State Council of Higher Education for Virginia

GUIDELINES FOR ASSESSMENT OF STUDENT LEARNING

Submitted to State Council for Action
SCHEV Task Force on Assessment
October 3, 2007

BACKGROUND

The assessment of student learning outcomes is not new to the Commonwealth's system of higher education. For more than two decades, the State Council of Higher Education for Virginia (SCHEV) has collaborated with Virginia's thirty-nine public colleges and universities on assessment as an integral component of a larger, shared commitment to institutional effectiveness and public accountability.

In 1985, Senate Joint Resolution 125 directed SCHEV to "investigate means by which student achievement may be measured to assure the citizens of Virginia the continuing quality of higher education in the Commonwealth." SCHEV responded by outlining its first assessment guidelines in Senate Document No. 14, "The Measurement of Student Achievement and the Assurance of Quality in Virginia Higher Education." In this document, SCHEV recommended that "all state-supported institutions of higher education establish procedures and programs to measure student achievement" and that these procedures and programs "recognize the diversity of Virginia's public colleges and universities, the tradition of institutional autonomy, and the capacity of faculty and

administrators to identify their own problems and solve them creatively." Further, the document called for institutions to furnish "concrete, non-anecdotal and quantifiable information on student achievement to the Council of Higher Education."

Building on this landmark work, in 1987, SCHEV worked with Virginia's public colleges and universities to develop guidelines for assessing general education. Leaving the specific procedures for the assessment to the individual institutions, the guidelines stipulated that each college or university may choose to employ either absolute assessment measures or those that demonstrate the value-added "contribution the institution has made to the student's development." The year also marked the first time that the General Assembly provided a line-item appropriation to institutions for assessment. Incorporated into their respective base operating budgets, the appropriation was typically used by institutions to cover the personnel costs associated with assessment.

Both SJR 125 and its ensuing guidelines framed the Commonwealth's approach to student assessment for the remainder of the twentieth century. Institutions evaluated the learning outcomes of students in the major and the general education core.

In 1999, the Governor's Blue Ribbon Commission on Higher Education affirmed this state-wide approach to student assessment and commended institutions for using the results of their assessment activities to improve student learning and inform academic program decision-making. The Commission further identified six areas of knowledge and skills that cross the bounds of academic discipline, degree major, and institutional mission to comprise basic competencies that should be achieved by all students completing a degree program at a Commonwealth institution of higher education—namely, Information Technology Literacy, Written Communication, Quantitative Reasoning, Scientific Reasoning, Critical Thinking, and Oral Communication. Terming them "areas of core competency," the group recommended that institutions conduct regular assessments of these areas, the results of which would be shared with the general public.

SCHEV subsequently adopted these six areas of core competency and established a timetable for institutional reporting, with two areas to be reported on biennially so that all six competencies would be cycled through within a four-year span. Four-year institutions and Richard Bland College submitted their initial reports in 2002 and their final reports in 2006. The Virginia Community College System began its cycle in 2003 and finished in 2007.

In 2005, as the Council anticipated completion of the first full round of reporting on the six areas of core competency by four-year institutions, it established the Assessment of Student Learning Task Force (ASLTF). The ASLTF's charge was to work directly with the Academic Affairs Committee on reviewing the experiences of institutions with the core competencies and developing guidelines for improving the process through the next cycle. This collaborative effort led to the following conclusions:

1. An important strength of the Commonwealth's higher education system lies in the diversity of its institutions, which affords Virginians a wide range of distinctive learning experiences. This strength must be preserved and nurtured.

2. Each institution should continue to be responsible for implementing an assessment program that is congruent with its mission and goals; provides the kind of data needed for informed decision-making about curricula; and offers both policymakers and the general public useful information on student learning.

3. Assessment should continue to *fit*, rather than *drive*, the institution. It should be reasonable in its requirements for time, resources, and personnel and should, ideally, be integrated with the institution's larger framework for continuous improvement and public accountability. It

should also employ both valid and reliable measurements of educational experiences and student learning.[1]

4. Assessment should continue to focus on the improvement of learning while providing meaningful demonstration of accountability. It should continue to employ the six core areas and explore options to address the Council's preferred "value-added" approach that speaks to demonstrable changes as a result of a student's collegiate experience.

