

No. 24

Assessment in WIA Youth Programs

Cheryl R. Sturko Grossman
Michael E. Wonacott
Diana Jackson

2007



Youthwork Information Briefs are sponsored by Ohio Department of Job and Family Services - ODJFS, Office of Workforce Development, Bureau of Workforce Services.

Why Learn about Assessment?

Assessment plays a key role in fulfilling the purpose of the Workforce Investment Act (WIA), which is to help at-risk youth make successful transitions to further education and employment in high-demand, high-wage occupations and careers.

In the first stage of a youth's involvement in WIA program activities, an objective assessment is mandated to identify important baseline information about the youth:

- Academic levels, including basic skill levels
- Skill levels, including occupational skills
- Supportive service needs
- Prior work experience, employability, interests, and aptitudes, including interests and aptitudes for nontraditional jobs
- Developmental needs

Assessment of WIA youth participants is not a one-time task, done and then forgotten. Rather, the results of assessment are used to design a specific program of WIA youth services, the Individual Service Strategy (ISS), that will address the youth's barriers and needs and lead to individualized educational and employment goals.

In addition, the results of objective assessment are a baseline against which the youth's progress can be measured; they serve as a pre-test. As the youth receives the services listed in the ISS, ongoing assessment identifies gains and improvements in academic levels, skill levels, and developmental needs. If there are gains and improvements, the ISS should evolve to meet the youth's evolving needs, ensuring that ongoing services remain appropriate to the youth.

Finally, statutory WIA Performance Measures and new Common Measures require periodic reassessment of youth skills:

- Younger Youth Skill Attainment Rate, under statutory Performance Measures, identifies and documents gains in basic skills, occupational skills, and work readiness skills by all in-school youth and by out-of-school youth assessed to be in need of those skills.
- Literacy/Numeracy Gains, under Common Measures, identifies and documents literacy/numeracy skill gains by out-of-school youth who are basic skills deficient.

Key Points

Useful Resource

LearningWork Connection's *Youth Assessment Guide: The Right Test at the Right Time* provides basic information about more than 300 tests to aid youth-serving professionals in the selection, administration, and interpretation of assessments.

Initial Assessment and Objective Assessment

Following intake, WIA youth staff gather and document information to identify the youth's barriers and to determine if WIA youth services are appropriate. Information is also gathered to help determine a youth's needs for supportive services such as transportation, child care, dependent care, medical services, housing, and work clothing or tools. This process is initial assessment, a form of informal assessment, and does not initiate participation in WIA.

Objective assessment, on the other hand, does initiate participation in WIA and involves formal assessment. There are many types of formal assessments used in WIA youth objective assessment. The most familiar type of formal assessment is the achievement test. This type of assessment measures what a youth already knows or can do. Examples include a math test or a basic skills test.

Another type of achievement test is a skill or performance test, which requires test-takers to demonstrate a skill or procedure, rather than just demonstrate knowledge about it. An occupational skills test is one example.

Other formal assessments measure traits, characteristics, interests, or aptitudes related to work and careers, like a career interest survey, a vocational aptitude battery, a life skills inventory, a learning style assessment, a measure of depression, or an intelligence scale.

The Ohio Department of Job and Family Services (ODJFS) mandates the use of specific basic skills assessments for objective assessment of WIA out-of-school youth; ODJFS also strongly recommends a specific instrument, TABE 9/10, for the required basic skills assessment for in-school youth. Local Workforce Investment Boards (WIBs) are free to select appropriate assessments for traits, characteristics, interests, or aptitudes related to work and careers. Assessments given by schools or other agencies up to six months prior to date of participation may be used for literacy/numeracy measures for both older and younger youth.

Key Points

Timing of Basic Skills Assessments

- Appropriate basic skill assessments given six or fewer months before the date of participation may be used for an objective assessment of literacy/numeracy skills for older youth.
- Objective assessment, such as for basic skills, should be conducted promptly after eligibility determination; objective assessment initiates program participation.

