NPEC Sourcebook on Assessment: Definitions and Assessment Methods for Communication, Leadership, Information Literacy, Quantitative Reasoning, and Quantitative Skills

**Report of the National Postsecondary
Education Cooperative Working Group on
Student Outcomes Sourcebook**

NPEC Sourcebook on Assessment: Definitions and Assessment Methods for Communication, Leadership, Information Literacy, Quantitative Reasoning, and Quantitative Skills

Report of the National Postsecondary Education Cooperative Working Group on Student Outcomes Sourcebook

The National Postsecondary Education Cooperative (NPEC)

NPEC is a voluntary partnership of representatives from postsecondary institutions, associations, government agencies, states, and other organizations with a major interest in postsecondary education. Its mission is to promote the quality, comparability, and utility of postsecondary education data and information that support policy development at the federal, state, and institution levels. The National Center for Education Statistics (NCES) established NPEC and provides resources to support its activities. NPEC receives guidance for its activities from various working groups and committees that are composed of individuals from throughout the United States.

NPEC publications do not undergo the formal review required for standard NCES products. The information and opinions published in them are the products of NPEC and do not necessarily represent the policy or views of the U.S. Department of Education or NCES.

June 2005

This publication and other NPEC publications may be found at the NCES website:

The NCES World Wide Web Home Page is *http://nces.ed.gov*
The NCES World Wide Web Electronic Catalogue is *http://nces.ed.gov/pubsearch*
The NPEC World Wide Web Home Page is *http://nces.ed.gov/npec*

Suggested Citation

National Postsecondary Education Cooperative. *NPEC Sourcebook on Assessment: Definitions and Assessment Methods for Communication, Leadership, Information Literacy, Quantitative Reasoning and Quantitative Skills* by Elizabeth A. Jones and Stephen RiCharde, NPEC 2005-0832. Washington, DC: 2005.

For ordering information on the report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

Or call toll free 1-877-4ED-PUBS; or order online at *http://www.edpubs.org*

Contact:
Nancy Borkow
202-502-7311
nancy.borkow@ed.gov

FOREWORD

Faculty, instructional staff, and assessment professionals are interested in student outcomes assessment processes and tools that can be used to improve learning experiences and academic programs. How can students' skills be assessed effectively? What assessments measure skills in communication? Leadership? Information literacy? Quantitative reasoning?

The NPEC Sourcebook on Assessment: Definitions and Assessment Methods for Communication, Leadership, Information Literacy, Quantitative Reasoning, and Quantitative Skills is a compendium of information about commercially developed instruments used to assess those skills, including costs, content, reliability and validity, strengths, and limitations of various assessments. In addition, the Sourcebook examines definitions and important outcomes in each of these areas and cites resources that provide more in-depth information about these issues.

The primary audiences for this publication are faculty, assessment professionals, institutional researchers, and others who are involved in selecting assessments and developing assessment processes. Policymakers, including professional accrediting agencies and state-level boards, may also find this to be a valuable resource.

NPEC's sourcebooks on student outcomes assessments have certain limitations. They describe tests that are designed primarily for traditional students and do not describe such "nontraditional" assessment methods as portfolios and competencies. The information in the sourcebooks is time sensitive and may change. For example, the costs of instruments will likely increase, and companies that publish instruments may merge resulting in different contact information. Additionally, evaluations of the tests are based on the way the developers market them and on third-party test reviews.

The tests and assessments that are reviewed in this sourcebook were identified by the authors through careful research and consideration. They are a sampling of the numerous possible instruments, rather than a comprehensive list of all that were available. In the views of the authors, they are representative of available tests and assessments because multiple sources have cited them as being useful to postsecondary institutions and   most relevant to the outcomes under consideration.

We would also like to emphasize that all comments about, and reviews of, particular tests or assessments in this publication are descriptive and based on available information. They were not intended, nor should they be construed, as a recommendation for any particular test or assessment.   Rather, a prospective user should take into account the characteristics of the tests and assessments as reported here and should judge their appropriateness and validity for his or her own particular circumstances.

NPEC has a number of other products that address student outcomes. *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing* (2000), which was created by T. Dary Erwin, is a compilation of assessments that measure three of these student outcome domains. This first volume is designed to help institutions and states select the appropriate methods that assess the relevant cognitive outcome. *The NPEC Sourcebook on Assessment, Volume 2: Selected Institutions Utilizing Assessment Results* (Erwin 2000) presents the results of assessment case studies at eight institutions. An exploratory framework is presented in *Student Outcomes Information for Policy-Making* (Terenzini 1997; see http://nces.ed.gov/pubs97/97991.pdf). Recommendations for changes to current data collection, analysis, and reporting on student outcomes are included in the paper *Enhancing the Quality and Use of Student Outcomes Data* (Gray and Grace 1997; see http://nces.ed.gov/pubs97/97992.pdf). *Defining and Assessing Learning: Exploring Competency-Based Initiatives* explores the use of competency-based assessments across postsecondary education and

details the principles that underlie successful implementation of such initiatives (Jones and Voorhees 2002; see http://nces.ed.gov/pubs2002/2002159.pdf).

The publication has gone through NPEC's extensive review process. This product was developed through the use of Working Groups composed of professionals from all sectors of postsecondary education. In addition, four external reviewers evaluated this product. For this Sourcebook, focus groups were held at California State University (Fullerton, California), Allegany College (Cumberland, Maryland), and the University of Delaware (Newark, Delaware). Additionally, for this Sourcebook, focus groups tested the content and organization of the draft Sourcebook on faculty who teach in one of the subject areas covered by the Sourcebook. All of these review activities were designed to create the most useful and accurate products possible.

Brenda Albright
National Postsecondary Education Cooperative

Roslyn Korb
Program Director
Postsecondary Cooperative Systems
Analysis and Dissemination Program, NCES

ACKNOWLEDGMENTS

NATIONAL POSTSECONDARY EDUCATION COOPERATIVE (NPEC)
NPEC STUDENT OUTCOMES SOURCEBOOK WORKING GROUP, 2002–2003

Members:

Julie Noble, Principal Research Associate
ACT, Inc. (Co-Chair)
Iowa City, IA

Dawn Geronimo Terkla, Executive Director
Institutional Research
Tufts University (Co-Chair)
Medford, MA

Trudy Banta, Vice Chancellor
Planning and Institutional Improvement
Indiana University-Purdue University Indianapolis
Indianapolis, IN

Alice Bedard-Voorhees, Founder/Principal
Trainers-Online.Com
Littleton, CO

Randall Bergen, Dean of Instruction
Greenville College
Greenville, IL

Jennifer Brown, Director
Institutional Research and Policy Studies
University of Massachusetts-Boston
Boston, MA

Harriott Calhoun, Director
Institutional Research & Information Services
Jefferson State Community College
Birmingham, AL

Stephen Chambers, Director
Institutional Research
University of Colorado at Colorado Springs
Colorado Springs, CO

Joe Hagy, Senior Director of Special Programs
Oklahoma State Regents for Higher Education
Oklahoma City, OK

Susan Jones, Senior Analyst, Academic Affairs
University of Wyoming
Laramie, WY

Cecilia Lopez, Vice President, Academic and
Student Affairs, Harold Washington College
Chicago, IL

Sherwyn Morreale, Associate Director
National Communication Association
Washington, DC

Rick Noeth, Principal Research Associate
Policy Research Department, ACT, Inc.
Iowa City, IA

Gary Pike, Assistant Vice Chancellor
Student Affairs
University of Missouri-Columbia
Columbia , MO

Linda Suskie, Coordinator of Assessment
Office of Institutional Research, Assessment and
Planning, Towson University
Towson, MD

Nancy Szofran, Chief Technology Officer
Idaho State Board of Education
Boise, ID

Norval Wellsfry, Professor
Cosumnes River College
Sacramento, CA

Grant Wiggins, President and Director of
Programs, Relearning by Design
Pennington, NJ

NATIONAL POSTSECONDARY EDUCATION COOPERATIVE (NPEC)
NPEC STUDENT OUTCOMES SOURCEBOOK WORKING GROUP, 2002–2003

Consultants to the Working Group

Elizabeth Jones, Associate Professor
Advanced Education Studies
College of Human Resources
West Virginia University
Morgantown, WV

Stephen RiCharde
Director of Institutional Research
Virginia Military Institute
Lexington, VA

NPEC Staff

Brenda Albright, NPEC Executive Director
Education Statistics Services Institute
Washington, DC

Nancy Borkow
NPEC Project Officer
National Center for Education Statistics
Washington, DC

NPEC Staff (continued)

Roslyn Korb, Program Director,
Postsecondary Cooperative System
Analysis and Dissemination Program
National Center for Education Statistics
Washington, DC 20006

Denise Glover, Senior Research Analyst
Westat (NPEC Contractor)
Rockville, MD

Hans L'Orange, Director
SHEEO/NCES Network
State Higher Education Executive Officers
Denver, CO

Alison Leff, Research Assistant
 Washington, DC

Robert Wallhaus, NPEC Consultant
Hilton Head, SC

EXECUTIVE SUMMARY

Faculty, instructional staff, and assessment professionals are interested in student outcomes assessment processes and tools that can be used to improve learning experiences and academic programs. How can students' skills be assessed effectively? What assessments measure skills in communication? Leadership? Information literacy? Quantitative reasoning?

To better understand the success of the learning process as well to respond to requests from accreditation agencies and other organizations that are seeking greater accountability for postsecondary education institutions, some colleges and universities are using assessment instruments.

Because of the importance of assessment, the National Postsecondary Education Cooperative (NPEC), with financial support from the National Center for Education Statistics, has sponsored the development of this Sourcebook. It is intended as a resource to assist individuals who are seeking information about the assessment process and assessment instruments in the areas of communication, leadership, information literacy, and quantitative reasoning.

This Sourcebook defines the most important outcomes in each of these critical domains. Assessment tools and resources are cited, including explanations of scope, availability, measurability, cost, and other methodological concerns. Research is drawn from numerous publications that include in-depth reviews of the assessments. Faculty and staff at colleges, accrediting agencies, federal and state government agencies, and other organizations—anyone who measures, reports, or is interested in information about student outcomes can benefit from this sourcebook.

This Sourcebook includes six chapters and five searchable database tables. The introductory chapter focuses on issues in accountability, internal motivations for institutional change, the background and purpose for this project, and the intended audiences. Chapter 2 outlines the steps that need to be taken when building an effective assessment process.

Chapter 3 focuses on the expectations for students' communication, interpersonal, and listening skills. In chapter 4, leadership traits as well as situational, and functional approaches to leadership outcomes are discussed. Key issues in assessing leadership outcomes in education, and in business, are examined, including distinguishing between management and leadership. Chapter 5 discusses the constructs of information literacy as they evolved in response to the changes in technology and library resources.

In chapter 6, the authors differentiate between the key concepts of quantitative reasoning and quantitative literacy, and the assessments associated with these skills. While quantitative literacy instruments are developed to measure the level of pure mathematical ability, quantitative reasoning instruments are developed to measure problem solving and critical thinking, using quantification skills as a medium. The differences between quantitative reasoning assessment in business and academic communities are also summarized.

All the tables are structured as searchable web database tools. Table A contains reviews of instruments that assess communication skills, including those developed commercially and those created by communication scholars. The assessments encompass communication competency, teamwork, interpersonal skills and conflict management appraisal and assessment, and listening skills. Ten leadership assessment instruments are reviewed in Table B, which provides a web database tool covering leadership skills for individuals and for teams.

The searchable database for Table C provides a detailed description of assessment instruments for information literacy including the psychometric properties of each instrument. Table D compares learning modules and tutorials developed by several colleges and universities, as well as commercial publishers to the Association of College and Research Libraries (ACRL) Information Literacy Competency standards. Table E reviews 18 instruments designed to assess quantitative reasoning and quantitative skills.

TABLE OF CONTENTS

LIST OF SEARCHABLE DATABASE TABLES

1. INTRODUCTION

1.1 Assessment for Improvement Purposes

Faculty, instructional staff, and assessment professionals are often responsible for articulating student outcomes for academic programs, designing the curricula on which the programs are based, delivering the curriculum, and determining the quality of these learning experiences. As they communicate their expectations and then regularly assess student outcomes, they derive information that can be used for internal purposes, as well as shared with external audiences. Through an effective assessment process, insights can be gleaned about the types of learning occurring in programs. this information can lead to more informed decisions about needed program changes. The foundation for effective assessment is created early when decisions are made about essential outcomes, assessment methods, and audiences. Questions that should be addressed at the beginning of the assessment process include the following: Who will review the assessment results? Who will make suggestions for planned changes based on those results? Who will be responsible for implementing these informed changes? One way to enhance the internal usefulness of assessment results is to link them with other major initiatives such as strategic planning.

As Barr and Tagg (1995) note, a paradigm shift from instruction to learning is occurring in higher education. For a long time, colleges and universities were mainly concerned with instruction where by faculty and staff sought to transfer knowledge to students and offer the appropriate courses and programs to realize this goal. In the new paradigm, some institutions are redirecting their efforts to produce learning. They seek to elicit discovery by asking students to construct knowledge through gathering and analyzing information and demonstrating their skills through communication and problem solving. Students are expected to be actively involved and take responsibility for their learning. An emphasis is placed on using and communicating knowledge to address real-world problems or issues with teams of students attempting to find solutions. Through this learning-centered paradigm, performance assessment is increasingly used. Students demonstrate their skills and knowledge through activities, including essays, presentations, products, and exhibits that are rated or scored by faculty (Palumba and Banta 2001).

1.2 Calls for Accountability

Since the 1970s, the resources available to higher education have not kept pace with rising costs and inflation, resulting in a financial crisis for higher education (Huba and Freed 2000). At the same time, the population of students attending college has become increasingly diverse, including more part-time students and adult students returning to postsecondary education for additional training or retraining. The external public began to voice concerns that college graduates did not possess the skills and abilities necessary to be successful in the workplace, and some policymakers even began to question the value of higher education. The need to reform higher education began and was expressed in numerous reports calling for major changes. These reports placed a renewed emphasis on curricular issues, reinvigorated discussions of academic standards, and highlighted academic effectiveness (Eaton 1991).

Various external audiences have influenced assessment through their external reporting needs and through their influence on how faculty and staff proceed internally (Palumba and Banta 1999). These efforts often provide a stimulus for campuswide efforts to design and implement assessment plans. Regional and professional accrediting agencies require institutions or programs to assess student achievement and to document the results through appropriate measures (Palumba and Banta 2001). These organizations expect clearly specified educational objectives and assessments of student learning.

Additionally, there has been some efforts among state-level policymakers to initiate legislative reforms by implementing performance-funding programs, which earmark some portion of public resources allocated for colleges and universities based on their ability to meet performance targets including retention rates, graduation rates, or demonstrations of student learning (Huba and Freed 2001; Ewell 1997).

This report is a compendium of information about commercially developed instruments that have been cited as assessing skills in these four areas. Information about costs, content, reliability and validity, strengths and limitations of the various tests and assessments is included. In addition, this Sourcebook examines definitions and important outcomes in each of these areas, and it cites resources that provide more in-depth information about these outcomes. The primary audiences are faculty, assessment professionals, institutional researchers, and others who are involved in selecting assessments and developing assessment processes. Policymakers, including professional accrediting agencies and state-level boards, may also find this resource of value.

The tests described in this Sourcebook and the two previous volumes[1] were identified through careful research and consideration. Not all assessments in a particular area are included. The authors believe that the assessments in the sourcebooks could be relevant to the interests of postsecondary education institutions. Additionally, all comments in the sourcebooks are descriptive and should not be construed as a recommendation for any particular assessment. The instruments reviewed were selected because they were often profiled in other resources. In addition, most of them have published evidence of reliability and validity However, in some cases, instruments may be new and have less psychometric evidence, but they were included because, in the authors' opinion, they have the potential to be useful.

Additionally, for some of the instruments discussed in this compendium, the reliability may not be as high as expected or no reliability or validity is reported. Many of these instruments were designed to gauge behaviors through the use of student perceptions, which are not always reliable. These instruments will have relatively low reliability estimates. Furthermore, in the case of instruments that examine communication or leadership styles, there is no commonly accepted norm or measure? against which to compare a given set of ratings or scores.