In Fall 2006, as SCHEV began to develop guidelines for the new cycle of assessment to begin in Fall 2007, the Council of Presidents asked to take a deeper look at the complexities of incorporating value-added elements meaningfully into the institutional assessment of student learning and offered to collaborate with the State Council's Academic Affairs Committee on a new set of assessment guidelines that could be implemented in Fall 2008.

As a result of this overture, SCHEV passed a formal resolution establishing the 2007 Task Force on Assessment and charging it with developing "process-oriented" guidelines for "incorporating value-added methods and interpretations into institutional assessment programs." More specifically, it delineated the 2007 Task Force's charge as follows:

> *The overarching goal of the Task Force is to steward a system wide approach to assessment that is focused on improvement while also demonstrating accountability; builds on successful value added aspects of current assessment practices; is meaningfully related to institutional*

---

[1] Throughout these guidelines, *educational experiences* are the independent or causal variables of concern and *student learning* is the dependent variable of interest. Inference, validity, and reliability are the critical issues. The concern with inference is that institutions draw good, reasoned, and empirically sound conclusions. Validity refers to the appropriateness of measures for representing intended concepts—in this case, educational experiences and student learning. Reliability refers to the consistency or replicability of the measurements.

missions, practices, and educational aspirations; and will be sustainable into the future.

PURPOSE

An outgrowth of the work of the 2007 Task Force on Assessment, *Guidelines for Assessment of Student Learning* serves as an institutional resource for implementing value-added assessment in Virginia's colleges and universities. The document defines both the general parameters for and operating conception of value-added assessment; clarifies the relationship between "competency-based" assessment and "value-added assessment"; and delineates the differing approaches, data collection methods, and measurement strategies for value-added assessment. As such, it offers a frame for assessment that, in the words of SCHEV's charge to the task force, focuses on "outcomes, student progress, and academic program improvement" while affording institutions "the flexibility to establish outcomes and concomitant methods of assessment that are appropriate to their student bodies and programmatic missions."

Council staff will employ both this frame and its nomenclature when issuing implementation guidance for the new cycle of assessment set to launch in Fall 2008. And any amendments to this document will be made in collaboration with appropriate institutional representatives.

SIX CORE AREAS

Virginia's colleges and universities will use the six core subject areas established by the 1999 Governor's Blue Ribbon Commission on the Future of Higher Education as the general parameters for conducting value-added assessment. These core areas are as follows:

- Written Communication
- Quantitative Reasoning
- Scientific Reasoning
- Critical Thinking
- Oral Communication
- Information Technology Literacy

Given the changing levels of baseline competency in the sixth core area, Information Technology Literacy, institutions have the option of substituting an emergent institution-specific area of interest for it or of assessing this core area purely as a competency.

RELATIONSHIP BETWEEN COMPETENCY-BASED AND VALUE-ADDED ASSESSMENT

Individuals with a history of assessment in the Commonwealth know that institutional assessment of student learning has concentrated on the *competency* of student performance in the core subject areas as delineated above. Of course, competency-based assessment offers an inherently different perspective from the *value-added* approach as specified by Council. That said, competency and value-added perspectives, though different, are not mutually exclusive. Institutions can and, perhaps, should continue to define, set, and measure standards of performance for their students within a competency framework—incorporating into it a value-added component that builds on what is already a quite strong assessment foundation.

Of course, as with value-added assessment which will be addressed later in this document, competency-based assessment has its own set of advantages and challenges. Highlights of some of them follow.

Competency-Based Assessment

Definition. Competency-based assessment attempts to determine whether or not students have achieved the skills, knowledge, and/or abilities that are considered important to an educational mission, goal, or curriculum. Competencies must be defined clearly and in ways that allow for valid and reliable measurement. An institution *can* develop clear and unambiguous standards for assessing levels of competence, but their development typically evolves over a period of time.

Methodological Implications. Assessment of competency does not require comparisons, either among groups or over time.

Advantages. Developing agreements about competencies requires considerable discussion with the multiple stakeholders but can increase the likelihood that the competencies and the measurement of competencies will be valid. The primary advantage of a competency-based approach is the likelihood that the competencies will be related intentionally to the educational experiences. Accordingly, competency-based assessments often inform decisions about teaching and learning. For example, competency-based measures such as course assignments and tests can be used to show what educational experiences result in the specified learning outcomes. Classification of individual students as having met or not met the required standard of competence may be useful in situations such as licensure exams, as well as to stakeholders such as employers. This approach can include threshold judgments of competence or can accommodate multiple standards such as "competent" and "highly competent."