Key Points

Acceptable Basic Skills Assessments

Out-of-School Youth. Unless local areas receive permission from ODJFS, they must use one of the following assessments that have been crosswalked to the National Reporting System (NRS) educational functioning levels (EFLs):

- Comprehensive Adult Student Assessment System (CASAS)
- Tests of Adult Basic Education (TABE) (7/8 or 9/10)
- Adult Basic Learning Examination (ABLE)
- WorkKeys
- Student Performance Level (SPL) – English as a Second Language (ESL) only
- Basic English Skills Test (BEST) – ESL only (TEGL 17-05, Attachment C)

In-School Youth. Any assessment can be used that measures the basic skills required. However, local areas are strongly recommended to use the TABE 9/10 because it is most in line with the Ohio General Educational Development Tests (GED) and also incorporates workplace literacy components.

Why Is Assessment a Challenge?

WIA youth staff may not be specialists in assessment, evaluation, or psychometrics and may find assessment to be a challenge. What is the right assessment to use? How should it be used? Who should administer it? How should it be administered? Who should interpret it? How should it be interpreted? How should it be used?

Fortunately, assessment developers and publishers provide a great deal of information and guidance for users of assessments such as WIA staff. The following are some key questions that should be addressed for each assessment to be conducted:

- Is this an appropriate assessment?
- How should this assessment be administered?
- How should this assessment be interpreted and used?
- How do I ensure that this assessment is conducted ethically?

Is This An Appropriate Assessment?

Formal assessments are used to measure current skill levels, determine employment-related interests and aptitudes, and diagnose developmental characteristics and needs.

- **To measure a youth's current skill level.** A youth's academic skills, basic skills, occupational skills, and work readiness skills should be measured:
 - At the beginning of program participation as the basis for planning services.
 - Periodically during program participation to identify skill gains.
 - After program participation to determine final outcomes.
- **To determine a youth's employment-related interests and aptitudes.** What kinds of work tasks appeal to a youth? What fields of work interest the youth? How about specific occupations? What is the youth good at? What characteristics and traits does the youth have, and what occupations or careers are suited to those? Answers to questions like those help youth select realistic educational and employment goals.
- **To diagnose a youth's disabilities or developmental characteristics and needs.** Physical, cognitive, and emotional/behavioral disabilities must be carefully accommodated in planning and providing WIA youth services. Other developmental

Key Points

Finding Out More About an Assessment

The most comprehensive source for information about an assessment is the publisher. Publishers employ professional psychometricians with expertise and experience in technical assessment issues and detailed knowledge of past and ongoing studies and research on the assessment in question.

You may also want to look at independent evaluations of an assessment, especially if you are concerned about the appropriateness of the assessment for the WIA youth population. One of the best sources for independent research on specific assessments is the Buros Institute of Mental Measurements and its publications *Tests in Print* and *Mental Measurements Yearbook*.

These guides periodically commission expert reviews that evaluate the reliability, validity, norming process, and appropriate uses of assessments. The guides also gather citations to papers in the research literature about assessments. Such papers often provide additional validity or reliability evidence or may uncover problems with an assessment.

characteristics or needs (e.g., life skills, maturity, self-confidence) can indicate a need for particular program services.

Although WIA youth staff may administer a pre-screening assessment for learning disabilities (with appropriate training), only qualified specialists should make a formal diagnosis of disabilities. In other cases that call for formal assessment, WIA youth staff can use the following guidelines to select an appropriate assessment:

- Use a criterion-referenced assessment, if possible.
- Use a standardized assessment.
- Use an assessment with evidence of reliability.
- Use an assessment with evidence of validity.
- Use an assessment designed for the appropriate target population.
- Use the same criterion-referenced assessment for pre-test and post-test.

Use a Criterion-Referenced Assessment, if Possible

Assessments are usually designed to be scored in one of two ways.

- **Criterion-referenced assessments** show how a test-taker scores against an absolute standard. Usually, that criterion consists of a whole set of detailed standards against which the test-taker is measured, like occupational certification examinations, state graduation exams, or a driver's license examination. Classroom tests that measure how well students have mastered the material taught are also criterion-referenced assessments. One example is WorkKeys.
- **Norm-referenced assessments**, on the other hand, show how one test-taker's results compare to the results of other test-takers or how one group of test-takers compare to another group. They rank test-takers by comparing one test-taker's performance to that of a sample population of similar test-takers (called a norm group). Some common norm-referenced tests include the Iowa Tests of Basic Skills, the California Achievement Test, the Stanford Achievement Test, the SAT (formerly the Scholastic Aptitude Test; since 2005, the SAT Reasoning Test), and the ACT (originally, the American College Testing Program).