1.3   Organization of this Report

This report is divided into six chapters and five searchable database tables. The introductory chapter focuses on accountability and internal motivations for change. Chapter 2 presents an overview of planning an effective assessment process. A brief summary of the critical steps includes developing statements of intended learning outcomes, selecting assessment measures, and reviewing and discussing assessment results to identify key improvements. Chapter 3 focuses on essential communication, interpersonal, and listening skills. The key definitions of these outcomes and issues in assessing these skills are reviewed. In chapter 4, the important leadership outcomes are discussed. Key issues in assessing leadership outcomes are outlined and examples in the business environment are presented. Chapter 5 focuses on information literacy and highlights key definitions promoted by national organizations. In

---

[1] NPEC has other related products that focus on student outcomes assessments. *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing* is a compendium of information about tests used to assess these three skills. *The NPEC Sourcebook on Assessment, Volume 2* provides eight case studies of institutions that have addressed policy-related issues through the use of the assessment methods presented in Volume 1.

chapter 6, key definitions of quantitative reasoning, literacy, and numeracy are discussed. Quantitative reasoning in the business community is also summarized.

All five tables are structured as searchable web database tools. Table A contains reviews for commercially developed instruments that measure outcomes in the areas of communication, interpersonal, and listening skills. In addition, instruments developed by communication scholars are examined. For each review, information about the publisher (including their telephone number and/or web site address when available), costs, and testing time are provided. The total score and subscores are defined by the publisher's categories. Information about reliability and validity are presented when available. If there are reports of association with other measures, that evidence is presented. Finally, the strengths and limitations are briefly summarized. Ten leadership assessment instruments are reviewed in table B, which provides a web database tool covering leadership skills for individuals and for teams.

The searchable database for table C provides a detailed description of assessment instruments for information literacy including the psychometric properties of each instrument. Table D compares learning modules and tutorials developed by several colleges and universities, as well as commercial publishers to the Association of College and Research Libraries (ACRL) Information Literacy Competency standards. Table E reviews 18 instruments designed to assess quantitative reasoning and quantitative skills.

2. PLANNING AN EFFECTIVE ASSESSMENT PROCESS

2.1   Introduction

As faculty create or revise their assessment plans, it is important to be clear about the purpose of assessment. Formative assessments are often "conducted during the life of a program (or performance) with the purpose of providing feedback that can be used to modify, shape, and improve the program (or student performance)" (Palumbo and Banta 1999, p. 7). Summative assessments are completed at certain points in time after a program has been implemented or at its conclusion in order to judge the quality of the program or student performance compared to defined standards. The results from these assessments may be used to make decisions about whether to continue the program or to repeat certain sets of activities. Overall, the findings may be used to make decisions about the future of the program.

The primary purpose of most assessment plans is the improvement of educational programs and student learning. However, there are times when summative assessments may be required by regional or professional accreditation agencies or state governments. Palumbo and Banta (2001) examine how accreditors have encouraged attention to the assessment of student learning and how faculty in professional preparation programs (including nursing, teacher education, business, engineering) respond to the actions of accreditors.

To build an effective assessment process, a series of important steps should be completed. These steps are briefly discussed below, and sources of in-depth guidance about building effective assessment plans are referenced.

2.2   Developing Statements of Intended Learning Outcomes

The first step in the assessment process is to define expectations for student learning (i.e., what should students know and be able to do with their knowledge). Although this may seem like an obvious initial step, statements of outcomes do not exist for many curricula and courses (Diamond 1998). Most college catalogues present institutional goals, purposes, or missions in the form of broad concepts, such as character development, appreciation of diverse cultures, or citizenship (Erwin 1991). These types of goals are ambiguous and broad. Ewell (1997, p. 3) notes that curriculum reform has been "implemented without a deep understanding of what collegiate learning really means and the circumstances and strategies that are likely to promote it." Therefore, as a beginning step, one must define specific learning outcomes or objectives. These statements represent what the faculty hope students will achieve, while the results from assessments document actual student outcomes.

It is helpful if the appropriate stakeholders fully participate in identifying, defining, and reaching a consensus about important outcomes. When there are clear statements of learning outcomes for student performance that are measurable, a more coherent curriculum can emerge. These outcomes provide direction for all instructional activity, inform students about the intentions of the faculty, and form the basis of assessment at the course, program, and institutional levels (Huba and Freed 2000).

Intended learning outcomes can address important aspects of learning that are considered important for students to be effective citizens and employees in the workplace. There are several taxonomies of educational objectives that faculty can review as they make decisions about the most important outcomes for their own students. Bloom's (1956) taxonomy of cognitive outcomes consists of six different levels ranging from recalling basic knowledge to evaluating information.

These types of outcomes are considered important for college students. Bloom's work continues to be cited by numerous assessment scholars including Palomba and Banta (1999) and Waterhouse (2005).

Listed below are Bloom's six original outcomes:

- **Knowledge.** Recognizing or recalling facts, terminology, principles, or theories. Includes behaviors such as describing, listing, identifying, or labeling.

- **Comprehension.** Understanding, ability to describe in one's own words, to paraphrase. Includes behaviors such as explaining, discussing, or interpreting.

- **Application.** Using material in a new way, applying concepts, laws, or theories in practical situations to solve problems. Includes behaviors such as demonstrating, showing, and making use of information.

- **Analysis.** Breaking down information into its component parts to see interrelationships and ideas. Includes behaviors such as differentiating, comparing, and categorizing.

- **Synthesis.** Combining the parts into a new whole, arranging or rearranging to get new patterns and structures. Includes behaviors such as using creativity to compose or design something new.

- **Evaluation.** Comparing material or ideas to known standards, judging or making decisions based on appropriate internal or external criteria. Includes behaviors such as concluding, criticizing, prioritizing, and recommending.

Knowledge and comprehension are often labeled "lower order" cognitive skills. Application, synthesis, analysis, and evaluation require students to use more advanced thinking skills and are often called "higher order" cognitive skills. Faculty articulate a combination of both lower order and more advanced reasoning skills as they define their expectations for student learning.

Anderson and Krathwohl (2001, p. 31) revised Bloom's original taxonomy. They wanted to refocus instructors' attention on the value of his work and add new knowledge and information into the framework. The taxonomy still consists of six major levels but the order of the last two levels has changed and terms have been revised as outlined below.

- **Remember.** Retrieve relevant knowledge from long-term memory. Involves recognizing and recalling information.

- **Understand.** Construct meaning from instructional messages including oral, written, and graphic communication. Involves interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining.

- **Apply.** Carry out or use a procedure in a given situation. Involves executing and implementing.

- **Analyze.** Break material into constituent parts and determine how the parts relate to one another and to an overall structure or purpose. Involves differentiating, organizing, and attributing.

- **Evaluate.** Make judgments based on criteria and standards. Involves checking and critiquing.

- **Create.** Put elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure. Involves generating, planning, and producing.

Affective outcomes are also important and include both values and attitudes. Individuals often possess deeply held beliefs, ideas, and assumptions about life goals and ways of living (Erwin 1991). These values may influence how an individual may behave. Attitudes consist of feelings toward people, ideas, and institutions (Palumbo and Banta 1999). Krathwohl, Bloom, and Masia (1964) have developed a taxonomy that defines affective objectives as follows:

- **Receiving.** Being aware or willing to attend to something; learner is passive but attentive, listening with respect.

- **Responding.** Complying to given expectations; learner participates actively by reacting as well as showing awareness.

- **Valuing.** Accepting importance of the attitude; learner displays behavior consistent with a belief or attitude though not forced to comply.

- **Organization.** Committing to a set of values; bringing together different values and resolving conflicts between them; building an internally consistent value system.

- **Characterization.** Behaving according to a characteristic life style or value system; maintaining a consistent philosophy regardless of surrounding conditions.

Finally, professional associations in various disciplines and program accrediting organizations often state specific learner outcomes—such as the achievement or mastery of an ability or skill, or the development of a value or attitude—that are important for graduates in their areas to achieve.

2.3   Selecting Assessment Measures

Through their specific objectives, faculty and instructional staff identify the most important priorities for student learning and development. These objectives serve as the basis for determining how to best collect, assess, and interpret the data to make improvements. The second essential step is to create or use existing instruments to determine if students are mastering these defined expectations. There is a wide array of methods and instruments that could be selected. However, assessment may be most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time (American Association of Higher Education 1992). This typically means that faculty and staff choose multiple methods to assess student learning. They closely review and make decisions about the strongest assessment instruments that will measure specific outcomes (U.S. Department of Education 2000).

To plan an effective assessment process, faculty gather evidence that is closely related to the defined learning outcomes (Palumbo and Banta 2001). There are numerous commercially developed instruments that assess outcomes deemed important by faculty. These instruments usually contain information about their reliability and validity, but often assess only some of the intended outcomes. It is difficult to find an instrument that will measure all stated outcomes. Therefore, faculty sometimes use commercially developed instruments supplemented with locally developed course-embedded assessments such as projects, papers, products, exhibitions, performances, or portfolios. These direct methods may be more relevant to the specific learning outcomes that faculty and staff want to examine and may be more

appropriate and effective because students are asked to demonstrate what they know and can do with the knowledge (Huba and Freed 2000).

As faculty develop course-embedded assessments, they also may design rubrics or rating scales to determine the quality of a student's performance. A rubric is usually based on a set of criteria used by an individual or multiple raters to judge student work. The criteria are ideally explicit, objective, and consistent with expectations for student performance. These rubrics articulate what knowledge, content, skills, and behaviors are characteristics of various levels of learning or mastery. Rubrics are meaningful and useful when they are shared with students before they are assessed so they better understand the expectations for their performance. Huba and Freed (2000) and Walvoord and Anderson (1998) provide helpful advice about how to develop and utilize rubrics for course-embedded assessments.

Individuals who decide to develop their own assessment methods, particularly tests, can consult with the *Standards for Educational and Psychological Testing* published by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (1999). These Standards provide critical information about test construction, evaluation, documentation, fairness in testing, and testing applications. In addition, the *Use of Tests as Part of High-Stakes Decision-Making for Students: A Resource Guide for Educators and Policy-Makers* (U.S. Department of Education 2000) provides guidance about the development and implementation of policies that involve the use of tests as part of making decisions with high-stakes consequences for students.

Faculty and staff may also consider self-report methods such as surveys that can be distributed to students and used both in individual courses and at the program or institutional levels to gain information about students' perceptions. Surveys of alumni and employers are examples of indirect measures that can provide useful data about perceptions regarding academic programs at the college or university. Suskie (1996) provides a thorough overview of planning, developing, and implementing locally developed surveys as well as processing the results, analyzing the data, and reporting the findings.

The overall purpose of an assessment program also influences the choice of instruments. For example, a statewide coordinating board may want to compare students' performance of quantitative reasoning skills with that of other college and university populations within the state. Sometimes, such organizations may require that all college students complete a particular commercial test so that comparisons can be made. However, faculty are often interested in assessing student learning so that improvements in the curriculum can be made.

Faculty who are interested in a group of commercial instruments can request examination copies from the publisher. Then groups of faculty can review each instrument and evaluate how closely it assesses each important student learning outcome. Through a formal review of each potential assessment instrument, faculty can discuss which methods seem the strongest and match their needs.

As faculty and staff review different instruments, there are numerous aspects that they will want to carefully evaluate. Many of these different aspects are discussed fully in NPEC's Sourcebook, Volume 1 (Erwin, 2000). These dimensions include the following:

- **Conceptual Considerations**

    - Relevancy of particular outcome to the issue or problem,

    - Utility of potential data,

    - Applicability of assessment measures,

- Interpretability of the test information,

- Credibility of the measure and resulting data, and

- Cultural fairness.

- **Methodological Considerations**

  - Scope of data needed,

  - Availability of outcome measures, and

  - How outcome is operationally defined and measured.

- **Test Properties**

  - Reliability,

  - Method design, and

  - Validity.

The authors adopted the terminology from NPEC's Sourcebook, Volume I so that there would be consistency in the usage of key terms. However, the term *validity* was modified by the American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (1999). In 1974, *content validity* referred to a kind or aspect of validity that was "required when the test user wishes to estimate how an individual performs in the universe of situations in which the test is intended to represent" (p. 24). In the current Standards, validity is viewed as a unitary concept in which "content validity" is now characterized as "evidence based on test content" (p. 174).

Another dimension is the cost of the actual instrument, the facilitator's guide, and the technical manual, which often includes information about reliability and validity. Some commercial testing companies also provide scoring services and will analyze and report the findings from an individual college or university assessment program. Other publishers require institutions to analyze and report the findings themselves.

Costs can relate to the amount of time that it takes the institution to either analyze the results or get the results back from the testing company. The amount of students' time required for each assessment can be another cost issue. Faculty or staff may be reluctant to set aside 2 hours of class time for an assessment to be administered. However, if the assessment takes only 15 or 20 minutes, they may be more open to freeing up some of their class time so that students can participate. Embedding assessments within existing courses or other activities can lead to greater participation from the students. For example, orientation programs for new students could include administering relevant assessments to determine their abilities, skills, and attitudes.

2.4   Reviewing and Discussing Assessment Results to Identify Key Improvements

Once faculty review the assessment results closely, it is critical that they identify potential key improvements. Faculty can discuss the results and then use them to make informed changes. Through these open discussions, faculty can gain insights into the type of learning occurring in the program and

better understand what students can do well and where they have not succeeded (Huba and Freed 2000). Such information can provide insights about where targeted improvements are needed. Assessment results can also be used for program review or strategic planning. In addition, these results can be shared with external audiences such as accreditors and state policymakers.

2.5   Summary

Some colleges and universities are seeking ways to build strong assessment programs that provide meaningful information about whether students are mastering important skills and outcomes. As plans are developed or revised, it is critical to define important outcomes that can be assessed. It is equally important to use multiple assessment methods to determine if students have achieved the desired outcomes. Finally, the review and discussion of assessment results by all relevant stakeholder groups is vital to ensure that a range of ideas about ways to strengthen student learning are generated based on the findings.

3. IDENTIFYING IMPORTANT ORAL COMMUNICATION, INTERPERSONAL, AND LISTENING SKILLS

3.1   Introduction

Many faculty, employers, and policymakers agree that college students should be skilled communicators and problem solvers (Jones 1997). The quantity and complexity of information has been increasing at a rapid rate and can be overwhelming for new college graduates who must learn to gather, organize, and manage it. An undergraduate education should provide students with the necessary skills, abilities, and values that are critical to successfully navigate the dynamic complexities in our diverse environments.

As Gabelnick notes, "the challenge of educating a committed citizenry is to change the societal and university paradigm from a strategy of competitiveness to one of collaboration, from a perspective of scarcity to one of sufficiency and inclusion, and from a stance that looks for expedient solutions to one that engages and commits to a series of values and way of life" (1997, p. 10). College students interact with a wide array of individuals and groups during their lifetime. One of the challenges for higher education is to prepare students with a set of strong communication and interpersonal skills. The ability to work in teams is crucial for college students as they participate in more collaborative environments. In addition, students are confronted with a multitude of messages through various forms of media. Through their learning experiences, they can become more adept at analyzing messages and drawing their own conclusions about very complex issues.

Increasingly employers are searching for employees who have strong abilities in such areas as problem solving, team work, communications, leadership, learning, and systems thinking (Carnevale 2000; Rao and Sylvester 2000; Oblinger and Verville 1998; Miles 1994). Although most employees enter new positions with adequate technical skills, it is the general skills (especially communications and problem solving) that count toward successful job performance over time, and it is these skills that are most often absent (College Placement Council 1994). In addition, college students view these skills as crucial and necessary to ensure their own career mobility (Education Commission of the States 1995). More recent studies continue to report that employers believe their new college graduates demonstrate weaknesses in "the ability to communicate orally and in writing, interpersonal and leadership skills, the capacity to contribute to and participate in teams, analytical ability and adaptability" (Business-Higher Education Forum 1997, p. 20). While some reports are very critical of the abilities of their new employees, others suggest that the gap exists increasingly today between ideal outcomes and actual performance because there have been considerable efforts to increase the skills of new workers through training across various sectors of the work place (Business-Higher Education Forum 1997; Carnevale, Gainer, and Meltzer 1990).