Challenges. As a result of the institutional distinctiveness of mission, goals, and learning experience, definitions of competencies may vary across institutions. Such variance is consistent with SCHEV's long-standing acknowledgement of institutional diversity as an important strength of the Commonwealth's higher education system. And, while measurement validity is potentially a strength of competency-based assessment, reliability is more problematic.

OPERATING CONCEPTION OF VALUE-ADDED ASSESSMENT

An institution's assessment plan should embody the following operating conception of value-added assessment:

> *Value-added assessment measures indicate progress, or lack thereof, as*
> *a consequence of the student's institutional experience.*

Within this context, the institution must define the methodologies it proposes to employ, providing substantive explanation and justification for each. It must also identify any specific challenges inherent in the proposed method(s), addressing the mechanism for

overcoming them so as to ensure the most rigorous results possible. The following section sets forth a description of methods that constitute the parameters according to which institutional plans may be justified.

SELECTED VALUE-ADDED ASSESSMENT METHODS: ANALYTICAL APPROACHES, DATA COLLECTION, MEASUREMENT STRATEGIES

Analytical Approaches

Professionals designing assessment programs draw from differing assessment approaches, data collection methods, and strategies for measurement. Assessment literature delineates three basic approaches—each responding to a slightly different question, each with its own set of methodological implications, advantages, and challenges. Of course, all three approaches are affected by institution-specific constraints—e.g., selectivity, graduation rates, transfer tendencies—that the college or university will have to consider as it develops its assessment plan.

Value-Added Approach One: Longitudinal

Definition. Longitudinal designs are useful for identifying and describing change over time. They are helpful but not sufficient for establishing causal relationship. (See the discussion of "Advantages" below.) A longitudinal repeated-measures approach occurs when measures are taken for a single sample of individual performances at two different points in time.

Methodological Implications. Longitudinal studies in the context of higher education typically involve pre- and post-test measures of some sample that represents a meaningful population. Ideally, the measurements are taken just before and immediately after the presumed experimental variable, in this case, some educational intervention.

Advantages. The primary advantage of longitudinal research is to increase confidence in conclusions about the effects of the intervention—again, some educational experience. However, in the absence of experimental design and valid and reliable

measurement of both the presumed causal variable (educational experience) and the intended outcome variable (student learning), no valid scientific inference can be made. Additionally, data can be used to assess both inter-individual and intra-individual change.

Challenges.   Attrition makes it difficult to track a single sample of students from freshman to subsequent years.   That problem is particularly severe for some universities and most community colleges. As noted above, longitudinal designs do not necessarily identify the causes of any change (educational interventions), nor do they necessarily involve tested (reliable and valid) measures of student learning.

Value-Added Approach Two: Cross-Sectional

Definition.   Cross-sectional designs involve measurements of samples from different populations at a single point in time. For example, an institution might take performance measures for a group of freshmen and, within the same semester, a group of seniors. In a comparison of the scores, the assumption is that differences are attributable to experiences between the freshman and senior years.

Methodological Implications.   Cross-sectional designs provide less control over sources of variation than do longitudinal designs.   Accordingly, inferences must be drawn more carefully.

Advantages.   Generally, cross-sectional studies are more economical than longitudinal ones. And results can be reported much more quickly because the institution does not need to follow a sample over time.

Challenges.   In a cross-sectional study, attrition is a concern. Furthermore, the cross-sectional design cannot be used to assess individual changes; instead, it can be used to make comparisons between/among groups. Since the groups are distinct, the observed differences may be attributable to differences in the samples or the populations. As with the longitudinal design, critical issues for inference include experimental control and the

validity and reliability of measures of the experimental variable (educational experience) and the dependent variable (student learning).

Value-Added Approach Three: Residual Analysis

Definition. With residual analysis, value added is inferred from the comparison of actual scores and scores that were predicted based on some previous measurement (for example, SAT or ACT).   Residual analysis is used by the Collegiate Learning Assessment (CLA).

Methodological Implications. The main advantage of residual analysis is the savings incurred in both time and money—the result of testing only one group of students. That said, measurement validity with this approach is particularly problematic, especially as currently proposed in the use of CLA and similar instruments. (See "Challenges" below.) Whether the accompanying design is longitudinal or cross-sectional, the challenges with inference are the same as those delineated in the preceding sections.

Advantages. Residual analysis controls for selection bias by using a measure such as the SAT to predict expected scores on a subsequent measure. This measure of change allows flexibility in the research design.   It also allows for comparisons across institutions.