If possible, WIA youth programs should use criterion-referenced assessments to identify youth's academic levels, including basic skills levels, and skill levels, including occupational skills. Criterion-referenced assessments provide detailed information on which specific skills youth have and which they don't have. Norm-referenced assessments, on the other hand, only tell you how one youth compares to other youth – they don't tell you which skills the youth has and which skills the youth

Key Points

Criterion- and Norm-Referenced Assessments

Use criterion-referenced assessments to:

- Determine specifically what a youth knows and is able to do.
- Determine whether youth have met standards.
- Place individual youth in programs based on instructional levels.
- Measure gains in achievement over time.
- Determine program effectiveness.
- Assess program effectiveness.

Use norm-referenced assessments to:

- Rank youth.
- Compare results across different groups of youth – e.g., compare individual or group performance to national or state norms.

lacks. So, norm-referenced assessments don't provide the detailed information necessary to identify specific services needed or to measure skill gains over time.

Some assessments, like the TABE 9/10, are both criterion-referenced and norm-referenced.

Use a Standardized Assessment

An assessment is standardized when every test-taker has an equal opportunity to accurately demonstrate his or her knowledge or other characteristic being measured. Developers achieve this, in part, by examining individual items and the assessment as a whole to eliminate any possible racial, ethnic, cultural, linguistic, or socioeconomic bias. The assessment is normed against a representative sample of the target population to verify that items are not biased and to provide statistical evidence of quality.

Developers also provide standard instructions for administering the assessment. Instructions are often elaborate, including scripts that must be read verbatim. This ensures that all test-takers are receiving the same directions – if directions are ambiguous, at least they are ambiguous for everyone!

Instructions may also recommend adjustments in assessment procedures for youth with specific disabilities or for youth who speak English as a second language (ESL), along with statistical evidence of the quality of such adjustments. Finally, specific and detailed instructions are provided on scoring the assessment and interpreting results.

Use an Assessment with Evidence of Reliability

Reliability is a critical technical characteristic of a good assessment; an assessment is reliable if it gives consistent results – over time, across different administrations, or across different raters. Positive correlations (+0.6 or greater) provide statistical evidence of reliability.

Key Points
Evidence of Reliability
For high-stakes outcomes such as placement in a program, the reliability of an assessment should be above +0.8, preferably in the +0.9 range.

Use an Assessment with Evidence of Validity

Validity is another critical technical characteristic of a good assessment; an assessment is valid if it accurately measures what it is intended to measure. Assessments are only valid for the purposes specified by the developer. An assessment that is valid for comparisons between individual test-takers may not be valid if used to compare programs unless the developer has validated it for that purpose.

Expert judgment. The assessment should be a representative sample of the domain of knowledge and skills being assessed, and that is a matter of expert judgment. So, look for an assessment developed or endorsed by authoritative, acknowledged content experts in the domain.

Statistical evidence. Statistical evidence can also demonstrate that an assessment is valid.

Key Points
Evidence of Validity
<ul style="list-style-type: none">• Developed or endorsed by acknowledged, authoritative experts in the area.• Statistical correlations between +0.6 and +1.0 with:<ul style="list-style-type: none">o Other valid, reliable assessmentso Desired future outcomes – e.g., academic success in college, successful performance on the job.• Statistical correlations between -0.6 and -1.0 indicate a negative relationship between two variables.• Logical, intuitive appeal – especially to youth who must take the assessment seriously.

A valid assessment should show a positive correlation (as one factor increases or decreases so does the other) between +0.6 and +1.0 with:

- Other valid, reliable assessments in the area.
- Desired future outcomes in the area – for example, job satisfaction might correlate with a career interest survey; academic success in college, with a basic skills assessment.