3.2   Communication Skills Concepts

Before any assessments of communication skills are planned and implemented, faculty and administrators can reflect upon several concepts. First, communication is inherently interactive and dialogic (Daly 1994). For example, two students can create meaning together by engaging in an interaction. Conceptually, the focus is on these interactions including interviews, debates, meetings, small group work, and presentations. A student speaker and his or her listener craft meaning from their interactions, which can create and maintain a social reality as long as they interact (Daly 1994).

Communication also occurs in real time. Preparation for most types of interaction (except for most formal public speaking) occurs simultaneously with production (Daly 1994). In discussions or meetings, students

seldom spend considerable time preparing before the interaction in terms of what they will say. This is a major distinction between speaking and writing. Students can have multiple opportunities to revise their writing, and the same is often expected from faculty. Many forms of written communication are usually time-delayed (Daly 1994). However, discussions require immediate interaction among the participants.

Communication is also embedded in different contexts. Students speak and listen within different situations that shape the sorts of interactions that occur and influence the interpretations made by participants, and affect the effectiveness of the communication. The structure, content, and style of a particular formal speech may be appropriate in one context and then inappropriate in another.

Finally, communication has crucial outcomes that can be attained by effective interactions. For example, participants in a discussion can leave that meeting with a different understanding of key issues and insights into the perspectives of others. There are consequences from all types of interactions. In addition, most individuals have goals that they hope to achieve through their communications.

3.3 Defining Important Speech Communication and Listening Outcomes

College students should achieve certain communication competencies to be fully prepared to effectively participate in the workplace and society. There are two major sets of competencies that are discussed in this section. The first set, developed by the National Communication Association (1998), defines speaking and listening skills. The major competencies are listed below.

Speaking Competencies

College students should be able to

- determine the purpose of oral discourse;

- choose a topic and restrict it according to purpose;

- fulfill the purpose of oral discourse by;

  - formulating a thesis statement,

  - providing adequate supporting material,

  - selecting a suitable organizational pattern,

  - demonstrating careful choice of words,

  - providing effective transition;

- employ vocal variety in rate, pitch, and intensity;

- articulate clearly;

- employ language appropriate to the designated audience; and

- demonstrate nonverbal behavior that supports the verbal behavior.

Listening Competencies

College students should be able to

- recognize main ideas,

- identify supporting materials,

- recognize explicit relationships among ideas,

- recall basic ideas and details,

- listen with an open mind,

- perceive the speaker's purpose and organization of ideas,

- discriminate between statements of fact and statements of opinion,

- distinguish between emotional and logical arguments,

- detect bias and prejudice,

- recognize speaker's attitude,

- synthesize and evaluate by drawing logical inferences and conclusions,

- recall the implications and arguments,

- recognize discrepancies between the speaker's verbal and nonverbal messages, and

- employ active listening techniques when appropriate.

A national study spearheaded by Jones (1997) identified the essential communication skills necessary for college graduates to be effective. The framework adapted for this study was drawn from work originally conducted by Bassett and colleagues (1978), as well as the foundation for an assessment instrument developed by Rubin (1982). A speech communications goals inventory was developed.

The goals inventory contains four major categories of essential skills. The initial grouping consists of basic communication skills relating to selecting and arranging elements to produce spoken messages. The second set is advanced communication skills that require students to use their analytic and reasoning skills including audience analysis. Examples of advanced skills include being able to understand people from other cultures, organizations, or groups, and adapting messages to the demands of the situation or context (Jones 1997). These skills are more than just knowing, doing, or feeling (Rubin and Morreale 1996). They are blends of knowledge, skills, and attitude; they require greater levels of behavioral flexibility as well as adaptability (Morreale, Rubin, and Jones 1998). Advanced skills in both public speaking and interpersonal communication are included.

The third category of skills focuses on interpersonal and group communication that relate to the development and management of human relations. The fourth set consists of communication codes relating to the ability to use and understand spoken English and nonverbal signs. The final grouping of

skills emphasizes the evaluation of oral messages and their effects. College students and graduates may use these skills in various communication contexts.

More than 600 faculty and policymakers rated the importance of these skills for their college graduates, while employers rated their importance for their new employees who had completed college degrees. These stakeholder groups reached a consensus about the importance of 87 percent of the speech communications skills. The speech communication framework and the entire set of specific skills are available in Jones, 1997.

3.4   Key Issues in Assessing Communication Skills

An assessment of communication skills should include a behavior sample (Daly 1994). Simply knowing how to structure a message or other elements of effective presentations does not mean that students will necessarily apply their knowledge to their actual performance. Therefore, a substantial behavior component is critical.

It is also important to consider other outcomes as well as the performance. An individual's presentation may be effective in terms of delivering an interesting message, using a varied tone, and providing strong evidence for his or her claims. However, when participants are asked if they were persuaded to change or adopt the speaker's position and none have changed their minds, the speaker did not achieve his or her intended goal. Often in assessments, there is a failure to consider whether or not an individual has actually achieved his or her interactive goal (Daly 1994). An assessment of communication must include the listener to determine if the message has had an impact. After the communication, did the participant have a stronger understanding of the message or key issues? Was the participant persuaded by the message? Most assessment instruments will include feedback from both the participants (audience) as well as the actual speaker who conducts a self-assessment.

Assessing oral communication skills is more challenging than assessing writing or reading skills. Part of this challenge stems from the need to assess nonverbal behaviors. Students in social interactions, meetings, presentations, and other types of exchanges use both their verbal messages and their actions to communicate (Knapp and Hall 1992). Listeners or participants usually interpret what is stated in the context of nonverbal behaviors. Faculty should carefully consider and identify the nonverbal behaviors that contribute to an effective message.

Assessments should also examine both typical performance competency and the maximum performance capability (Daly 1994). The way a student may typically participate in discussions and what he or she is actually capable of doing could be very different when one examines oral communication and listening. Regular or routine conversations elicit different behaviors than working in a team on a formal project where the individual is seeking a specific goal or the interaction has some degree of uncertainty.

Important decisions need to be made about whether the evidence will be based upon an individual student and reported back to that student or be aggregated and reported for groups of students (Daly 1994). This consideration will shape what information is collected, how it is collected, how interpretations are made, and how the results are shared. Individual-level scores can offer valuable information upon which to base decisions about placement into the appropriate level of courses and allow diagnosis that would indicate strengths and areas for improvement.

Finally, decisions need to be made about the appropriateness of holistic versus atomistic judgments of communication performance (Daly 1994). An instructor may want to assess very specific behaviors or actions, such as how many times a particular student participated in the class discussion. On the other

hand, the instructor may want to focus on the quality of the individual's contribution and assess whether the student offered strong evidence or reasoning to support his or her position.

3.5   Assessment Guidelines for Oral Communications

The National Communication Association (1998) recommends assessment criteria that cut across several areas. First, they suggest general criteria that focus on the purpose of assessment and the multiple, interactive dimensions of communication. Second, they believe the content of assessment should include communications in more than one setting. Third, they outline specific criteria for selecting assessment instruments that are consistent with the criteria outlined in this Sourcebook. Finally, they conclude with criteria regarding assessment procedures and administration, assessment frequency, and the use of assessment results.

The majority of the instruments reviewed for communication are inventories asking students to rate their own skills. Other instruments ask students to rate their skills and then ask the same questions to observers (an instructor or peers). Comparisons can be made to see if there is congruence between an individual's self-rating and the judgments of other raters.

Three clusters of instruments are reviewed in table A. The first cluster focuses on methods to assess an individual student's competence in communication. Very few commercially developed instruments exist that are designed to actually measure students' competence in communication. The second group consists of measures to assess teams or groups. The third set focuses on measures to assess interpersonal skills such as conflict management. A sampling of instruments, rather than an extensive review of all available instruments, is reviewed in this Sourcebook. Rubin, Palmgreen, and Sypher (1994) review additional instruments developed by communication scholars and include the contents of the actual instruments in their profile. Morreale and Backlund (1996) review instruments for assessing oral communication in kindergarten through 12th grade as well as in higher education.

The instruments in table A were selected because they were profiled in several resources such as the two documents listed above. In addition, most have some evidence of reliability and validity. In some cases, instruments are new and have less psychometric evidence, but they were included because the authors believe they have the potential to be useful.

3.6   Assessment Within the Workplace

Increasingly, corporations are finding that they must offer formal training programs including courses, workshops, and seminars to help their employees develop stronger leadership, team building, project management, writing, interpersonal communications, problem solving, negotiation, and conflict management skills. Within the business environment, the evaluation of these training programs is a focus of concern among managers and executive leaders who want productive organizations.

The most common and frequently used framework to evaluate employee performance and the impact of training is a model created by Kirkpatrick (1994). There are four different levels that should be evaluated. At Level 1, employees' reactions to formal training typically focus on how much they liked the course. The assessment tools most frequently used are locally developed rating sheets. The second level is the actual amount of learning gained by employees. Often tests and simulations are tools used to evaluate the extent of learning. At Level 3, the behaviors are examined and tools are used to measure how well employees apply what they learned to their jobs or positions. Locally developed performance measures

tend to be used to address this issue. Finally, Level 4 emphasizes results and the determination of the return on the investment of training. Usually cost-benefit analyses are conducted to address this last level.

Data are easiest to gather and interpret at Levels 1 and 2. The effects from Level 1 tend to be short-term; as employees progress through the levels, the results are more lasting. An estimated 85 percent or more of all training programs evaluate Level 1, which decreases by ascending levels to fewer than 10 percent being measured at Level 4 (Parry 1997).

Carnevale, Gainer, and Meltzer (1990) examined successful workplace training programs in oral communication. They found that companies such as Xerox, Digital Corporation, and IBM have large, advanced training delivery systems. Training is "provided through in-house trainers, external consultants (sole practitioners and companies), associations, local colleges and universities, and vendors of interactive videos, computer programs, and workbooks" (Carnevale, Gainer, and Meltzer, p. 139). Companies often view communication as a critical competency that is the heart of businesses that rely on their employees' interpersonal skills. The specific nature of training programs within the corporate setting may be outlined in a general nature, but the specific details are usually not provided. "More training in communication skills is being provided as the sources of organizational return on investment are better documented" (Carnevale, Gainer, and Meltzer, p. 142). Companies now embrace the belief that the more communication training is offered, the greater the potential that employees will develop over time rather than be fired or released from their positions.

Portnoy (1986) developed the relationship life cycle model that is frequently used to illustrate how people develop working relationships and either learn to cope with differences or end relationships because of conflicts that could not be resolved. Using this model, it is possible to identify essential skills that lead to greater competency in communication and interpersonal skills. According to the Portnoy model, individuals initially establish credibility that can be formal or informal within their workplace. Carnevale, Gainer, and Meltzer (1990, p. 299) note that areas of training related to this stage of Portnoy's model include "cross-cultural awareness (differences among people from dissimilar backgrounds), job skills knowledge (competent job performance), and written communication (accurate expression on paper)." The second stage in the Portnoy model is getting acquainted as each individual or group evaluates the other. Carnevale and his colleagues (1990, p. 299) note that essential areas of training related to this stage in Portnoy's model include "interview skills (asking and responding to questions), active listening (responding nonjudgmentally to a speaker's content and feeling, thereby building rapport), values clarification (discovering what is important to a person), interest identification (learning what each person likes), learning styles, and nonverbal communication."

In Portnoy's third stage, attachments are formed as the relationship unfolds. Carnevale and his colleagues believe critical areas of training are as follows: "disclosure (opening up to another), process observation (understanding and describing the action in the immediate environment), feedback (giving and receiving responses), oral communication skills, and self-insights (understanding the thoughts, feelings, and motives of oneself)" (p. 300). The fourth stage in the Portnoy model focuses on an individual's participation in determining the functions that are to be performed by each person in the relationship. Carnevale and his associates note the areas of training that can address this stage in the Portnoy model include "negotiation skills, role negotiating (positioning with regard to others), modeling (setting an example and learning from others' examples), mental flexibility (adapting to the needs of the moment, and goal setting" (p. 300). In the next stage of the Portnoy model, members of a functioning unit become closely connected to their groups and their organization. According to Carnevale and his colleagues, the main areas of training stress "group processing skills, group dynamics, and coaching" (p. 301). As the groups become stable over time, they can proceed in a smooth way. Under these conditions additional training would include "learning to work together in teams, group growth skills (enhancing group performance), risk taking (performing creatively), and consensus building (gaining support for ideas and

actions)" (p. 301). Disruptions in relationships occur because of tensions that may arise and when one's behavior is not consistent with another's expectations (Portnoy 1986). Areas of training needed in this case emphasize "patience and flexibility (learning how to deal with discontinuity of expectations), brainstorming, and tolerance for ambiguity" according to Carnevale and his colleagues (p. 301). During the stage of instability in Portnoy's model, a relationship becomes very strained because an individual's or group's needs are not being fulfilled. Carnevale and his colleagues (p. 302) believe areas of training related to this stage include "conflict management, repair strategies, and force field analysis (a problem-solving technique for analyzing and dealing with resistance to change)."

In this Sourcebook, we review some assessment tools that have been used in the workplace that may have potential applications in postsecondary education (see table A). Often these tools are self-inventories, asking participants to reflect on their own abilities and skills to judge their own capabilities. Sometimes these instruments include an "other" rater who is familiar with the individual and can assess the skills. These tools may be useful in the context of higher education.

4. IDENTIFYING ESSENTIAL LEADERSHIP SKILLS

4.1 Introduction

Many colleges and universities believe that the development of leadership is important and articulate their vision within mission statements. However, for the most part, development of leadership skills in college graduates has been viewed as a by-product of the education process and not the results of well-designed deliberate learning experiences (Schwartz and Lucas 1998). Some faculty may believe that leadership can not be taught or that leaders are born. Despite these myths, in the past several years the leadership mission of colleges and universities has been resurrected in programs designed for deliberate cultivation of leadership abilities and habits (Schwartz and Lucas 1998). One aspect of this is the move toward more collaborative models of shared leadership or self-empowered teams. While leadership development programs continue to experience some growth, there is a definite need to assess students' leadership capabilities both as individuals and within teams.

4.2 Defining Important Leadership Outcomes

There are four primary approaches for understanding and explaining leadership. Researchers who conducted the early studies believed that leaders were born with specific traits that predisposed them to positions of influence (Hackman and Johnson 2000). This **traits approach** was the focus of many research studies between the early 1900s and the end of World War II. Stogdill's (1948) review of 124 published studies that examined traits and personal factors related to leadership did not find a strong connection between physical traits and leadership (Hackman and Johnson 2000).

Research studies have found certain competencies or skills to be related to enhanced leadership effectiveness in different contexts. For example, interpersonal competencies from "skill-based behaviors, such as the ability to present an effective oral presentation or manage conflict, to more individual-based approaches such as emotional stability and self-confidence" are related to leadership effectiveness (Hackman and Johnson 2000, p. 65). The authors also found that effective leaders have stronger abilities in making decisions, solving problems, and thinking critically. They are particularly effective at solving problems that are complex, embedded within ambiguous circumstances, and require more creative solutions.

According to scholars who advocate the traits approach, successful leaders are "better at planning and organizing and are generally well versed in the methods, processes, and procedures, and techniques required for the completion of tasks performed by their followers" (Hackman and Johnson 2000, p. 65).

The **situational approach** was the next focus of research studies on leadership. Researchers believed leadership was contingent upon variations in the situation. Differences in leadership style "might be attributed to task and relational structure, superior-subordinate interactions, the motivation of followers, or any one of a number of other situational factors" (Hackman and Johnson 2000, p. 66). Although the situational theory was a different view on leadership, it was not more comprehensive than the traits approach in measuring leadership.

While the traits and situational approaches focused on the individual characteristics of leaders and followers, the **functional approach** emerged with an emphasis on the leaders' behavior. The underlying assumption was that leaders perform certain functions that allow a group or organization to operate effectively (Hackman and Johnson 2000). This perspective was used primarily to study group leadership.

James MacGregor Burns (1978) initially examined the **transformational approach** by comparing it with the more traditional "transactional" form of leadership. The motivational appeals of the transactional leader are designed to satisfy basic human needs, while the transformational leader strives to satisfy a follower's higher level needs (Hackman and Johnson 2000, p. 88). The transformational leader moves beyond concern with basic needs and works toward getting his or her followers to develop strong self-esteem and self-actualization. These leaders would seek to empower and inspire their followers in an effort to promote higher levels of motivation and ethical behavior (Hackman and Johnson 2000).