Challenges. With a residual analysis design, pre-test scores are usually estimated by instruments such as the SAT or ACT. The challenge is that these instruments do not measure general education and, as such, might be misleading. Attrition is also a challenge with this value-added approach, as students who dropped out before the post-testing cannot be sampled. Furthermore, if the goal is to compare scores across institutions, the participating schools would need to employ identical methods.   And, in terms of value-added interpretation, selective institutions may experience a ceiling effect with students who had initial scores at the very top of the range not demonstrating much change on standardized tests. Finally, the residual score is simply not as clear as a difference score in longitudinal or cross-sectional design and requires significant additional analyses.

Data Collection

An institution's assessment approach poses no limit on its data collection method. That said, the methods vary in terms of practicality and generalizability back to the institution. Three basic methods for collecting data, along with their inherent advantages and challenges, are delineated below.

Data Collection Method One: Course-Embedded

Definition. Data collected within the classroom could include course assignments and tests or measures designed for specific assessment purposes. As such, these data can serve a three-fold function: provide the traditional basis for faculty to "grade" their students' skills and knowledge, give faculty a built-in vehicle for evaluating their approaches to instruction and course design, and document student learning for broader program and institutional assessment purposes.

Advantages. Course-embedded data collection requires relatively modest faculty time to organize for program or institutional level analyses. Faculty and students can receive feedback to improve teaching and learning. In addition, if the course-embedded assessments are also linked to course grades, student motivation for performance is likely to be relatively high. A course-embedded approach overcomes many of the limitations of proprietary instruments administered outside of the classroom. The course-embedded approach also connects learning experiences directly to evaluation methods. Capstone course assignments and student portfolios are two common examples of course-embedded assessment data. Capstone courses provide program information and institutional assessment data through assignments and exams, while portfolios are collections of artifacts from course experiences.

Challenges. Faculty must develop shared standards and measures. Course-embedded data often are scored or rated by more than one reviewer (e.g., team or department) to enhance validity. To ensure inter-rater reliability, the raters must be trained. Course-embedded data address course-level objectives, whereas program-level outcomes may require modification or elaboration of the measures. Program-level

course-embedded assessments are often most appropriate when embedded in capstone courses or identified through curricular mapping.

Data Collection Method Two: Single Setting

Definition. A single setting may be used to collect data on large group samples or a census of students. Some institutions conduct an "assessment day" on which students across the institution participate in assessment activities.   With the single setting method, the environment for data collection is controlled, and the data collected often extends across courses, sections, and/or disciplines.

Advantages. Large amounts of data can be collected from many students in a brief period of time. Furthermore, administration procedures are uniform within a defined time period and location. This collection method can be used to optimize sampling, whether random or purposive. And when using selected-response tests, scoring can be completed quickly with high levels of reliability. In short, the primary advantages of this method are ease of administration and reliability of measures.

Challenges. The major challenge of the method is to develop valid measures without compromising the advantages cited above. Collecting data in a single setting may not be suitable for all learning objectives and may be more conducive to selected response formats. Care must be taken to align test items with the intended educational experiences and student learning goals.   And there is some tension between maximizing participation in the testing and optimizing the students' stake in the outcomes.   More specifically, if single setting assessments are not linked to student evaluations, student motivation for performance is likely to be relatively low and needs to be monitored.

Data Collection Method Three: Portfolio

Definition.   Portfolio assessment involves collecting and reviewing students' work for evidence of learning. Work samples often are collected across time, and there is an opportunity to assess a student's development across the scope of course or program work.

Advantages. Portfolios involve multiple measures and indicators, enhancing measurement validity. By directly involving faculty in assessment, portfolios facilitate linkages between teaching and learning. Portfolios can be linked to one or more courses and used to assess achievement with respect to learning goals.

Challenges. Portfolio assessment does not emphasize standardization. Faculty members and others can develop rubrics, and raters can be trained to use those rubrics to enhance reliability. This data collection method requires substantial faculty commitment and time, but it yields information that is relevant directly to teaching and learning.

Measurement Strategies

As with data collection, an institution's assessment approach poses no limit on the ways in which educational experiences and student learning are measured. The critical issues continue to be inference, validity, and reliability. At the broadest level, the institution can consider quantitative, qualitative, or hybrid approaches.

Measurement Strategy One: Selected-Response Test

Definition. The selected-response strategy requires respondents to select an answer from a set of alternatives.