On the other hand, a negative correlation, between -0.6 and -1.0, would occur if one factor increases (for instance, a score on a coping skills assessment) as another decreases (for instance, aggressive behavior). Such a negative correlation can also be evidence for validity.

Logical and intuitive appeal. Youth sometimes may not see the connection between the assessment items and the knowledge or skills being assessed. If the items don't have a logical, intuitive appeal to youth, youth may:

- Be suspicious of the assessment, deny its importance, and refuse to accept its results.
- Be frightened and experience anxiety – “test anxiety” – that can artificially lower their results.
- Not take the assessment seriously and not try to do well, leading to lower results.

An assessment administrator should explain the underlying rationale for the assessment to help youth accept its validity.

Use an Assessment Designed for the Appropriate Target Population

Any assessment selected should be designed specifically to be used by youth like the WIA youth population – ages 14 -21, from a variety of racial, ethnic, cultural, and socioeconomic backgrounds, and with barriers to success.

For example, consider a basic skills reading assessment. If the assessment is written for adults, many items on the assessment may involve aspects of independent adult living – voting, purchasing a house, personal finances, parenting, and so on. A

16-year-old high school dropout might score low on the assessment because the content of the items is unfamiliar – not because he or she doesn't have the required reading skills.

Finally, some assessments are available in both a short form and a long form; using the short form can help youth perform better. With a long form, youth can lose concentration and become fatigued – or even become so discouraged they exit the program prematurely, especially if they have had bad testing experiences in the past.

Use the Same Criterion-Referenced Assessment for Pre-Test and Post-Test

A pre-test is used to identify a youth's skill level at one point in time; a post-test identifies the youth skill level at a later point. The difference between pre-test results and post-test results is the increase in the youth's skill – the skill gain. In order for those results and gains to be meaningful, the same assessment must

Key Points

Assessments and Populations

- If possible, use an assessment that has been normed for the specific youth to be assessed – e.g., race, ethnicity, age, language.
- Use an assessment with norming data no more than 15 years old.
- Use a short form if one is available.

Key Points

Pre-Test and Post-Test

- Always use the same assessment for both.
 - Consistent with good assessment methodology.
 - Consistent with WIA requirements.

be used for both pre-test and post-test. That assessment must also be criterion-referenced in order to provide specific details on skills gained.

How Should This Assessment Be Administered?

Staff who administer an assessment must be familiar with the administration manual and follow explicit instructions carefully. The instructions are designed to ensure valid and reliable results. To avoid biasing assessment results, all youth must always receive the same exact, full instructions.

Format, Materials, and Equipment

Youth should have an opportunity to familiarize themselves with the format of the assessment and its items and any materials or equipment being used (calculator, scientific apparatus, computer). Computers are increasingly being used for assessment and may be an attractive option for youth.

Most youth are comfortable with computers, but keyboarding, mouse use, and reading lengthy material from a screen may all detract from the knowledge and skills being assessed. If computers are used, policies and procedures should be in place to deal with software, hardware, or power failures.

Assessments typically provide sample items that test-takers can review, so administrators should go over those sample items to familiarize youth with format and equipment use. However, actual items should never be reviewed – sample items only!

Environmental Factors

Environmental factors can interfere with test-takers' ability to demonstrate their true knowledge or skills so the physical environment should be controlled as much as possible:

- Set the thermostat at a comfortable level – not too hot, not too cold.
- Eliminate noise or other distractions.
- A clock should be visible during a timed assessment.
- Make snacks and water available during a long assessment.

Administration procedures should be adjusted per assessment instructions to accommodate the needs of youth with disabilities or ESL youth; however, such accommodations should be made in assessment only

Key Points

Conducting Assessments

- Always follow instructions to the letter when administering, scoring, and interpreting assessments.
- Adjust administration procedures as appropriate for youth with disabilities and ESL youth.
- Aside from accommodating youth with disabilities and ESL youth, always treat all test-takers the same – same time, same instructions, same information, same conditions.
- Explain the underlying rationale for the assessment, especially if it does not have logical and intuitive appeal to youth.
- Allow youth to get familiar with the assessment formats, materials, and equipment.
- Review sample items – not actual items.
- If allowed by standardized instructions, break a long assessment into shorter segments to avoid discouragement and fatigue.
- Adjust environmental factors to maximize success:
 - o Temperature and ventilation
 - o Noise
 - o Clock on the wall
 - o Snacks
- Be sure youth know how to prepare:
 - o A good night's sleep
 - o A good breakfast
 - o Low stress level

if they have also been made during instruction. Additional information on accommodations for youth with disabilities is provided in TEGL 17-05.