Researchers have investigated the characteristics of transformational leaders (e.g., Peters and Waterman 1982; Peters and Austin 1985; Peters 1992; Bennis and Nanus 1985; Kouzes and Posner 1995). They have found that leaders are change agents who try new ideas and challenge the status quo by experimenting with new strategies to perform tasks. They consistently seek opportunities to grow, innovate, and improve. These leaders take risks and learn from mistakes (Kouzes and Posner 1995).

Transformational leaders communicate their ideas through images, metaphors, and models that organize meaning for their followers (Hackman and Johnson 2000). Successful leaders are aware of the needs and motivations of their followers. They encourage open communication and facilitate the exchange of ideas and insights from their followers. Effective leaders empower others. They build individual or team capabilities by helping them develop competence, sharing power, providing choices, and offering visible support (Kouzes and Posner 1995).

Effective leaders also communicate a vision that provides followers with a sense of purpose and encourage commitment. A vision is a concise statement or description of the direction in which an individual, group, or organization is expected to strive toward over time (Hackman and Johnson 2000). Nanus (1992) finds that an effective vision attracts commitment, energizes and creates meaning for people, establishes a standard of excellence, and builds a connection between the present and the future.

Transformational leaders demonstrate a high level of passion and personal enthusiasm for their work that motivates others to perform to their highest levels (Hackman and Johnson 2000). They regularly recognize individual contributions to the success of projects and celebrate team accomplishments (Kouzes and Posner 1995).

4.3   Key Issues in Assessing Leadership Outcomes

In examining potential assessment instruments, it is important to determine whether an assessment instrument is measuring leadership or management skills. Sometimes these terms are used interchangeably in the literature and in documentation for various survey instruments. The instruments reviewed in this section focus primarily on leadership rather than management. Leaders are more concerned with the direction of the group while managers are more focused on the status quo (Hackman and Johnson 2000). John Kotter (1990) outlines the differences that he perceives between management and leadership by examining three main activities: creating an agenda, developing a network for achieving the agenda, and implementing the agenda. Managers who work on creating an agenda focus on planning and budgeting including details about time frames, analyses of potential risks, and resource allocations. However, leaders design an agenda by creating a direction and communicating long-term goals. "The presence of a shared and meaningful vision is a central component of effective leadership" (Hackman and Johnson 2000, p. 13).

Once the agenda is created, managers work on getting individuals with the appropriate training to carry out the plan. They focus on organizing and staffing. Leaders seek to align people by concentrating on integration, forming teams, and gaining commitment.

As managers implement the agenda, they focus on containment, control, and predictability (Hackman and Johnson 2000). Leaders implement their agenda by motivating and inspiring others. They seek to stimulate empowerment and creativity. Managers seek to produce orderly results while leaders seek to bring about appropriate changes.

Two major groups of instruments are examined in this Sourcebook (see table B). One set focuses on leadership as an individual process. In this case, a leader rates his or her style and then an observer also rates the same dimensions. In many cases, certain competencies are embedded within the different leadership styles assessed through the instruments. Styles reflect different leadership orientations rather than one set of ideal effective leadership skills. Very few commercially developed instruments assess different levels of leadership performance where particular advanced skills build on more foundational skills. As the results from these leadership inventories are analyzed, one can examine the degree of consistency in the ratings between the observer and the leader. The Student Leadership Practices Inventory is a commercially developed instrument specifically designed for college students. Other instruments focus on leadership as an individual process and were developed for training and assessment purposes in the workplace. These instruments are reviewed as well because the majority of them consist of items that could be used to assess college student leadership. The second set of instruments is designed to assess team leadership. No commercially developed instruments exist that are designed specifically to assess college students. However, other instruments are reviewed because they contain items that would allow undergraduates to make judgments about team leadership. The individual survey items in most cases are applicable and relevant for the college student population.

A variety of undergraduate leadership courses, concentrations, minors, and majors are profiled in *Leadership Education: A Sourcebook of Courses and Programs* (Center for Creative Leadership 1998). This resource provides information in the format of abbreviated syllabi that outline the course content, structure, and general description of the assessment methods. Most often, the course-embedded assessment methods are locally developed and may include tests, quizzes, article critiques, case studies, simulations, analytic or reflection papers, and action research projects.

5. DEFINING ESSENTIAL SKILLS FOR INFORMATION LITERACY

5.1   Introduction

The term *information literacy* (IL) was first coined in 1974 and attributed to Zurkowski (Doyle 1992). Since that time, the concept has gone through several iterations and has emerged in its current form over the last decade and a half. The construct known as information literacy evolved in increments that can be tied both to changes in emphasis in cognitive psychology and to advances in technology. As computer technology advanced, the information available to students exploded. It became necessary to wed existing psychological constructs across all domains to new skills expected of students at every level.

As a psychological construct, information literacy has been most closely associated with critical thinking (see Bloom's Taxonomy of Educational Objectives described in Section 2.2 of this Sourcebook), and constructivist-based education (Loertscher and Wolls 1997). The cognitive skills include identifying necessary information, extracting the required information, evaluating information critically, and using information from a wide range of resources. Out of the technological advances, the additional expectation has arisen that students on college campuses be able to identify and extract information from vast resources and databases worldwide at the touch of a few computer keys.

In 1968 there was a widespread assumption that a digital utopia was just around the corner (Dupuy 1968, p. 7). In 2000, the wave of anticipated technological advancement created a backlash of concerns based on the simple fact that computers are digital but people are analog (Crawford 1999, pp. 1-4). There is currently a new set of problems for the multimedia specialist working to support student learning that has been created by the very innovations created to solve all our information problems. These latest problems revolve around the question of how one teaches an analog student to impose order in a world of seeming digital chaos, and they result in the need to consider information management and information literacy as two sides of the same coin.

It is out of this need to impose order on the digital cacophony that the concept of information literacy has emerged. In 1989, Breivik and Gee's *Information Literacy: Revolution in the Library*, a book that solidified the construct of information literacy, was published. In this work, the authors began to chisel the construct of information literacy from an amalgam of competing and overlapping attributes such as computer literacy, problem solving, critical thinking, and reading. As a result of this and similar works during the 1990s, definitions of information literacy began to take shape.

Information literacy is continually being defined by new resources and studies. Grassian and Kaplowitz (2002) have written a useful resource for individuals considering instruction in this area. They provide a broad discussion of IL, beginning with the history of IL instruction and review theoretical issues in instruction (learning style, critical thinking, and active learning.). The authors then cover instructional design and assessment issues, ending with several chapters on instructional delivery systems for IL. While these topics are not covered in depth, these disparate topics are brought to focus on instruction in information literacy. In the assessment chapter, there is a broad and, in this instance, rather thorough discussion of assessment to include a differentiation of summative and formative evaluation, reliability and validity, and many related topics addressing IL. Unfortunately, there is no discussion of what existing instruments or methods are available to assess it.

As a final note, the National Forum on Information Literacy has a comprehensive web site devoted to this burgeoning area of research (see http://www.infolit.org/publications/index.html). The National Forum discusses the background of IL reviews definitions and relevant publications and provides numerous useful resources.

5.2 Defining Important Information Literacy Outcomes

Definitions of information literacy generally fall into two categories: broad definitions that are global in nature but provide little operational specificity, and narrow definitions that are at times overly concrete but delineate measurable operational skills. Broad definitions usually encompass three general types of skills. *Apprehensive* skills emphasize the ability to recognize the need or become aware of the need for information. *Conative* skills include the range of purposive behaviors that lead to the accession of information. *Comprehensive* skills lead to the understanding of information from varying resources. These broad-based skills together begin to provide a clearer framework for the concept of information literacy, but fall short of providing sufficient specificity to convert the concept into a measurable construct.

Narrow definitions of information literacy begin to stratify the concept into more specific operational skills. These fall into the four categories listed below:

- Variations on scientific or problem-solving models, including the following:

  - The desire or need to know;

  - Hypothesis formation including convergent or analytic skills necessary to formulate questions and the identification of research methodologies and/or sources and the organizational/application skills to organize the information into a usable form;

  - Critical-thinking or synthetic skills necessary to evaluate the results of an information search and integrate it into existing knowledge;

  - A divergent/creative component for formulating alternative or diverse search procedures; and

  - Values clarification as information extends existing values and attitudes.

- Developmental processes as a variation on the scientific model including the following:

  - Assimilative skills to place information into existing cognitive structures;

  - Accommodative skills to extend the knowledge base by the acquisition of new cognitive structures as the result of new kinds of information;

  - The development of executive or metacognitive processes to guide the entire process from the recognition of the need for information through the search-formulation process and the "weeding-out" process and, finally, to the extraction and incorporation of information into existing or new cognitive structures; and

  - The development of values and ethics as the student incorporates information in an existing values structure and learns to use information ethically.

- The technical aspects of information searches, including the following:

  - Conventional IL skills, including the use of a library card catalogue or books in print;

- Computer/IL, including knowledge of search engines, specialized search techniques (e.g., know-bots, and competency with use of the Internet and the World Wide Web);

- Other computer-based search skills including knowledge of proprietary software used in specific libraries and search skills in specific academic disciplines (e.g., Psychinfo, ERIC, e-journals, etc.); and

- Alternative media-based search skills such as television.

- Learner-based and/or personality attributes, including the following:

  - Learning competency and/or independence;

  - Active learning;

  - Self-efficacy and/or self-confidence in learning skills;

  - Internal locus of control in various learning environments;

  - Communication skills;

  - Strong work ethic and high standards of learning;

  - Adaptability and need for change; and

  - Ethical standards of use of information.

In the mid-1980s, just as the technological revolution was leading to the development of the concept of information literacy, the assessment movement as part of educational reform was also emerging. It was only natural that the two movements would cross paths, as accountability became an increasing concern on college campuses. As the assessment movement matured, it sparked competency-based measures on college campuses as a means of tying curricular development with measurable student outcomes. As a result, attempts to define information literacy have also led to competency guidelines for college campuses.

Following a model set forth for public school students, in January 2000, the Association of College and Research Libraries (ACRL) published a set of competency standards for higher education that were subsequently endorsed by the American Association for Higher Education. The five basic competencies for information literacy as they appear in ACRL's publication, *Information Literacy Competency Standards for Higher Education* (2000) are listed below:

- The information literate student determines the nature and extent of the information needed.

- The information literate student accesses needed information effectively and efficiently.

- The information literate student evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.

- The information literate student, individually or as a member of a group, uses information effectively to accomplish a specific purpose.

- The information literate student understands many of the economic, legal, and social issues surrounding the use of information, and accesses and uses information ethically and legally (Breivik 2000).

It is clear from the many facets of existing definitions of information literacy that the concept is both complex and overlaps with a multitude of related constructs. Efforts to measure information literacy have emerged from this complex set of definitions, but they too are many and varied. Any attempt to provide an overview of these measures should provide a systematic categorization of these tests to give the reader a clear picture of precisely what one is trying to measure. To assist the effort to systematize the assessment of information literacy, ACRL also provided guidelines for performance measures in the *Information Literacy* publication. These standards are based on the five competencies listed above and can serve as a guide to examine assessment tools.

The competency standards established by the ACRL are a conglomerate of attributes and skills. They encompass generalized attributes that include components of critical thinking, as well as task-specific measures such as self-efficacy, and cover several of the psychological domains including cognition, conation, and affect. This wide range of included topics may pose a problem for establishing assessment tests in that it will be virtually impossible to find instruments that will cover all of the ground established in the competency standards. However, because the standards are thorough, they may be used as a yardstick against which to measure various assessment approaches.

The standards, according to the ACRL, are purposely varied and cut across all disciplines and domains of academic life. They encompass the concepts of lifelong learning and metacognition (learning to learn or knowing what one knows). They lend themselves to various teaching and learning styles ranging from student-centered learning through problem-based and evidence-based learning to inquiry learning and incorporate elements of critical thinking and problem solving. They include developmental models such as Piaget's stages based on assimilation and accommodation, and they include the entire range of Bloom's Taxonomy from knowledge to evaluation.

The standards are hierarchical in that each of the five standards listed above has been given specificity through several performance indicators that, in turn, have each been given further specificity through several assessment outcomes. Table C provides reviews of assessment instruments for information literacy.

5.3   Key Issues in Assessing Information Literacy

As reflected throughout this Sourcebook, assessment and accountability differ in that the focus of assessment is on improving student learning. This narrows and limits the approach necessary to get at the construct of information literacy because it enables us to think about "power" testing rather than measures of student perception. In their 2002 publication, Hernon and Dugan discuss the differences between assessment and accountability, citing Frye's *Assessment, Accountability, and Student Learning Outcomes* to assist in this task. Frye states that "assessment is a set of initiatives the institution takes to review the results of its actions, and make improvements; accountability is a set of initiatives others take to monitor the results of the institution's actions, and to penalize or reward the institution based on the outcomes" (as cited in Hernon and Dugan 2002).

Methods designed to assess some aspect of information literacy seem to fall into one of two broad categories. First, there are a large number of instruments aimed at measuring some facet of human-computer interaction. These tests are designed either to measure computer skills or the affect generated in subjects when interacting with computers (e.g., computer efficacy). Second, there is an emerging class of

interactive tutorials or learning modules directly aimed at teaching or measuring information literacy as described by the Association of College and Research Libraries (2000).

An interesting epistemological paradox has arisen from these two diverse approaches. On the one hand, tests of computer literacy or skills are, by their very nature, easy to administer and subject to the scrutiny of standardization procedures but do not adequately measure the broad construct encompassed by the concept of information literacy. Even when these computer literacy tests are embedded in computer training modules (e.g., TekExam and SmartForce), they do not focus on the construct of information literacy; rather, their focus is on computer skills. On the other hand, interactive tutorials and learning modules are not, by their very nature, easily validated by standardization procedures and are more complex to administer, yet they are clearly aimed at measuring the diverse construct that is information literacy. It will be necessary to examine instruments from both categories to cover the array of possible measures, but the reader is encouraged to keep the paradox in mind while working through the tables and templates in table C. Furthermore, there is at least one emerging methodology that may lay the paradox to rest. James Madison University is constructing a web-based teaching tool with built-in testing modules that are being psychometrically scrutinized. This system is included in the templates.

It may be helpful to the reader to review the table of standards and outcomes of the ACRL by placing some of the learning modules/tutorials on the horizontal axis. This analysis is provided in table D as a guide to the extent of coverage of this form of instrumentation. It should be clear that this form, while perhaps lacking the depth of normative standardization, provides a broad approach to skill development in information literacy.

5.4   Information Literacy in the Business Community

This chapter would be incomplete without mentioning the burgeoning role of information literacy in the business community, which is becoming increasingly reliant on immediately accessible, current information. However, because IL skills are differentially distributed among decision makers, there may be poor decisions even in the presence of good information. Individuals interested in new concerns in information literacy can read the minutes of the January 25, 2002, meeting of the National Forum on Information Literacy for a complete transcript of Christopher Burn's presentation on information in the private sector (http://www.infolit.org/documents/literary_transcript.doc).

A current concern is that even when data are accurate (as in the three cases cited above), operating in the virtual world of information—where the real events reflected by that information are invisible to the decision maker—may lead to situations where safeguards are ignored and decisions are made in an artificial vacuum.

On the other hand, skilled decision makers check for correspondence between information and the real world, listen for noise in data, and have a healthy respect for uncertainty. These skills then become essential elements of information literacy in business and industry. The problem from an assessment standpoint is that there is little research on information literacy in the workplace in the United States. Without research to define the parameters of the construct in this particular context, it is impossible to develop measures of what one might call "industrial information literacy."

6. QUANTITATIVE LITERACY: QUANTITATIVE REASONING, NUMERACY, AND QUANTITATIVE SKILLS

6.1   Introduction

The purpose of this chapter is to help assessment professionals differentiate among the concepts of quantitative skills, quantitative literacy, and quantitative reasoning. Quantitative skills are generally viewed as pure mathematical abilities (the ability to manipulate mathematical symbols according to rules), that are usually attained by mathematics, science, and engineering majors and that are based in advanced algebra and calculus. Quantitative literacy, on the other hand, is viewed as the minimum knowledge of applied mathematics that all college graduates should have, knowledge attained by students who have eschewed the study of pure mathematics. This concept of quantitative literacy has been most closely associated with the construct of quantitative reasoning as opposed to the purely mathematical interpretation of quantitative skills. In other words, what some call quantitative literacy others call quantitative reasoning and still others call numeracy. It is important to sort these and other related terms and determine reasonable definitions before proceeding to discuss measurement.