Advantages. Selected-response tests can be administered efficiently and results can be quantified easily. As a measurement strategy, selected-response has primary strengths for reliability. Scoring is fast and cheap.

Challenges. The development of valid measures requires significant time and conceptual clarity. Tests must be developed carefully to reduce the effects of "guessing" and to ensure that there is only one "correct" answer. It takes skill to develop such items, particularly where they must tap into higher orders of learning. Moreover, item security is a concern, especially over time. The link between educational experiences and student learning must be induced and is not readily demonstrated. Selected response tests also limit expressions of learning and, as with

any single measure, this approach advantages some students and disadvantages others.

Measurement Strategy Two: Constructed Response Format

Definition. Constructed response formats require students to respond to a stimulus such as a question or a prompt or an assignment. Typical constructed response formats include performance assessments, open-ended tests, essay questions, or products from a senior design course or art portfolio. Constructed response format assessment is dependent on well constructed rubrics or scoring guides that lay out the specific expectations and behaviors for an assignment, a project, or a presentation, for example.

Advantages. Constructed response assessment activities often closely parallel the intended student learning. Accordingly, they are particularly useful for assessing complex learning. Performance observations may reveal how students reasoned to a correct (or incorrect) answer, and results can be used to improve teaching and learning. Well-written prompts and open-ended tests are particularly useful in assessing writing for evidence of critical thinking and good written communication skills. Further, a primary advantage to using rubrics is that they help to make explicit the learning expectations for students in similar learning environments. Rubrics are particularly useful for assessing multiple sections of a course and for courses that cross discipline boundaries.

Challenges. Developing good rubrics requires significant faculty input and instrument refinement. Writing clear, succinct, and behaviorally anchored criteria for each learning objective is typically an iterative process. Questions regarding missing data must be addressed. Raters must be trained, and often the instrument itself must be modified based on rater experience. Rater error is an important issue with this method. While developing and using the strategy requires time and close collaboration with faculty, the results inform teaching and increase the likelihood of connecting assessment and educational decision-making.

CLOSING THE LOOP: USING ASSESSMENT RESULTS TO ENHANCE LEARNING AND DEMONSTRATE ACCOUNTABILITY

Guidelines for Assessment of Student Learning provides a general introduction to assessment's most fundamental issues—namely, inference, validity, and reliability. As such, colleges and universities can use it to develop the value-added component of their assessment plans. And SCHEV staff can use it as the context for reviewing and providing feedback to institutions on their plans.

In July 2007, the State Council released *Advancing Virginia: Access, Alignment, Investment: The 2007-13 Plan for Higher Education in Virginia*. The plan's tenth goal speaks specifically to "strengthen[ing] academic program quality and accountability through assessment" as follows:

> Virginia's colleges and universities must rigorously document whether students are taught what they need to learn in order to be successful in their chosen fields. In the case of degree programs, assessment should proceed according to standards established by accrediting bodies, both disciplinary and regional, and by the institutions' own internal standards and expectations. Information about degree program assessments should be made available to the public.

To be sure, the State Council requirement that Virginia's colleges and universities demonstrate the "added value" of their respective educational experiences to student knowledge, skills, and abilities has the potential to build on the solid foundation in formative assessment that institutions have laid over the past twenty-plus years. And, clearly, the public has a right to know that its higher education institutions are making the best possible use of taxpayer dollars. But care must be taken to ensure that institutions not sacrifice the proven formative value of their existing assessment strategies for ensuring learning and improving the quality of teaching and learning. Indeed, the institutional blueprint for incorporating a value-added component into

assessment's foundational structure must be both careful and deliberate—else nobody wins.

APPENDIX ONE: SELECTED ANNOTATED BIBLIOGRAPHY ON ASSESSMENT

Angelo, T.A. and Cross, K.P. *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco, CA: Jossey-Bass, 1993.

> A classic text for faculty. Offers ideas for targeted classroom assessment.

Astin, A. *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*. Portland, OR: Onyx in collaboration with American Council on Education, 1996.

> An excellent discussion of the philosophical and conceptual underpinnings of assessment.

Banta, T. *Assessment in Practice: Putting Principles to Work on College Campuses*. San Francisco, CA: Jossey-Bass, 1996.

> Edited volume focused on best practices in assessment.

Diamond, R. *Designing and Assessing Courses and Curricula: A Practical Guide*. San Francisco, CA: Jossey-Bass, 1998.

