision & Curriculum Development.
- Educational Measurement: Issues and Practices*, Washington, DC: National Council on Measurement in Education.
- Journal of Educational Measurement*, 1230 17th Street, Washington, DC, 20036-3078; National Council on Measurement in Education.
- Phi Delta Kappan*, 6th & Union, Box 789, Bloomington, IN. 47402.

Newsletters

Assessment in Practice (Mathematical Sciences Education Board) National Research Council, 2101 Constitution Ave. NW, HA 476, Washington, DC, 20418.

CRESST Line (Center for Research on Evaluation, Standards & Student Testing), UCLA Graduate School of Education, 405 Hilgard Ave., Los Angeles, CA, 90024-1522.

FairTest Examiner (The National Center for Fair & Open Testing), Cambridge, MA. (617)854-4810.

Horace, Providence, RI., The Coalition of Essential Schools.

News on Tests, (Newsletter), Princeton, NJ., Educational Testing Service. (609) 734-5686.

Northwest Evaluation Association Newsletter. Lake Oswego, OR: Northwest Evaluation Association. (503) 624-1951.

Portfolio Assessment Newsletter. Lake Oswego, OR: Northwest Evaluation Association. (503) 624-1951.

Portfolio News. (Newsletter) Encinitas, CA: Portfolio Assessment Clearinghouse.

Report on Educational Research. P.O. Box 1453, Alexandria, VA: Capitol Publications (800) 327-7203.

The Standard Deviation (Newsletter), Washington Educational Research Association, Gig Harbor, WA.

Networks

Authentic Assessment Network (ASCD group), c/o Judith Dorsch Backes, Carroll County Board of Education, 55 N. Court Street, Westminster, MD 21157.

Classroom Assessment (AERA Special Interest Group), c/o Richard Stiggins, P.O. Box 19794, Portland, OR. (503) 452-9734.

International Thinking Assessment Network, c/o Sally Duff, Coordinator, Coppin State College, 2500 West North Ave., Baltimore, Maryland, 21216.

Resources

Bowman, R. P. (1987) Test Buster Pep Rally. Minneapolis, MN: Educational Media Corporation.

Grady Profile (1991). Aurbach & Associates, Inc., 8233 Tulane Ave., St. Louis, MO, (314) 726-5933.

Learner Profile, the observational assessment tool (1993) Wings for learning, PO Box 660002, Scotts Valley, CA, 95067, (800) 321-7511.

Training

Assessment Training Institute (Rick Stiggins), P.O. Box 19794, Portland, OR. (503) 452-9734.

Meisels, S. J. (1992). The Work Sampling System (Training and developmental checklist). Ann Arbor, MI: Center for Human Growth and Development, University of Michigan, (313)-747-1084.

Northwest Regional Educational Laboratory, Classroom Assessment Workshops, Videos sold through: IOX Assessment Associates, Los Angeles, CA.

Redesigning Assessment Series (Videos). Alexandria, VA: Association for Supervision & Curriculum Development.

Portfolio Bibliography

- Batzle, J. (1992). *Portfolio Assessment and Evaluation: Developing and using portfolios in the K-6 classroom*, Cypress, CA: Creative Teaching Press.
- Belanoff, P. & Dickson M. (Eds.) (1991). *Portfolios: Process and Product*. Portsmouth, NH: Boynton/Cook Publishers..
- Camp, R., Seidel, S. Wolf, D. Zessoules, R. (1992). *Arts Propel: A Handbook for Imaginative Writing*. Princeton, NJ: Educational Testing Service.
- DeFina, A.A. (1992). *Portfolio Assessment*. NY, NY: Scholastic Inc.
- Glazer, S.M. & Brown, C.S. (1993). *Portfolios and Beyond: Collaborative assessment in reading and writing*, Norwood, MA: Christopher-Gordon Pub.
- Grady, E. (1992). *The Portfolio Approach to Assessment* (Fastback #341), Bloomington, IN: Phi Delta Kappan.
- Graves, D.H. & Sunstein, B. S. (Eds.) (1992). *Portfolio Portraits*, Portsmouth, NH: Heinemann Educational Books, Inc.
- Jasmine, J. (1992), *Portfolio Assessment for your Whole Language Classroom*. Huntington Beach, CA: Teacher Created Materials, Inc.
- Paulson, Paulson & Meyer (1991). What makes a portfolio a portfolio? *Educational Leadership*, February, 48(5), 60-64.
- Tierney, R. J., Carter, M. A. & Desai, L. E. (1991). *Portfolio Assessment in the Reading-Writing Classroom*. Norwood, MA. Christopher-Gordon Publishers.
- Valencia, S. W. (1990). A portfolio approach to classroom reading assessment: The whys, whats and hows. *The Reading Teacher*, January, 338-340.