How Should This Assessment Be Interpreted And Used?

Assessments should be scored and results should be interpreted and reported as prescribed in the developer’s instructions.

Criterion-Referenced Assessment Results

Raw score. The basis for assessment interpretation is the raw score – usually how many items were answered correctly. The raw score provides only limited information; a score of 20 may be very good or very bad.

Percentage score. In order to make judgments about performance, raw scores are converted into percentages for criterion-referenced assessments. A percentage score is calculated using the number of correctly answered items divided by the total number of items; for instance answering 18 of 20 items correctly achieves a 90 percent score.

The percentage shows how closely an individual approaches the criterion or standard.

Cut score. Cut scores, or passing scores, specify the score that is considered proficient for the assessment – a “passing grade,” so to speak. Cut scores may be set using raw scores, percentages, or percentiles.

Key Points

Common Scores

- Raw score – the number of items correct
- Percentage – the proportion of all items answered correctly
- Cut score – required level of proficiency, a “passing grade”

Norm-Referenced Assessment Results

Norm-referenced assessments convert raw scores into percentiles, stanines, and grade equivalents.

Percentile. The percentage differs from a percentile, or percentile rank, used with norm-referenced assessments. A percentile is calculated by determining the percent of scores within a group that fall below a given raw score; for instance, a score in the 90th percentile indicates that 90 percent of other test-takers received the same or lower score. Percentile ranks are not always evenly distributed and do not necessarily indicate quality of response. A test-taker could be in the 99th percentile and have correctly answered only 50 percent of the items, if everyone else in the norm group answered 50 percent or fewer of the items correctly.

Stanines. Because percentiles are not necessarily evenly distributed, scores are often reported in stanines. A stanine is based on a bell curve, or normal distribution of scores, divided into nine units (1 to 9); a score of five indicates average performance. Like all norm-referenced scores, stanines compare an individual score to the scores of other test-takers in a reference group.

Grade equivalent. Another way that scores are expressed is the grade equivalent – for example, a

Key Points

Norm-Referenced Scores

- Percentiles place an individual’s score within a rank order with other individual scores.
- Stanines group scores to ease comparisons.
- Grade equivalents compare scores with average scores for reference groups in each grade and month.

7.6 grade equivalent score on a math test. However, a 7.6 grade equivalent score does not mean that the youth has math skills appropriate to the 6th month of grade 7. Rather, it means that the youth's raw score was average for youth in the norm group who were in the 6th month of grade 7 – whether those youth had math skills appropriate to that grade level or not.

Educational Functioning Levels

Under Common Measures, annual literacy and numeracy skills gains for out-of-school youth must be reported in educational functioning levels (EFLs). Each EFL is equal to approximately 2-3 grade levels. WIA youth programs must use basic skills assessments crosswalked to EFLs for Common Measures reporting. Please see TEGL 17-05, especially Attachment C, for more information regarding EFLs.

Interpreting Assessment Results

Be wary of both perfect scores and scores that approach chance (scores that could be achieved by random guessing); neither of those scores provides an accurate result. A youth who receives a perfect score may be taking an assessment that is too easy; a youth whose scores hover around 50 percent may be taking an assessment that is too difficult. Reassessment with the appropriate level will provide better information about the youth's specific needs. Some assessments, such as the TABE 9/10, offer brief Locator Tests that help determine the appropriate level to administer.

In some case, it is legitimate to question the accuracy of a youth's score. If the results of an assessment are inconsistent with other measures of performance such as school grades or another assessment, or if you know that the youth was ill or unduly stressed during the assessment, interpret the results with caution, or, if possible, allow the youth to retake the assessment.