> A step-by-step approach to assessment.

Erwin, T. *Assessing Student Learning and Development*. San Francisco, CA: Jossey-Bass, 1991.

> Guide for developing an assessment program in both the academic and student affairs areas.

Fulcher, K. and J. Willse. "Value-added: Changing Back to Basics." *Assessment Update* (forthcoming).

> Explores issues of reliability and validity in assessing change.

Jaeger, R., ed. *Complementary Methods for Research in Education*. 2nd ed. Washington, DC: American Educational Research Association, 1997.

> Outlines various approaches to research education.

Klein, S., R. Shavelson, and R. Benjamin. "Setting the Record Straight." *Inside Higher Education (February 8*, 2007*)*. http://www.insidehighered.com/views/2007/02/08/benjamin (accessed September 2007).

> Explores issues related to the use of portfolios.

Menard, S. *Longitudinal Research.* Newbury Park, CA: SAGE Publications, Inc., 1991.

> Primer on longitudinal research design.

Palomba, C. and T. Banta. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco, CA: Jossey-Bass, 1999.

> Useful guide for developing an assessment program.

Pastor, D., P. Kaliski, and B. Weiss. "Examining College Students' Gains in General Education." *Research and Practice in Assessment* (June 2007). http://www.virginiaassessment.org/rpa/2/Pastor%20Kaliski%20Weiss.pdf (accessed September 2007).

> Good example of applied value-added assessment.

Stevens, D., and A. Levi. *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning*. Sterling, VA: Stylus Publishing, 2005.

> Good primer on assessment, with examples for developing course scoring guides.

Suskie, L. *Questionnaire Survey Research: What Works?* Tallahassee, FL: Association for Institutional Research, 1996.

> An excellent primer for conceptualizing, creating, and conducting surveys.

_____. *Assessing Student Learning: A Common Sense Guide*. Bolton, MA: Anker Publishing Co., 2004.

> A comprehensive review of assessment. Includes very good examples for assessing student learning.

APPENDIX TWO: ASSESSMENT FRAMEWORK

Definition of Terms

1. Identify the core area that is to be assessed. How is it defined? How does it relate to the institution's mission and goals?

2. Describe learning objectives for the defined core area.

Methodology

1. Describe the measurement strategy (e.g., selected-response test).

2. Describe the expected relationship between the educational experiences (independent variable) and student learning (dependent variables).

3. Discuss the reliability and validity of the measures.

4. Describe the data collection method (e.g., course-embedded or single setting).

5. Describe the analytical approach (e.g. longitudinal or cross-sectional) and justify the methodological choices. If a competency approach is employed, describe the procedure used to determine the standards.

Process Evaluation

1. Describe the results and how they were disseminated.

2. Describe how the results have been and will be used to improve educational experiences (independent variables) and advance the mission of the college or university.

3. Estimate the costs of the assessment.

APPENDIX THREE: SELECTED WEB RESOURCES

1. http://www.schev.edu/higherEd/AssessmentResources.asp

    A list of assessment Websites for public colleges and universities in Virginia.

2. http://www.ncsu.edu/assessment

    Assessment Website for North Carolina State University, with many links to assessment resources.

3. http://www.aacu.org/gex/paa/assessment.cfm

    Accreditation and assessment Website of the American Association of Colleges and Universities.

4. http://www.aera.net

    Website of the American Educational Research Association.

5. http://www.ncme.org

    Website of the National Council on Measurement in Education.

6. http://www.PAREonline.net/Home.htm

    A peer-reviewed electronic journal entitled *Practical Assessment, Research and Evaluation*. Is a highly credible source for information on virtually every aspect of assessment. Includes sections on authentic assessment, as well as on rubrics and validity, scoring rubrics, and guidelines for multiple choice.

7. http://www.virginiaassessment.org/RPAJournal.php

    Virginia Assessment Group's *Research and Practice in Assessment Journal*.

8. http://www.rdg.ac.uk/ssc/publications/guides/toptgs.html

    A statistical "good practices" guide published under the auspices of the University of Reading's Statistical Services Centre in the United Kingdom.

9. http://edweb.sdsu.edu/webquest/rubrics/weblessons.html

    A web lesson in rubrics as a tool of authentic assessment.

10. http://www.chebucto.ns.ca/~LilleyS/tips.html

    Ten tips for evaluators on "how to deliver negative evaluation results constructively."