By distinguishing quantitative skills from quantitative reasoning, it is possible to establish a clearer link to the different types of assessment instruments that might be used to measure each construct. Quantitative reasoning assessments would be developed to measure problem solving or critical thinking, perhaps using low-level quantification skills as a medium. Quantitative skills assessments, on the other hand, would be developed to measure various levels of pure mathematical ability. For the most part, one would not see a calculus problem on a test of quantitative reasoning unless the test was an ambitious attempt to measure both quantification constructs.

Another distinction is between mathematics reasoning and quantitative reasoning. In seeking to distinguish between the requirements of mathematical reasoning and quantitative reasoning, Devlin (2000) notes critical distinctions between levels of abstraction. According to Devlin (2000, pp. 120–123) there are four levels of abstraction that are progressively removed from real objects in the world so that at the deepest level (Devlin's level 4 abstraction), thought takes place in a state of complete abstraction. His discussion of abstraction provides a method for distinguishing between pure mathematics, with its modern branches such as complexity theory and fractal geometry, quantitative or mathematical skills, with its emphasis on the manipulation of mathematical symbols according to rules, and quantitative reasoning, with its current definition revolving around the logic of mathematics independent of the manipulation of mathematical symbols according to rules. In this lexicon, quantitative reasoning is synonymous with the concept of quantitative literacy, that is, competences that all college students should have upon graduation.

Some researchers in this area, however, have begun to question what they consider to be artificial distinctions that have detached quantitative literacy from mathematics. These researchers consider mathematical reasoning and quantitative literacy as the same thing. As Joan Richards (2001) states, "the concept of quantitative literacy is rooted in the connection between mathematics and reason…When teaching mathematics is seen as a way of teaching people how to think, it can no longer be isolated. Its implications spread throughout the curriculum and it has a place in every class." Indeed, as it is practiced on most college and university campuses, the distinction between quantitative skills and quantitative reasoning is that quantitative skills involve problem solving with mathematics and quantitative reasoning is problem solving without mathematics.

6.2   Defining Important Quantitative Reasoning, Literacy, and Numeracy and Quantitative, and Mathematical Outcomes

A quotation from Lynn Arthur Steen (2001, p. 9) provides further evidence of the complexity of these concepts:

> These elements of basic skills illuminate but do not resolve the linguistic confusions that permeate discussions of quantitative literacy. Sometimes the terms "quantitative" and "mathematical" are used interchangeably, but often they are used to signify important distinctions—for example, between what is needed for life (quantitative) and what is needed for education (mathematics), or between what is needed for general school subjects (quantitative) and what is needed for engineering and physical science (mathematics). For some the word "quantitative" seems too limiting, suggesting numbers and calculation rather than reasoning and logic, while for others the term seems too vague, suggesting a diminution of emphasis on traditional mathematics. Similarly, the term "literacy" conveys different meanings: for some it suggests a minimal capacity to read, write, and calculate, while for others it connotes the defining characteristics of an educated (literate) person.

Definitions also depend on differing expectations for different populations (Gal 1993, p. 4). For example, a 1986 assessment of adult literacy defined quantitative literacy as the knowledge and skills needed to apply arithmetic operations, either singly or sequentially, that are embedded in printed materials, such as in balancing a checkbook, figuring out a tip, completing an order form, or determining the amount of interest from a loan advertisement. The 1991 National Literacy Act defined literacy as an individual's ability to read write and speak in English, and to compute and solve problems at levels of proficiency necessary to function on the job and in society, to achieve one's goals, and develop one's knowledge and potential.

The 1991 report *What Work Requires of Schools* suggested a set of quantitative skills necessary to succeed in the workplace. These included quantitative abilities: arithmetic, seeing with the mind's eye, and reasoning (U.S. Department of Labor 1991). *Arithmetic* was defined as "performs basic computations and approaches practical problems by choosing appropriately from a variety of mathematical techniques." *Seeing with the mind's eye* was defined as "organizes and processes symbols, pictures, graphs, objects, and other information," and *reasoning* was defined as "discovers a rule or principle underlying the relationship between two or more objects and applies it when solving problems."

Faculty have taken yet another route to outlining and defining quantitative literacy. In 2001, the Committee on the Undergraduate Program in Mathematics (CUPM) of the American Mathematical Association determined that the quantitatively literate undergraduate college student should be able to perform the following tasks: (1) interpret mathematical models such as formulas, graphs, tables, and schematics and draw inferences from them; (2) represent mathematical information symbolically, visually, numerically, and verbally; (3) use arithmetic, geometric, and statistical methods to solve problems; (4) estimate and check answers in order to determine reasonableness, identify alternatives, and select optimal results; and (5) recognize the limitations of mathematical and statistical methods (Quantitative Literacy and Service Course Subcommittee 2001). Because this Sourcebook is aimed at those faculty and staff members working in colleges and universities who are attempting to define and measure essential skills, it is reasonable to adopt the definition of quantitative literacy espoused by the CUPM.

In this definition, quantitative literacy is viewed as the minimum knowledge of applied mathematics that all college graduates should have, and this general education approach to quantitative literacy is closely

associated with the construct of quantitative reasoning as opposed to the purely mathematical interpretation of quantitative skills.

For assessment purposes, if one wishes to measure quantitative skills, one must procure an instrument that focuses on the manipulation of mathematical symbols according to rules (e.g., an algebra and/or calculus test that is designed at a sufficiently sophisticated level). In contrast, quantitative reasoning instruments are developed from the precept that one is measuring problem solving or critical thinking skills.

Additionally, the definition used for the array of terms, be it quantitative, mathematical, literacy, reasoning, numeracy, or skills, should reflect the curricular approach of the specific institution. The first question one must ask is, "how broadly are we defining all quantitative/mathematical constructs at my institution?" If the answer is very broadly, that is, quantitative literacy encompasses all mathematical and quantitative reasoning and skills to include abstract mathematical thought, then one would choose a very different approach to assessment than an institution that takes the narrower approach. The narrower approach is the most common approach in that most institutions seem to define quantitative reasoning in terms of problem solving without the computational skills and applied to practical, everyday problems while defining quantitative/mathematical skills in terms of the attributes that involve the manipulation of mathematical symbols according to rules and mathematical thought in terms of abstractions requiring the manipulation of those mathematical symbols.

6.3   Key Issues in Assessing Quantitative Reasoning, Quantitative Literacy, and Related Skills

Individuals involved in setting assessment standards for mathematics and related skills were among those who reevaluated assessment procedures following the call for accountability that began in the early and mid-1980s. By the late 1990s, mathematics professionals had turned their attention to quantitative literacy in its many facets. The National Council of Teachers of Mathematics (NCTM) and the Mathematics Association of America have made an effort to extract a cohesive set of assessment principles from the variety of methods and procedures extant in schools and colleges in much the same way that the ACRL took on the task of establishing clear outcomes for information literacy.

The NCTM established three basic tenets for the assessment of mathematics in the public schools: (1) assessment should reflect the mathematics that is important for individuals to learn; (2) assessment should enhance learning; and (3) assessment procedures should allow students to demonstrate knowledge in a variety of ways. These principles were designed to prevent a disconnect between instruction and assessment (Brosnan and Hartog, 1993; Mathematical Sciences Education Board and National Research Council, 1993).

In a similar fashion, the Mathematics Association of America (MAA), in its five-part document on quantitative literacy (Part IV of which is titled "Assessment") (MAA 1998), made an effort to establish firm assessment guidelines and good assessment practices. The guidelines included the following: (1) conducting and acting on assessment procedures should be a normal part of the teaching-learning process; (2) assessment must be sensitive to reality; and (3) assessment should be based on what is understood about the learning process. In this document, the third principle refers to five aspects of intellectual competency that the MAA felt were involved in quantitative literacy: resources, problem-solving strategies/heuristics, control, beliefs about mathematics and problem solving, and practices. *Resources* referred to knowledge of concepts or facts; *problem-solving strategies* encompassed the metacognitive component that guides problem solving; *control* is related to metacognition and referred to executive-control processes that guide how and when to use resources and problem-solving strategies; *beliefs* about mathematics and problem solving included the affective components of acquiring quantitative literacy

such as fear of mathematics and motivational components; and *practices* meant acquiring good mathematics habits and dispositions.

Good assessment practices were outlined in a seven-step sequence:[1]

- Review the goals set forth which the quantitative literacy program seeks to help students accomplish,

- Review the instructional strategies which resulted in the design of the quantitative literacy program,

- Review the performance standards which have been openly developed and communicated in linking the goals and strategies,

- Choose assessment methods to measure student learning resulting from the instructional strategies,

- Once assessment methods have been executed, summarize what is working, what is not working, and what could be working better in the learning-teaching match,

- Determine changes in courses, experiences, or placement processes, or in the program as a whole which can be implemented to lead to a more effective learning and teaching match, and

- Institute changes and begin the cycle again.

Regardless of how one defines the mathematically related constructs under question, the assessment procedures outlined by the MAA are sound. The sequence reflects a comprehensive assessment program focused on the teaching-learning process.

There has also been a philosophical shift in the way mathematics professionals believe mathematics should be assessed. For example, the SAT now includes items in quantitative reasoning that are not in the traditional multiple-choice format. For the performance section of the test, students are given extensive time to solve a single mathematical problem or task and must provide written justification for their solution.

However, while the current philosophical and theoretical environment seems to favor task-based assessment methods such as course-embedded assessment and the portfolio method, the dominant theme and the most readily available instruments for measuring quantitative reasoning for college or university assessment programs are tests in the traditional multiple-choice format. This Sourcebook looks at examples of both types of tests—multiple choice assessments and task-based assessments.

6.4   Assessment of Quantitative Literacy in the Workplace

Quantitative literacy is generally approached differently in the workplace than in the academic community. The workplace seems more interested in quantitative reasoning with problem-solving at its core, while the academic community seems more interested in literacy tied to traditional mathematical/algebraic computational skills. However, the target population for industry is not college

---

[1] Part IV, p. 2 of the MAA Standards for Quantitative Literacy.

students, but primarily those in vocational programs (high school or other) and adults already employed who may be retraining or changing occupations. Concurrently, collegiate assessment programs are now being asked to focus on quantitative reasoning rather than computational skills, especially in liberal arts programs, though there are few instruments aimed directly at this population. This means adaptations must be made by assessment professionals including using existing tests designed for different audiences upon which norming studies have to be performed.

The measurement of quantitative reasoning for the workplace focuses on aptitude testing; this means that task-specific, criterion-related validity is of prime concern. However, as one examines the literature in this area it quickly becomes apparent that many of the tests used in industrial settings are psychometrically immature in that there are few studies to support the use of a given test. Furthermore, as one might expect, tests used in business settings are more expensive than those designed for use in educational settings. For example, in Great Britain, where quantitative reasoning is dubbed "numeracy," there are a variety of tests used in business settings, but costs run into the hundreds of dollars per test for instruments with no evidence of reliability or validity reported. The Employee Aptitude Survey is an exception in terms of psychometric properties and is included in this Sourcebook.

There are tests included in the templates in table E that are used by both the educational and business communities and that are psychometrically sound. For example, both the Test of Adult Basic Education (TABE) and the Adult Basic Learning Examination (ABLE) have been used by commercial organizations to measure basic adult quantitative reasoning. However, these tests are primarily aimed at placement programs for continuing education rather than as aptitude tests for criterion-specific industrial tasks. The Differential Aptitude Test is designed for a broad range of uses, including "the selection of employees." Again, as the name asserts, it is designed as an aptitude rather than an achievement test. However, any assessment professional whose task includes measuring the readiness for employment in the content areas covered by this test may find it to be a useful instrument.

6.5   Suggested Sources for Further Review

A more detailed examination of quantitative skills and quantitative reasoning, can be found in Devlin (2000), Richards (2001), and the Mathematics Association of America (1998).

Information about sequencing of mathematical training the point of departure for quantitative literacy can be found in a series of surveys and reports by the Committee on the Undergraduate Program in Mathematics of the Mathematics Association of America (particularly Steen's Quantitative Literacy for College Students (available at http://www.stolaf.edu/other/ql/cupm.html).

REFERENCES

Alberti, R.E. (1977). *Assertiveness: Innovations, Applications, Issues*. San Luis Obispo, CA: Impact Publishers.

Allport, G.W., and Odbert, H.S. (1936). Trait names: A psycho-lexical study. *Psychological Monographs, 47*, 211.

American Association for Higher Education (1992). *Principles of good practice for assessing student learning.* Washington DC: American Association for Higher Education.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Anderson, L W., and Krathwohl, D. R, editors. (2001). *A Taxonomy for Learning, Teaching, and Assessing: a Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Association of College and Research Libraries, Information Literacy Competency Standards for Higher Education. (2000). Available: www.ala.org/Content/NavigationMenu/ACRL/Standards_and_Guidelines/Information_Literacy_Competency_Standards_for_Higher_Education.htm

Ayres, J. (1989). The Impact of Communication Apprehension and Interaction Structure on Initial Interactions. *Communication Monographs*, *56:*75-88.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Barr, R.B., and Tagg, J. (1995). From Teaching to Learning: A New Paradigm for Undergraduate Education. *Change, 27*:12-25.

Barry, D. (1992). Managing the Bossless Team: Lessons in Distributed Leadership. *Organizational Dynamics*, *20*(4):31-47.

Bassett, R.E., Whittington, N., and Stanton-Spicer, A. (1978). *The basics in speaking and listening: What should be assessed? Communication Education, 27(4)*: 621-627.

Beatty, M.J. (1988). Situational and Predispositional Correlates of Public Speaking Anxiety. *Communication Education*, *37*:28-39.

Beatty, M.J., and Friedland, M.H. (1990). Public Speaking State Anxiety as a Function of Selected Situational and Predispositional Variables. *Communication Education*, *39:*142-147.

Beatty, M.J., and Payne, S.K. (1985). Is Construct Differentiation Loquacity? A Motivational Perspective. *Human Communication Research, 11(4)*:605-612.

Beatty, M.J., Balfantz, G.L., and Kuwabara, A.Y. (1989). Trait-Like Qualities of Selected Variables Assumed to be Transient Causes of Performance State Anxiety. *Communication Education*, *38*:277-289.

Beatty, M.J., Forst, E.C., and Stewart, R.A. (1986). Communication Apprehension and Motivation as Predictors of Public Speaking Duration. *Communication Education, 35*:143-146.

Bell, R.A., Tremblay, S.W., and Buerkel-Rothfuss, N.L., (1987). Interpersonal Attraction as a Communication Accomplishment: Development of a Measure of Affinity-Seeking Competence. *Western Journal of Speech Communication*, *51*:1-18.

Bennis, W.G., and Nanus, B. (1985). *Leaders: The Strategies for Taking Charge*. New York: Harper & Row.

Bernardin, J., and Cook, D.K. (1995). Review of Leader Behavior Analysis II. In Conoley, J.C. and Impara, J.C., *The Twelfth Mental Measurements Yearbook*. Lincoln, NE: The Buros Institute of Mental Measurements.

Blake, R.R. and Mouton, J.S. (1964). *The Managerial Grid*. Houston, TX: Gulf.

Blanchard, K., Zigarmi, P., and Zigarmi, D. (1985). *Leadership and the One Minute Manager*. New York: William Morrow.

Bloom, B.S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1—The Cognitive Domain*. New York: McKay.

Brandt, D.R. (1979). On Linking Social Performance With Social Competence: Some Relations Between Communication Style and Attributions of Interpersonal Attractiveness and Effectiveness. *Human Communication Research*, *5*:223-237.

Breivik, P.S. (2000). Information literacy and the engaged campus. Giving students and community members the skills to take on (and not be taken in by) the Internet. *AAHE (American Association of Higher Education) Bulletin, 53*:1, 3-6.

Breivik, P.S., and Gee, E.G. (1989). *Information Literacy: Revolution in the Library.* New York: Macmillan Publishing Company.

Brosnan, P.A., and Hartog, M.D. (1993). *Approaching Standards for Mathematics Assessment* [ERIC Document Reproduction Service No. ED359069]. Columbus, OH: Clearinghouse for Science Mathematics and Environmental Education.