Mathematical Assessment References

Alternative Assessment: Evaluating student performance in elementary mathematics (1993) P. & Alto, CA: Dale Seymour Publications.

Assessment Standards Working Groups for the National Council of Teachers of Mathematics (1993). *Assessment Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Cawley, J.F. (Ed.) (1985). *Cognitive Strategies and Mathematics for the Learning Disabled*. Rockville, Maryland: Aspen Systems Publications.

Curriculum and Evaluation Standards for School Mathematics, (1989). Reston, VA: National Council of Teachers of Mathematics.

Kulm, G. (Ed.) (1991) *Assessing Higher Order Thinking in Mathematics*. Washington, D.C.: American Assoc. for the Advancement of Science.

Lesh, R., Lamon, S.J. (Ed.) (1992) *Assessment of Authentic Performance in School Mathematics*. Washington, D.C.: American Association for the Advancement of Science.

Mathematical Sciences Education Board (National Research Council) (1993). *Measuring Up: Prototypes for Mathematics Assessment*. Washington, DC: National Academy Press.

Mathematical Sciences Education Board (National Research Council) (1993). *Measuring What Counts: A Conceptual Guide for Mathematics Assessment*. Washington, D.C.: National Academy Press.

Stenmark, J. K., (Ed.) (1991). *Mathematics Assessment: Myths, models, good questions, and practical suggestions*, Reston, VA: National Council of Teachers of Mathematics.

Webb, N.L., (Ed.) (1993) *Assessment in the Mathematics Classroom*. Reston, VA: National Council of Teachers of Mathematics.

Reading/Literacy Assessment References

Anthony, R.J., et al. ((1991). *Evaluating Literacy: A Perspective for Chan.* . Portsmouth, NJ.: Heinemann.

Brown, H. and Cambourne, B. (1987). *Read and Retell.* Portsmouth, N.H. : Heinemann.

Harp, B. (Ed.). *Assessment and Evaluation in Whole Language Programs.* Norwood, MA: Christopher-Gordon Publishers, Inc.

Rhodes, L.K. (Ed.). (1993). *Literacy Assessment: A Handbook of Instruments.* Portsmouth, NJ: Heinemann.

Rhodes, L.K. and Shanklin, N. *Windows into Literacy: Assessing Learners K-8.* Portsmouth, NJ: Heinemann.

Writing Assessment References

Cooper, C. R. & Odell, L. (Eds.) (1977) *Evaluating Writing: Describing, measuring and judging*, State University of New York at Buffalo, NY: National Council of Teachers of English.

Greenberg, K.L., et al. (1986). *Writing Assessment: Issues and Strategies*. New York: Longman.

Spandel, V. (1994) *Seeing With New Eyes: A guidebook on teaching and assessing beginning writers.*, Portland, Or: Northwest Regional Educational Laboratory.

Spandel, V. & Culham, R. (1994) *Picture Books: An annotated bibliography for use with the 6-trait analytic model of writing assessment and instruction.*, Portland, Or: Northwest Regional Educational Laboratory.

Spandel, V. & Stiggins, R. J. (1990) *Creating Writers: Linking assessment and writing instruction*, White Plains, NY: Longman Publishers.

White, E.M. (1985). *Teaching and Assessing Writing*. San Francisco: Jossey-Bass.