Using Assessment Results

Use both formal and informal assessment results to aid in developing an ISS and to help youth make educational and career plans. Informal assessment can be incorporated into all staff interactions with youth. Basic skills assessments must be given at least annually; other objective assessments may be given when youth reach an appropriate goal or milestone.

Remember that an assessment, no matter how rigorously developed, administered, and interpreted, only provides limited information about a youth. **Never make important decisions solely on the basis of the results of a single assessment;** also use other methods and results to makes important decisions about a youth's capabilities, interests, and needs.

Key Points

EFLs

Educational functioning levels (EFLs) must be used to report literacy/numeracy skill gains for out-of-school youth under Common Measures.

Key Points

Interpreting Assessment Results

- Don't interpret scores in a vacuum.
- Perfect or nearly random scores may indicate that the assessment is too easy or too difficult for the youth.
- Scores that differ greatly from grades or other indications of ability should be interpreted with caution.

Key Points

Using Assessment Results

- Use both informal and formal assessments to develop an ISS. Assessment should take place as needed thereafter.
- Never make important decisions based solely on the results of a single assessment.

How Do I Ensure That This Assessment Is Conducted Ethically?

Assessment results have consequences for WIA youth programs. Local area youth programs are held to performance standards, under both original WIA Performance Measures and new Common Measures. Two of those performance standards, the Younger Youth Skill Attainment Rate and Literacy/Numeracy Gains, involve specific skill gains that are identified through assessment. Since local areas can be sanctioned for poor performance or rewarded for good performance, temptations can occur.

In addition, it is only human to want youth to succeed, and that understandable desire can lead to “help” that defeats the purpose of the assessment. For example, teaching narrowly to the test may improve results in the short term, but important knowledge and skills for long-term success may be missed. WIA youth staff should never engage in unethical practices in assessment, such as:

- **Excluding poor performers.** Not assessing youth who are expected to do poorly can raise average scores. But those averages are artificial and dishonest, and excluded youth don't get the benefit of a good assessment to tell them how they are performing.
- **Providing assistance during the assessment.** Assistance should only be provided exactly as specified in the instructions. Unless instructions specifically allow it, staff may not define words or make comments about the quantity or quality of youths' work.
- **Cheating.** Cheating includes providing answers, changing response sheets, copying and sharing assessment materials, and offering false or misleading interpretations of results. These practices are often illegal as well, and sanctions for assessment administrators can include loss of job, criminal fines, and jail time.
- **Copyright infringement.** Copyrighted assessments may not be copied and used without the permission of the publisher. Publishers can and do sue to protect copyrights.

Remember that the ultimate purpose of assessment is to accurately demonstrate what youth know and can do. Information from assessment results can help improve instruction offered to youth and other programming. The best way to ensure good assessment results is to provide an effective program.

Key Points

Ethics and Assessment

- Instructors should not teach to the test.
- Assess all youth as appropriate – even if that leads to lower average scores.
- During an assessment, never provide assistance unless specifically allowed by the assessment instructions.
- Never allow youth to cheat.
- Never help youth cheat – no passing out answers or sharing assessments beforehand.
- Never change youth's responses sheets.
- Never misrepresent assessment results or interpret them falsely.
- Don't use copyrighted assessments without the publisher's permission.