Buerkel-Rothfuss, N.L., and Bell, R.A. (1987). Validity of the Affinity-Seeking Instrument. *Communication Research Reports*, *4*(2):24-30.

Buller, D.B., and Aune, R.K. (1988). The Effects of Vocalics and Nonverbal Sensitivity on Compliance: A Speech Accommodation Theory Explanation. *Human Communication Research*, *14*:301-332.

Buller, D.B., and Burgoon, J.K. (1986). The Effects of Vocalics and Nonverbal Sensitivity on Compliance: A Replication and Extension. *Human Communication Research, 13*:126-144.

Buller, D.B., Strzyzewski, K.D., and Comstock, J. (1991). Interpersonal Deception: I. Deceivers' Reactions to Receivers' Suspicions and Probing. *Communication Monographs*, *58*:1-24.

Buller, D.B., LePoire, B.A., Aune, R.K., and Eloy. S.V. (1992). Social Perceptions as Mediators of the Effect of Speech Rate Similarity on Compliance. *Human Communication Research*, *19*:286-311.

Burgoon, J.K., and Hale, J.L. (1987). Validation and Measurement of the Fundamental Themes of Relational Communication. *Communication Monographs*, *54*:19-41.

Burgoon, J.K., and Hale, J.L. (1988). Nonverbal Expectancy Violations: Model Elaboration and Application to Immediacy Behaviors. *Communication Monographs*, *55*:58-79.

Burgoon, J.K., and Newton, D.A. (1991). Applying a Social Meaning Model to Relational Message Interpretations of Conversational Involvement: Comparing Observer and Participant Perspectives. *Southern Communication Journal*, *56*:96-113.

Burgoon, J.K., Coker, D.A., and Coker, R.A. (1986). Communicative Effects of Gaze Behavior: A Test of Two Contrasting Explanations. *Human Communication Research*, *12*:495-524.

Burgoon, J.K., Olney, C.A., and Coker, R.A. (1987). The Effects of Communicator Characteristics on Patterns of Reciprocity and Compensation, *Journal of Nonverbal Behavior*, *11*:146-165.

Burgoon, J.K., Walther, J.B., and Baesler, E.J. (1992). Interpretations, Evaluations, and Consequences of Interpersonal Touch. *Human Communication Research*, *19*:237-263.

Burns, C. (2002). Information Literacy in the Enhancement of Business and Industry Success. Presentation at the National Forum on Information Literacy, Washington, D.C.

Burns, J.M., (1978). *Leadership.* New York: Harper & Row.

Business-Higher Education Forum. (1997). Spanning the Chasm: *Corporate and Academic Program Cooperation to Improve Work-Force Preparation.* Washington, DC: Author.

Canary, D.J., and Spitzberg, B.H. (1987). Appropriateness and Effectiveness Perceptions of Conflict Strategies. *Human Communication Research*, *14:*93-118.

Canary, D.J., and Spitzberg, B.H. (1989). A Model of Perceived Competence of Conflict Strategies. *Human Communication Research*, *15*:630-649.

Canary, D.J., and Spitzberg, B.H. (1990). Attribution Biases and Associations Between Conflict Strategies and Competence Outcomes. *Communication Monographs*, *57:*139-151.

Carnevale, A.P. (2000). *Community Colleges and Career Qualifications*. Washington, DC: American Association of Community Colleges.

Carnevale, A.P., Gainer, L.J., and Meltzer, A.S. (1990). *Workplace Basics: The essential skills employers want.* San Francisco: Jossey-Bass Publishers.

Cattell, R.B. (1971). *Abilities: Their Structure, Growth, and Action*. Boston, MA: Houghton Mifflin.

Cegala, D.J. (1981). Interaction Involvement. A Cognitive Dimension of Communication C*ompetence. Communication Education, 30:*109-121.

Cegala, D.J., Savage, G.T., Brunner, C.C., and Conrad, A.B. (1982). An Elaboration of the Meaning of Interaction Involvement: Toward the Development of a Theoretical Concept. *Communication Monographs*, *49:*229-248.

Chen, G-M. (1989). Relationships of the Dimensions of Intercultural Communication Competence. *Communication Quarterly*, *37*:118-133.

Cline, C.P. (1982). *A Calculating People: The Spread of Numeracy in Early America*. Chicago: University of Chicago Press. Available: http://www.stolaf.edu/other/ql/cohen.html

College Placement Council. (1994). *Developing the Global Workforce: Institute for Colleges and Corporations*. Bethlehem, PA: College Placement Council.

Crawford, W. (1999**)**. *Being Analog: Creating Tomorrow's Libraries*. Chicago: American Library Association.

Cupach, W.R., and Spitzberg, B.H. (1983). Trait Versus State: A Comparison of Dispositional and Situational Measures of Interpersonal Communication Competence. *Western Journal of Speech Communication*, *47:*364-379.

Dalton, M. (1999). *Learning Tactics Inventory*: Facilitator's Guide. San Francisco: Jossey-Bass/Pfeiffer.

Daly, J.A. (1994). Assessing, Speaking and Listening: Preliminary Considerations for a National Assessment. *The National Assessment of College Student Learning: Identification of Skills to be Taught, Learned, and Assessed; A Report on the Proceedings of the Second Study Design Workshop* (NCES 94-286). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Devlin, K. (2000). *The Math Gene*. London: Weidenfeld and Nicolson.

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

Downs, A. (1991). The Relationship Between Communication and Organizational Commitment in Two Australian Organizations. Unpublished master's thesis, University of Kansas, Lawrence.

Downs, C.W., and Hazen, M. (1977). A Factor Analytic Study of Communication Satisfaction. *Journal of Business Communication*, *14:*63-73.

Doyle, C.S. (1992). *Outcomes Measures for Information Literacy Within the National Education Goals of 1990*. Final Report to National Forum on Information Literacy. (ERIC Document Reproduction Service No. 351 033).

Dupuy, T.N. (1968). *Modern Libraries for Modern Colleges*. Washington DC: Communication Service Corp.

Duran, R.L. (1992). Communication Adaptability: A Review of Conceptualization and Measurement. *Communication Quarterly*, 40, 253-268.

Duran, R.L., and Kelly, L. (1985). An Investigation into the Cognitive Domain of Communication Competency. *Communication Research Reports*, *2:*112-119.

Duran, R.L., and Kelly, L. (1988). The Influence of Communicative Competence on Perceived Task, Social, and Physical Attraction. *Communication Quarterly*, *36:*41-49.

Duran, R.L., and Zukahi, W.R. (1984). Competency or Style: What's in a Name. *Communication Research Reports*, *1:*42-47.

Eaton, J.S. (1991). *The Unfinished Agenda: Higher Education and the 1980s*. New York: American Council on Education and the Macmillan Publishing Company.

Education Commission of the States (1995). *Making Quality Count in Undergraduate Education*. Denver, CO: Education Commission of the States.

Educational Testing Service. (1991). *Measures of Speech-Communication*. Princeton, N.J.: Author.

Erwin, T.D. (1991). *Assessing Student Learning and Development*. San Francisco: Jossey Bass.

Erwin, T.D. (2000). *The NPEC Sourcebook on Assessment, Volume 1: Definitions and Assessment Methods for Critical Thinking, Problem Solving, and Writing*. (NCES 2000–195). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative and National Center for Education Statistics. Available: http://nces.ed.gov/pubs2000/2000195.pdf

Erwin, T.D. (2000). *The NPEC Sourcebook on Assessment, Volume 2: Selected Institutions Utilizing Assessment Results*. (NCES 2000–196). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative and National Center for Education Statistics. Available: http://nces.ed.gov/pubs2000/2000196.pdf

Ewell, P. (1997). Organizing for Learning: A New Imperative. *American Association for Higher Education Bulletin. 50(4):*3-9.

Frye, R. (1999). Assessment, Accountability, and Student Learning Outcomes. Available: http://www.ac.wwu.edu/%7Edialogue/index.html

Gabelnick, F. (1997). Educating a Committed Citizenry. *Change*, 30-35.

Gal, I. (1993). *Issues and Challenges in Adult Numeracy*. [Technical Report TR93-15]. Philadelphia, PA: University of Pennsylvania.

Gal, I., and Schmidt, M.J. (Eds). (1994). Reflecting About the Goals of Adult Numeracy Education [PR94-02]. *Proceedings of the 1994 Adult Mathematical Literacy Conference*. Philadelphia, PA: University of Pennsylvania.

Glaser, R. (1994). *Building Negotiating Power*. King of Prussia, PA: Organization Design and Development.

Grassian, E.S., and Kaplowitz, J.R. (2002). *Information Literacy Instruction: Theory and Practice*. New York: Neal-Shuman Publishers, Inc.

Gray, M., and Grace, J. (1997). *Enhancing the Quality and Use of Student Outcomes Data*. (NCES 97–992). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative and National Center for Education Statistics. Available: http://nces.ed.gov/pubs97/97992.pdf

Gregson, T. (1987). An Empirical Investigation of the Relationship Between Communication Satisfaction, Job Satisfaction, Turnover, and Performance for Public Accountants. Ph.D. dissertation., University of Arkansas, 1987. Abstract in *Dissertation Abstracts International*, 48:1254A.

Hackman, M.Z., and Johnson, C.E. (2000). *Leadership: A Communication Perspective*. Prospect Heights, IL: Waveland Press.

Hall, J. (1980). *The Competence Process: Managing for Commitment and Creativity*. The Woodlands, TX: Teleometrics International.

Hall, J. (1992). *The Executive Trap: How to Play Your Personal Best on the Golf Course and on the Job*. New York: Simon & Schuster.

Hallam, G., and Campbell, D. (1994). *TDS Campbell-Hallam Team Development Survey Manual*. Minneapolis, MN: National Computer Systems, Inc.

Hallam, G., and Campbell, D. (1999). *TLP Campbell-Hallam Team Leader Profile: User's Guide*. Minneapolis, MN: National Computer Systems, Inc.

Hargrove, R. (2001). *The Masterful Coaching Feedback Tool: Facilitator Guide*. San Francisco: Jossey-Bass/Pfeiffer.

Hazelton, V., Jr., and Cupach, W.R. (1986). An Exploration of Ontological Knowledge: Communication Competence as a Function of the Ability to Describe, Predict, and Explain. *Western Journal of Speech Communication*, *50:*119-132.

Hecht, M.L. (1978). The Conceptualization and Measurement of Interpersonal Communication Satisfaction. *Human Communication Research*, *4*:253-264.

Hersey, P., and Blanchard, K. (1982). *Management of Organizational Behavior*. Englewood Cliffs, NJ: Prentice Hall.

Hernon, P., and Dugan, R.E. (2002). *An Action Plan for Outcomes Assessment in Your Library*. Chicago and London: American Library Association.

Horney, K. (1970). *Neurosis and Human Growth*. New York: Norton.

Huba, M.E., and Freed, J.E. (2000). *Learner-Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning*. Needham Heights, MA: Allyn and Bacon.

Jessup, H.R. (1990). New Roles in Team Leadership. *Training and Development Journal*, *44:*79-83.

Jones, E.A. (1997). *Goals Inventories: Writing, Critical Thinking, Problem-Solving, Speech Communications, Listening, and Critical Reading*. University Park, PA: National Center on Postsecondary Teaching, Learning, and Assessment.

Jones, E.A., and Voorhees, R.A. (2002). *Defining and Assessing Learning: Exploring Competency-Based Initiatives*. (NCES 2002–159). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative and National Center for Education Statistics. Available: http://nces.ed.gov/pubs2002/2002159.pdf

Jones, J.E. (1973). A Model of Group Development. In J.E. Jones, J.W. Pfeiffer, and L.D. Goodstein (Eds.), *The 1982 Annual Handbook for Group Facilitators*. San Diego, CA: University Associates.

Jones, J.E. (1974). Group Development: A Graphic Analysis. In J.W. Pfeiffer, and J.E. Jones (Eds.), *A Handbook of Structured Experiences for Human Relations Training, Vol. II*. San Diego, CA: University Associates.

Jones, T.S., and Brunner, C.C. (1984). The Effects of Self-Disclosure and Sex on Perceptions of Interpersonal Communication Competence. *Women's Studies in Communication*, *7:*23-37.

Kelley, C. (1979). *Assertion Training: A Facilitator's Guide*. San Diego, CA: University Associates.

Kelley, D.L., and Burgoon, J.K. (1991). Understanding Marital Satisfaction and Couple Type as Functions of Relational Expectations. *Human Communication Research*, *18:*40-69.

Kinlaw, D.C. (1989). *Coaching for Commitment: Managerial Strategies for Obtaining Superior Performance*. San Francisco: Jossey-Bass/Pfeiffer.

Kinlaw, D.C. (1990). *Developing Superior Work Teams*. San Francisco: Jossey-Bass/Pfeiffer.

Kinlaw, D.C. (1999). *Coaching for Commitment: Trainer's Guide*. San Francisco: Jossey-Bass/Pfeiffer.

Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels.* Alexandria, VA: American Society for Training and Development.

Knapp, M.L., and Hall, J.A. (1992). *Nonverbal Communication in Human Interaction.* Fort Worth, TX: Holt Rinehart and Winston.

Kolb, D. (1984). *Experiential Learning.* Englewood Cliffs, NJ: Prentice-Hall.

Kotter, J.P. (1990). *A Force for Change: How Leadership Differs from Management.* New York: Free Press.

Kouzes, J.M., and Posner, B.Z. (1995). *The Leadership Challenge: How to Get Extraordinary Things Done in Organizations.* San Francisco: Jossey-Bass.

Krathwohl, D.R., Bloom, S.B., and Masia, B.B. (1964). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 2: Affective Domain.* New York: McKay.

Larson, C.E., and LaFasto, F.M.J. (1989). *TeamWork: What Must Go Right/What Can Go Wrong.* Newbury Park, CA: Sage.

Leonelli, E., and Schwendeman, R. (1994). *Adult Basic Math Instruction: Massachusetts Practitioners' Viewpoints on the ABE Learner and the Instructional Environment.* Proceedings of the 1994 Adult Mathematical Literacy Conference, March 20-22, Arlington, VA.

Lewin, K. (1936). *Principles of Topological Psychology.* New York: McGraw-Hill Book Company.

Loertscher, D., and Woolls, B. (1997). *The Information Literacy Movement of the School Library Field: A Preliminary Summary of the Research.* Paper presented at the Annual Conference of the International Association of School Librarianship, Vancouver, BC. (ERIC Document Reproduction Service No. ED 412 972)

Likert, J.G., and Araki, C.T. (1986). Managing Without a Boss: System 5. *Leadership and Organization Development Journal, 73:*17-20.

Maduschke, K.M., and Grummon, P.T.H. (1996). *Technical Documentation to Working.* Lansing, MI: Public Policy Research, Development, and Evaluation.

Manz, C.C., and Sims, H.P., Jr. (1980). Self-Management as a Substitute for Leadership: A Social Learning Theory Perspective. *Academy of Management Review, 5*(3):361-367.

Manz, C.C., and Sims, H.P., Jr. (1984). Searching for the "Unleader." Organization Member Views on Leading Self-Managed Groups. *Human Relations, 37*(5):409-424.

Mathematical Sciences Education Board. (1991). *For Measure: Principles and Goals for Mathematics Assessment.* Washington, DC: Report of the National Summit on Assessment of Mathematics.

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

Mathematics Association of America (MAA). (1998). *Quantitative Reasoning for College Graduates: A Complement to the Standards.* Available: http://www.maa.org/past/ql/ql_toc.html

McCroskey, J.C., and McCain, T.A. (1974). The Measurement of Interpersonal Attraction. *Speech Monographs, 41*:261-266.

McCroskey, J.C., Richmond, V.P., Daly, J.A., and Cox, B.G. (1975). The Effects of Communication Apprehension on Interpersonal Attraction. *Human Communication Research*, *2*:51-65.

McGregor, D. (1960). *The Human Side of Enterprise*. New York: McGraw-Hill.

McLaughlin, M.L., and Cody, M. J. (1982). Awkward Silences: Behavioral Antecedents and Consequences of the Conversational Lapse. *Human Communication Research*, *8*:299-316.

Mezirow, J. (1990). *Fostering Critical Reflection in Adulthood: A Guide to Transformative and Emancipatory Learning*. San Francisco: Jossey-Bass.