Sources

- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: author.
- Baker, E.L., Linn, R.L., Herman, J.L., & Koretz, D. (2002). *Standards for educational accountability systems*. Policy Brief 5. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved March 15, 2006, from <http://www.cse.ucla.edu/products/newsletters/polbrf54.pdf>
- CTB/McGraw-Hill. (2003). *Beyond the numbers: A guide to interpreting and using the results of standardized achievement tests*. Monterey, CA: Author. Retrieved February 20, 2007, from http://www.ctb.com/media/articles/pdfs/ResearchArticles/BeyondtheNumbers.pdf?FOLDER%3C%3Efolder_id=2534374302134983&bmUID=1171993798059
- Employment and Training Administration, U.S. Department of Labor. (2006, February 17). *Common Measures policy for the Employment and Training Administration's (ETA) performance accountability system and related performance issues*. Training and Employment Guidance Letter 17-05. Washington, DC: Author. Retrieved March 22, 2006, from http://wdr.doleta.gov/directives/corr_doc.cfm?DOCN=2195
- Ellis, B.B., & Raju, N.S. (2003, August). Test and item bias: What they are, what they aren't, and how to detect them. Chapter 7 in J.E. Wall & G.R. Walz, Eds., *Measuring up: assessment issues for teachers, counselors, and administrators*. Greensboro, NC: ERIC Counseling and Student Services Clearinghouse. Retrieved April 17, 2006, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/23/2d/c4.pdf
- Gronlund, N.E. (2006). *Assessment of student achievement*. (8th ed.) Boston, MA: Pearson.
- Harcourt Assessment, Inc. (2005). *Some things parents should know about testing: A series of questions and answers*. San Antonio, TX: Author. Retrieved November 17, 2005, from <http://harcourtassessment.com/haiweb/Cultures/en-US/Resources/Additional+Resources/Some+Things+Parents+Should+Know.htm>
- McMillan, J. H. (2000, November). *Basic assessment concepts for teachers and school administrators*. ERIC/AE Digest. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation. Retrieved April 20, 2006, from <http://edweb.sdsu.edu/Courses/Ed791BCF02/ED447201.pdf>.
- Friedman, M.I., Hatch, C.W., Jacobs, J.E., Lau-Dickinson, A.C., Nickerson, A.B., & Schnepel, K.C. (2003). *Educators' handbook on effective testing*. Columbus, SC: Institute for Evidence-Based Decision-Making in Education, Inc.
- Hammill, D.D., Brown, L., & Bryant, B.R. (1992). *A consumer's guide to tests in print*. (2nd ed.). Austin, TX: PRO-Ed, Inc.
- Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: author. Retrieved March 21, 2006, from <http://www.apa.org/science/fairtestcode.html>
- Keith, T.Z., & Reynolds, C.R. (2003). Measurement and design issues in child assessment research. Chapter 4 in C.R. Reynolds & R.W. Kamphaus, Eds., *Handbook of psychological and educational assessment of children: Intelligence, aptitude and achievement*. (2nd ed.). New York: Guilford Press.
- Power, M.A., ed. (2001). *Ethical standards in testing: Test preparation and administration*. (Rev. ed.). WERA Professional Publications Volume 1. University Place, WA: Washington Educational Research Association. Retrieved March 23, 2006, from http://www.wera-web.org/pages/publications/WERA_Test_Ethics.pdf
- Rotherham, A. J. (2006, July). *Making the cut: How states set passing scores on standardized tests*. Washington, DC: Education Sector. Retrieved July 28, 2006, from http://www.educationsector.org/analysis/analysis_show.htm?doc_id=385844
- Rudner, L.M., & Schafer, W.D. (2001, April). *Reliability*. ERIC Digest. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation. Retrieved March 24, 2006, from <http://www.ericdigests.org/20022/reliability.htm>
- Rumsey, D. (2003). *Statistics for dummies*. Hoboken, NJ: Wiley Publishing.
- Salkind, N.J. (2006). *Tests and measurement for people who (think they) hate tests and measurement*. Thousand Oaks, CA: Sage Publications.

- Stuart-Hamilton, I. (1995). *Dictionary of psychological testing, assessment, and treatment*. London: Jessica Kingsley Publishers.
- Walsh, W.B., & Betz, N.E. (2001). *Tests and assessment*. (4th ed.) Upper Saddle River, NJ: Prentice-Hall.
- Zenisky, A.L., & Sireci, S.G. (2005, April 15). *No adult left behind either: Creating large-scale computer-based tests for adult basic education students*. Center for Educational Assessment Report No. 563. Paper presented at the annual meeting of the American Educational Research Association. Amherst, MA: Center for Educational Assessment. Retrieved March 21, 2006, from http://people.umass.edu/azenisky/AERA_2005_Z&S_final.pdf

Inspiring great youth work

LEARNINGWORK CONNECTION

LearningWork Connection is an initiative of the
Center for Learning Excellence at The Ohio State University.

614/292-8665 | e-mail: learningwork@osu.edu
www.learningworkconnection.org