Miles, C. (1994). *The Mindful Worker: Learning and Working in the 21st Century*. Clearwater, FL: H&H Publishing Company.

Morreale, S.P., and Backlund, P. (Eds.). (1996). *Large Scale Assessment in Oral Communication: K-12 and Higher Education*. Annandale, VA: National Communication Association.

Morreale, S.P., Rubin, R.B., and Jones, E. (1998). *Speaking and Listening Competencies for College Students*. Available: http://www.natcom.org/instruction/assessment/collegecomp/college%5fcompetencies.htm

Murphy, K.R. (1995). Review of Management Inventory on Leadership, Motivation and Decision-Making. In Conoley, J.C. and Impara, J.C., (Eds.), *The Twelfth Mental Measurement Yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.

Nanus, B. (1992). *Visionary Leadership*. San Francisco: Jossey-Bass.

National Communication Association. (1998). *Speaking and Listening Competencies for College Students*. Washington, DC: National Communication Association.

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

National Forum on Information Literacy. Minutes of the January 2003 meeting.

Oblinger, D.G., and Verville, A.L. (1998). *What Business Wants from Higher Education*. Phoenix, AZ: The Oryx Press.

O'Connor, L.G., Radcliff, C.J., and Gedeon, J.A. (2002). *Applying Systems Design and Item Response Theory to the Problem of Measuring Information Literacy Skills*. College and Research Libraries, 63(6), 528-543.

Palumbo, C.A., and Banta, T.W. (1999). *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco: Jossey-Bass.

Palumbo, C.A., and Banta, T.W. (Eds.). (2001). *Assessing Student Competence in Accredited Disciplines: Pioneering Approaches to Assessment in Higher Education*. Sterling, VA: Stylus Publishing.

Parker, G.M. (1990). *Team Players and Teamwork: The New Competitive Business Strategy*. San Francisco: Jossey-Bass.

Parker, G.M. (1996). *Team Players and Teamwork*. San Francisco: Jossey-Bass/Pfeiffer.

Parry, S.B. (1997). *Evaluating the Impact of Training: A Collection of Tools and Techniques*. Alexandria, VA: American Society for Training and Development.

Perotti, V.S., and DeWine, S. (1987). Competence in Communication: An Examination of Three Instruments. *Management Communication Quarterly*, *1*:272-287.

Peters, T. (1992). *Liberation Management*. New York: Ballantine.

Peters, T.J., and Waterman, R.H., Jr. (1982). *In Search of Excellence*. New York: Harper and Row.

Peters, T.M., and Austin, N.K. (1985). *A Passion for Excellence: The Leadership Difference*. New York: Warner Books.

Portnoy, F.C., and Russell, J.F. (1986). *Leadership: What Every Leader Should Know About People*. Englewood Cliffs, N.J.: Prentice-Hall.

Potvin, T.C. (1992). Employee Organizational Commitment: An Examination of its Relationship to Communication Satisfaction and an Evaluation of Questionnaires Designed to Measure the Construct. Ph.D. Dissertation, University of Kansas, 1991. Abstract in *Dissertation Abstracts International*, *52*:4147A.

Quantitative Literacy and Service Course Subcommittee. (July 2001). *On the Curriculum for Students in Non-Mathematically Intensive Disciplines.* Available: http://www.valpo.edu/home/faculty/rgillman/ql/

Query, J.L., Parry, D., and Flint, L.J. (1992). The Relationship Among Social Support, Communication Competence, and Cognitive Depression for Nontraditional Students. *Journal of Applied Communication Research*, *20*:78-94.

Quianthy, R.L. (1990). *Communication is Life: Essential College Sophomore Speaking and Listening Competencies.* Annandale, VA: National Communication Association.

Rao, M. and Sylvester, S. (2000). Business and Education in Transition: Why New Partnerships are Essential to Student Success in the New Economy. *AAHE Bulletin*, *52*(8):11-13.

Redding, J.C. (2000). *The Radical Team Handbook*. San Francisco: Jossey-Bass/Pfeiffer.

Rees, F. (1997). *Teamwork From Start to Finish*. San Francisco: Jossey-Bass/Pfeiffer.

Revans, R.W. (1980). *Action Learning*. London: Blond and Briggs.

Richards, J.L. (2001). *Connecting Mathematics With Reason*. Available: http://www.woodrow.org/nced/mathematicsdemocracy.html

Riechmann, D. (1998). *Team Performance Questionnaire.* San Francisco: Jossey-Bass Publishers.

Roberts, C.V. (1986). The Question of Validity and Application: How Do We Know How Well We Are Listening? Paper presented at the 77th annual meeting of the Eastern Communication Association. Atlantic City, NJ.

Rubin, R. (1982). Assessing Speaking and Listening Competence at the College Level: The Communication Competency Assessment Instrument. *Communication Education. 31(1)*:19-32.

Rubin, R.B., and Graham, E.E. (1988). Communication Correlates of College Success: An Exploratory Investigation. *Communication Education*, *37*:14-27.

Rubin, R.B., and Morreale, S.P. (1996). Setting Expectations for Speech Communication and Listening. In E.A. Jones (Ed.). Preparing Competent College Graduates: Setting New and Higher Expectations for Student Learning. *New Directions for Higher Education. 96*, 19-29.

Rubin, R.B., Palmgreen, P., and Sypher, H.E. (Eds.). (1994). *Communication Research Measures: A Sourcebook*. New York: The Guilford Press.

Schraw, G., and Dennison, R.S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*:460-475.

Schwartz, M.K., Axtman, K.M., and Freeman, F.H. (1998). *Leadership Education: A Sourcebook of Courses and Programs.* Greensboro, NC: Center for Creative Leadership.

Schwartz, S.W., and Lucas, N. (1998). Leadership Education in the Senior Year Experience. In Gardner, J.N., Van der Veer, G. and Associates, *The Senior Year Experience: Facilitating Integration, Reflection, Closure, and Transition*. San Francisco: Jossey-Bass.

Shea, G. (1983). *Creative Negotiating*. Boston: CBI.

Snavely, W.B. (1981). The Impact of Social Styles Upon Person Perception in Primary Relationships. *Communication Quarterly*, *29*:132-143.

Spitzberg, B.H., and Cupach, W.R. (1988). *Handbook of Interpersonal Competence Research*. New York: Springer-Verlag.

Steen, L.A. (1989). *Quantitative Literacy for College Students: A Memorandum for CUPM.* Available: http://www.stolaf.edu/other/ql/cupm.html

Steen, L.A. (2001). *Mathematics and Democracy: The Case for Quantitative Literacy.* Princeton, NJ: The Woodrow Wilson National Fellowship Foundation, National Council on Education and the Disciplines.

Stogdill, R.M. (1948). Personal Factors Associated With Leadership: A Survey of the Literature. *Journal of Psychology*, *25*:35-71.

Street, R.L., Jr., Mulac, A., and Weimann, J.M. (1988). Speech Evaluation Differences as a Function of Perspective (Participant Versus Observer) and Presentational Medium. *Human Communication Research*, *14*:333-363.

Suskie, L. (1996). *Questionnaire Survey Research: What works?* (2nd edition). Tallahassee, Florida: Association for Institutional Research.

Terenzini, P. (1997). *Student Outcomes Information for Policy-Making*. (NCES 97–991). U.S. Department of Education. Washington, DC: National Postsecondary Education Cooperative and National Center for Education Statistics. Available: http://nces.ed.gov/pubs97/97991.pdf

Urbina, S. (1998). Review of INSIGHT Inventory. In Impara, J.C. and Plake, B.S. (Eds.), *The Thirteenth Mental Measurements Yearbook*. Lincoln, Nebraska: The Buros Mental Measurement Yearbook.

U.S. Department of Education. (2000). *The Use of Tests as Part of High-Stakes Decision-Making for Students: A Resource Guide for Educators and Policy-Makers*. Washington, DC: United States Department of Education.

U.S. Department of Labor. (1991). *What Work Requires of Schools*. Washington, DC. Available: http://wdr.doleta.gov/SCANS/whatwork/whatwork.html

Walther, J.B., and Burgoon, J.K. (1992). Relational Communication in Computer-Mediated Interaction. *Human Communication Research*, *19*:50-88.

Walvoord, B.E., and Anderson, V.J. (1998). *A Tool for Learning and Assessment*. San Francisco: Jossey-Bass.

Waterhouse, S. (2005). *The Power of E-learning: the Essential Guide for Teaching in the Digital Age*. Boston: Allyn and Bacon.

Weimann, J.M. (1977). Explication and Test of a Model of Communicative Competence. *Human Communication Research*, *3*:195-213.

Wheeless, L.R., and Grotz, J. (1977). The Measurement of Trust and its Relationship to Self-Disclosure. *Human Communication Research*, *3*:250-257.

Wheeless, L.R., Frymier, A.B., and Thompson, C.A. (1992). A Comparison of Verbal Output and Receptivity in Relation to Attraction and Communication Satisfaction in Interpersonal Relationships. *Communication Quarterly*, *40*:102-115.e

White, S. (1981). *The New Liberal Arts*. New York: Alfred P. Sloan Foundation.

Willmington, S.C. (1992, October). *The Validity and Reliability of a Performance Test of Interpersonal Communication Proficiency*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Wilson, C.L. (1981). *Survey of Management Practices (Form SMP-SE)*. New Canaan, CT: Clark Wilson Publishing.

TABLE A

**ASSESSMENT REVIEWS FOR COMMUNICATION,
LISTENING, AND INTERPERSONAL SKILLS**

Table A-1.   Reviews of assessments of communication competency in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Presentation Skills Profile 24 items **Author** Ian MacDonald **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Date** 1997 **Testing Time** 10 to 15 minutes scoring time 60 to 90 minutes interpretation and discussion **Cost** $63.00 Participant guide, feedback form, and facilitator guide	Total score and six subscores: (1) Objectives (2) Audience (3) Structure (4) Impact (5) Visual aids (6) Stage	 Stating goals and evaluating presentation. Analyzing your audience. Designing a clear and logical structure. Creating aural impact and creating visual impact Effectively using visual aids. Preparing and delivering high-performance presentations.	None reported.	None Reported.	Pilot tested with business professionals and business college students.	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. In higher education, it could be used in courses to help students better understand effective practices in preparing and delivering high quality presentations. Instrument is easy to administer and score. **Limitations** Instrument has little psychometric evidence.

Table A-1.   Reviews of assessments of communication competency in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Communication Competency Self-Report Questionnaire (CCSR) 19 items **Author** R.B. Rubin **Original Instrument Located In:** R.B. Rubin. (1985). The Validity of the Communication Competency Assessment Instrument. *Communication Monographs*, 52, 173-185. **Testing Time** 15 to 20 minutes	Total score Examines abilities– (1) Public speaking (2) Interaction (3) Listening	Designed to assess college students' self-perceptions of their own communication competence.	Alpha coefficient of .87 was reported.	None re+ported.	CCSR reflects the 19 functional communication competencies approved by the National Communication Association.	The CCSR correlated with the Personal Report of Communica-tion Apprehension.	**Strengths** The 19 items provide information about the students' perceptions of communication abilities in several situations. Could be used as a pre- and postmeasure for a course or major and may help to identify changes in perceptions (Morreale and Backlund 1996). Administration and scoring are highly standardized (Morreale and Backlund 1996). **Limitations** Validity studies indicate that self-perceptions may not be accurate measure of communication competence (Morreale and Backlund 1996).

Table A-1.   Reviews of assessments of communication competency in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
The Competent Speaker Speech Evaluation Form 19 items **Authors** S.P. Morreale, K.P. Moore, D.S. Tatum and R. Hulbert-Johnson **Publisher** National Communication Association 1765 N Street NW Washington, DC 20036 **Testing Time** Length of assigned speech plus an additional 10 minutes. Requires approximately 2 hours of training. **Cost** $17.50 for members of National Communication Association $22.50 for nonmembers	Eight competencies are assessed: (1) Topic (2) Purpose (3) Supporting material (4) Organization (5) Language (6) Voice (7) Usage (8) Physical behaviors	Designed to assess college students' public speaking performance. Training manual includes discussion of each competency and an explanation of how each would be demonstrated at excellent, satisfactory, and unsatisfactory levels. Instructions are included for preparing a video tape to demonstrate different levels of student performance along with information about how other speech communication educators evaluated sample speeches.	High inter-rater reliability was reported after training of assessors.	None reported.	Developers conducted extensive literature review to determine appropriate competencies and criteria. Panel of 11 speech communication educators was involved in final version.	A positive correlation was reported with seven public speaking items on Communication Competency Assessment instrument.	**Strengths** Instrument has multiple purposes, including evaluate informative and persuasive speeches in class, testing-in or testing-out placement purposes, tool for instructing and advising students, and generate assessment data for departmental or institutional accountability. **Limitations** Current instrument does not link with higher order skills, such as critical thinking, because major components of preparing speeches and delivering them.

Table A-1.   Reviews of assessments of communication competency in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Basic Course Communication Competency Measure 24 items **Authors** W.S.Z. Ford and A.D. Wolvin **Original Instrument Located In:** W.S.Z. Ford and A.D. Wolvin. (1993). The Differential Impact of a Basic Communication Course on Perceived Communication Competencies in Class, Work, and Social Contexts. *Communication Education*, 42, 215-223. **Testing Time** 15 to 20 minutes	Total score Self-perceptions about: (1) Public speaking (2) Interpersonal communication (3) Interviewing (4) Group communication (5) Listening self-confidence	Designed to assess college students' self-perceptions of their own communication competence.	Cronbach alpha coefficients ranged from .93 to .95 on the three context subscales—class, social/family, and work.	None reported.	Students who completed the instrument on a pre- and postcourse basis demonstrated significantly higher scores after completion of the course.	None reported.	**Strengths** This instrument is a good indicator of perceived change in communication abilities. It can be used for the communication course (Morreale and Backlund 1996). **Limitations** Because instrument is self-report, the outcomes may not correspond with the actual development of these communication abilities (Morreale and Backlund 1996).

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Communicative Competence Scale (CCS) 36 items **Author** J.M. Wiemann **Original Instrument Located In:** J.M. Weimann. (1977). Explication and Test of a Model of Communicative Competence. *Human Communication Research*, 3, 195-213. **Testing Time** Less than 5 minutes	Total score and five subscores: (1) General Competence (2) Empathy (3) Affiliation/ Support (3) Behavioral Flexibility (4) Social Relaxation	Used to assess dimensions of interpersonal competence. Instrument assesses another person's communicative competence by responding to items using Likert scales that range from strongly agree "5" to strongly disagree "1". Can also be used as self-report.	.96 coefficient alpha (and .74 magnitude of experimental effect) (Wiemann 1977). .94 to .95 overall alpha with subscale scores ranging from .68 to .82 (Jones and Brunner 1984). .84 alpha (Street, Mulac, and Wiemann 1988). On self-report version, alpha of .90 (Cupach and Spitzberg 1983). Alpha of .91 (Hazleton and Cupach 1986). Alpha of .86 (Query, Parry, and Flint 1992).	None reported.	Evidence of construct validity (McLaughlin and Cody 1982; Street, Mulac, and Wiemann 1988).	All three dimensions of Interaction Involvement Scale scores positively correlated with CCS (Cegala et al. 1982). Strongly correlated with communication adaptability and trait self-rated competence (Cupach and Spitzberg 1983).	**Strengths** Scale used with college students only. Instrument can be completed quickly. Instrument can help college students understand their communication competence. It has strong reliability data. **Limitations** Perotti and DeWine (1987) recommend that instrument be used as composite measure of communicative competence rather than breaking scores into the five subareas. There are some issues with the factor structure.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Team Development Survey 12 items **Author** Glenn M. Parker **Publisher** Consulting Psychologists Press, Inc. 3803 East Bayshore Rd. P.O. Box 10096 Palo Alto, CA 94303 (800) 624-1765 http://www.cpp-db.com **Date** 1992 **Testing Time** Administration time not reported. **Cost** $46 per package includes leader's guide, 10 team member surveys, and team scoring form.	Means reported for each item. Author suggests that means of 3.0 or higher be considered "high" and those below should be considered "low." No rationale or supporting data given to justify these suggestions.	Two Likert scales per item, one "description" scale ranging from "1" (strongly disagree) to "5" (strongly agree)—the degree to which the statement accurately describes the respondent's work team; second scale is focused on "importance" ranging from "1" (unimportant) to "5" (critically important)—the extent to which characteristic is judged to be important for success of the team.	No reliability data reported.	None reported.	Evidence only for content validity is based on instrument's derivation from several sources, including work done by McGregor (1960) and Parker (1990).	None reported.	**Strengths** Instrument can be useful to stimulate discussion about team members' perceptions on various topics. This tool could be an instructional aid. **Limitations** Author provides evidence of content validity but lacks evidence of reliability and other types of validity.

Table A-2. Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Effective Team Member Profile 36 items **Author** HRDQ Research and Development Team **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 624-1765 http://www.hrdq.com **Date** 2001 **Testing Time** 15 minutes; 1 to 1½ hours interpretation and discussion **Cost** $73 Preview Pack (participant guide, feedback form, and facilitator guide)	Six subscores: (1) Understanding team direction (2) Clarifying team roles (3) Showing commitment (4) Encouraging open communi-cation (5) Learning continuously (6) Sharing leadership responsibilities	Clear understanding of team's purpose and direction. Understanding how a team member's efforts contribute to the team's achievement of a goal. Commitment to team's purpose. Demonstration of respect and value of others' opinions and perspectives by truly trying to understand what someone has to say. Learning new skills, behaviors, and information. Looking for ways to help out others and take action when it is necessary to bolster individual and team performance.	Split half coefficient is .87	Factor analysis was conducted and coeffici-ents for each factor are: Direction, .708 Structure, .669 Commit-ment, .526 Commun-ication, .452 Learning .689 Leader-ship, .650	Face validity determined by expert panel of training and development practitioners. Extensive details about model that was developed based on previous research conducted by Larson and LaFasto (1989), Parker (1996), Riechmann (1998), Redding (2000), and Rees (1997).	None reported.	**Strengths** Assessment designed for individuals who are members of teams. Primary purpose is to enable individuals to determine their strengths and areas for improvement as team members. Although reviewed by training professionals, items in survey are relevant and could be applied to college students working in teams. Further refinements of communication factor will occur in future revisions of instrument.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Group Development 40 items **Authors** John E. Jones and William L. Bearly **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Dates** 1985, 1986, 1993 **Testing Time** 15 minutes **Cost** $36.00 per pack (5 instruments in each pack)	Eight subscores: Task behaviors (1) Orientation (2) Organization (3) Open Data Flow (4) Problem Solving Process behaviors (5) Dependency (6) Conflict (7) Cohesion (8) Interdependence	Learning what is expected of the group. Making choices about how to organize the work. Sharing information that is relevant to the task. Using the information to make decisions. Group members are dependent on designated leader. Group members experience difficulty with each other. Group members are open with each other. Team members organize themselves in highly flexible ways.	None reported.	None reported.	Instrument based on theory of group development originally proposed by Jones (1973, 1974).	None reported.	**Strengths** Individuals assess their own group. All items are relevant for college students who work in teams. Primary purpose is to enable individuals to determine the strengths and areas for improvement for their team. **Limitations** Instrument has little reliability and validity evidence.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Team Effectiveness Survey 12 items **Author** Jay Hall **Publisher** Teleometrics International 1755 Woodstead Court The Woodlands, TX 77380-0964 http://www.teleometrics.com **Date** 1968-86 **Testing Time** Administration time not reported. **Cost** $8.95 per test booklet.	Total score plus four subscores for both individual and team: (1) Exposure (2) Feedback (3) Defensive Climate (4) Supportive Climate	Instrument designed for organizational development purposes. Tendency to engage in open expressions of one's feelings and knowledge. Tendency to solicit information from others about their feelings and knowledge. Tendency to have a constraining effect on team. Foster feelings of insecurity, vulnerability, and lack of trusts among members. Tendency to help team work effectively and encourages feelings of well-being and warmth.	No reliability evidence reported.	None reported.	None reported.	None reported.	**Strengths** Instrument can be useful to stimulate discussion about team members' perceptions on various topics. This tool could be an instructional aid. Discussion of findings from this instrument should be guided by experienced facilitator. **Limitations** Instrument has little psychometric evidence.

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Team Performance Questionnaire 32 items **Author** Donna Riechmann **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 1998 **Testing Time** Administration time not reported. **Cost** $14.00 Team Performance Questionnaire and Team Development Workbook	Total score plus six subscores for team performance: (1) Goals and results (2) Collaboration and involvement (3) Competencies (4) Communication process (5) Emotional climate (6) Leadership	Instrument designed for organizational development purposes. Roles and responsibilities are clearly defined and agreed upon. Goals and outputs are evident to team members. Members feel a sense of belonging to team. There is open and honest exchange of ideas. Members draw on each other's strengths. They develop their skills and abilities. Interpersonal and team interactions are open and effective. People feel strong sense of commitment to team. Provides a vision and communicates it.	Test-retest correlations for six subscales reported ranging from .71 for goals and results to .90 for collaboration and involvement. Internal consistency ranged from .77 to .89.	Initial factor analysis—eight redundant or confusing items were dropped. Final factor analysis, remaining 32 items were organized into six factors.	Content validity—expert panel reviewed all items in instrument. Instrument builds on research studies conducted by Dr. Riechmann.	None reported.	**Strengths** Instrument can be useful to stimulate discussion about team members' perceptions of their work group characteristics and level of performance. The instrument can also help teams identify opportunities for improvement. This tool could be an instructional aid. **Limitations** Discussion of findings from this instrument should be guided by experienced facilitator.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Assessing Your Team 32 items **Authors** Dick Richards and Susan Smyth **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 1994 **Testing Time** Administration time not reported. **Cost** $15.00 Team Member's Manual and Instrument	Total score plus seven subscores for team performance: Purpose (1) Role (2) Strategy (3) Processes (4) People (5) Feedback (6) Interfaces	Team is clear about their purpose. Identification of team's role. Team members clear about how to proceed to achieve their purpose. Team has support mechanisms in place to pay attention to interactions of its members. Members feel valued and well used. Members give and receive feedback. Aware of groups and individuals important to team's success.	Split-half reliability test is high (.96).	Test of item discrimination showed that 28 items on the assessment discriminate positively among teams that score high in team well-being and those that score low.	Ratings by team participants and by independent raters demonstrated significant correlations.	None reported.	**Strengths** Instrument can be useful to stimulate discussion about team members' perceptions of their work group characteristics and level of performance. The instrument can also help teams identify opportunities for improvement. This tool could be an instructional aid. **Limitations** Discussion of findings from this instrument should be guided by experienced facilitator.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
WORKKEYS— Teamwork Two-part assessment contains 12 scenes each followed by three questions **Publisher** ACT—WORKKEYS 2201 North Dodge Street P.O. Box 168 Iowa City, IA 52243-0168 (800) 967-5539 http://www.act.org/workkeys/ contacts.html **Testing Time** 65 minutes (recommend splitting total time into two sessions) **Cost** $6.00 per student (includes scoring)	Level 6 Level 5 Level 4 Level 3	Identify focus of team activity and choose approaches that encourage a team to act as a unit. Determine best use of team talents. Consider and evaluate possible effects of alternative behaviors on team relationships and completion of tasks. Identify the organization of tasks and time schedule and recognize need for commitment to quality and sensitivity to customers. Identify team goals and ways to work with team members to accomplish team goals and recognize need for trust and dependability in a team environment.	Reliability =.59 Consult ACT.	Consult ACT.	All test items submitted to external consultants for content and fairness reviews. Consult ACT.	Consult ACT.	**Strengths** Focuses on skills deemed important by employers. Actual scripts focus on workplace situations. Provides students with recommendations about how to reach higher level of teamwork. **Limitations** Difficult to identify based upon limited information provided by ACT.

Table A-2.   Reviews of assessments of teamwork in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
TDS Campbell-Hallam Team Development Survey 72 items **Authors** Glenn Hallam and David Campbell **Publisher** National Computer Systems, Inc. Workforce Development Group 9701 West Higgins Rd. Rosemont, IL 60018 (800) 221-8378 http://pearsonreidlondonhouse.com/tests/tds.htm **Dates** 1992, 1994 **Testing Time** Administration time not reported. **Cost** $180.00 Preview package $15.00 Member survey and report	Total score plus subscores for team performance: (1) Resources (2) Efficiency (3) Improvement (4) Success	Instrument designed for organizational development purposes. Skills, commitment, information, time and staffing, material resources. Team unity, individual goals, empowerment, team coordination. Team assessment, innovation, feedback, leadership, rewards. Satisfaction, performance, and overall index.	Test-retest correlations ranged from .69 to .90 with a median of .80. Internal consistency reliabilities averaged .73 with a median of .69. (Hallam and Campbell 1994).	Unavail-able	The team performance scale score was correlated with the team leader's performance score and the observer's performance score at .70 (Hallam and Campbell 1994).	Unavailable	**Strengths** Extensive psychometric data reported in technical manual, including scale scores by demographic subgroups (sex, team role, length of time on team) Instrument can be useful to stimulate discussion about team members' perceptions of their work group characteristics and level of performance. The instrument can also help teams identify opportunities for improvement. This tool could be an instructional aid. **Limitations** Discussion of findings from this instrument should be guided by experienced facilitator.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Conflict Management Appraisal (CMA) —Assessment by Others 60 items **Author** Jay Hall **Publisher** Teleometrics International 1755 Woodstead Court The Woodlands, TX 77380-0964 http://www.teleometrics.com **Date** 1986 **Testing Time** 30 minutes **Cost** $7.95 per instrument	Total score for each of the five conflict management styles Across the four contexts, raw scores are summed to obtain total raw scores on the five styles. Rate each item regarding conflict by using a 10-point scale ranging from completely uncharacteristic "1" to completely characteristic "10." Ratings may be compared with self-ratings obtained on the Conflict Management Survey (a companion instrument).	9/9 Synergistic; 5/5 Compromise; 1/9 Yield/Lose; 9/1 Win-Lose; 1/1 Lose-Leave. First number in each pair represents degree of concern for personal goals and second number represents the degree of concern for relationship. Twelve items for each overall style score are evenly distributed across four contexts: personal orientation, interpersonal relationships, small group relationships, and intergroup relationships. A rater evaluates an associate (coworker, a personal or social acquaintance, or family member) in terms of how the individual handles conflict.	"Testing for internal consistency of CMA, an item analysis yielded a mean Cronbach Alpha of .81; the median alpha was .85" (test manual, p. 20).	None reported.	Background of test developer affects content validity. Author has significant experience in the field for developing the instruments as highlighted by two books, *The Executive Trap* (1992) and *The Competence Process* (1980).	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. It could be used to stimulate discussion about team building or improving relations. **Limitations** It lacks a technical manual that addresses important issues, such as validity studies, interscale correlations, and interrater agreement.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Conflict Management Survey, Self-Assessment 60 items **Author** Jay Hall **Publisher** Teleometrics International 1755 Woodstead Court The Woodlands, TX 77380-0964 http://www.teleometrics.com **Dates** 1969-1986 **Testing Time** 30 minutes **Cost** $8.95 per instrument	Total score for each of the five conflict management styles Across the four contexts, raw scores are summed to obtain total raw scores on the five styles. Rate each item regarding conflict by using a 10-point scale ranging from completely uncharacteristic "1" to completely characteristic "10." Ratings may be compared with self-ratings obtained on the Conflict Management Survey (a companion instrument).	9/9 Synergistic; 5/5 Compromise; 1/9 Yield/Lose; 9/1 Win-Lose; 1/1 Lose-Leave. First number in each pair represents degree of concern for personal goals; second number represents the degree of concern for relationship. Twelve items for each overall style score are evenly distributed across four contexts: personal orientation, interpersonal relationships, small group relationships, and intergroup relationships. A rater evaluates an associate (coworker, a personal or social acquaintance, or family member) in terms of how the individual handles conflict.	Split-half coefficients ranged from .70 to .87 reported for the five styles.	None reported.	Background of test developer affects content validity. Author has significant experience in the field for developing the instruments as highlighted by two books, *The Executive Trap* (1992) and *The Competence Process* (1980).	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. It could be used to stimulate discussion about team building or improving relations. **Limitations** Instrument lacks a technical manual that addresses important issues such as validity evidence, including construct validity.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Working: Assessing Skills, Habits, and Style 50 items **Authors** Curtis Miles and Phyllis Grummon **Publisher** H and H Publishing Company, Inc. 1231 Kapp Drive Clearwater, FL 33765-2116 (800) 366-4079 http://www.hhpublishing.com **Date** 1996 **Testing Time** 30 minutes **Cost** 1-99 instruments=$4.00 per instrument or 100+=$3.50 per instrument	Nine subscores: (1) Taking responsibility (2) Working in teams (3) Persisting (4) A sense of belonging (5) Life-long learning (6) Adapting to change (7) Permanent problem solving (8) Information processing (9) Systems thinking	Instrument designed for students' self-assessment. Taking personal responsibility for assigned tasks. Paying attention to goals of group and social processes used to accomplish goals. Staying with a task until completion. Taking pride in one's work. Learning throughout their lifetime. Level of comfort with changes. Interest and skill at using systematic problem solving skills. Using multiple strategies and managing their own learning. Understanding of relationship among parts in a system and effects of actions within a system.	Reliability for each construct ranged from .52 for systems thinking to .75 for persisting.	None reported.	Initial review identified competencies through numerous resources, including SCANS. Instrument field tested in 13 different colleges. Teachers returned their assessment of students (using same instrument), which was matched with individual student's perceptions. Significant correlations with student and teacher perceptions on all scales, except adapting to change.	None reported.	**Strengths** Instrument can be a diagnostic measure to identify areas in which students could benefit most from educational interventions (Maduschke and Grummon 1996). Results can help inform the development of individual plans to address perceived weaknesses and build on strengths. It can be used by both faculty and counselors to help students become better prepared for the workplace. Instrument has been field tested and used with college students in community colleges, technical colleges, and state universities. **Limitations** None

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
INSIGHT Inventory 32 items **Author** Patrick Handley **Publisher** WISEWORK (816) 587-3881 http://www.wisework.com and http://www.wisework.com/insight.htm **Dates** 1988-1995 **Testing Time** 15 to 50 minutes **Cost** $6.50 per inventory $250.00 for comprehensive trainer's guide, skill-building activities, overhead transparencies, and technical manual	Eight profile scores for each area work style and personal style Getting one's own way Responding to people Pacing activity Dealing with details	Consists of two lists of 32 adjectives each, presented side by side. Individuals complete instrument by indicating on a four-point scale extent to which adjective describes the way they are at work (or for students the way they are at college) and in the second list of same adjectives indicate the way they are at home. Direct versus indirect Outgoing versus reserved Steady versus urgent Precise versus unstructured	Coefficient alpha ranged from .71 to .85 with median of .77 (Urbina 1998). Stability over 6 weeks by means of test-retest ranged from .54 to .82 with median of .755 (Urbina 1998).	Began with initial set of adjectives from Allport and Odbert's (1936) list. After some initial modifica-tions, list was adminis-tered to sample of adults and then factor analyses conducted. Based on these analyses, list was altered again and given once more to samples of adults, high school, and college students. Resulting factor loadings were used to create the four scales, each consisting of eight	Author cites the work of Kurt Lewin (1936), Gordon Allport and Odbert (1936), and Raymond Cattell (1971) as source material for development of this inventory.	Evidence of validity for insight is limited to comparisons between its scores and three other well-known self-report inventories, self-directed search, the Sixteen Personality Factor Questionnaire, and Myers-Briggs Type Indicator.	**Strengths** This inventory could be used to stimulate discussion about interpersonal relationships especially in the context of team building. Simplicity and directness of measure likely to increase its usefulness and appeal (Urbina 1998). **Limitations** This instrument is a learning tool rather than a formal test.

Table A-3.　Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
INSIGHT Inventory (continued)				adjectives represent-ting four factors. Norms presented separately for female, male, and combined gender groups of adults and students drawn from a variety of settings.			

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Coaching Process Questionnaire 40 items **Author** McBer and Company **Publisher** McBer and Company 116 Huntington Ave. Boston, MA 02116 (800) 729-8074 (617) 425-4588 http://trgmcber.haygroup.com **Date** 1992 **Testing Time** Administration time not reported. **Cost** $65.00 per package of 10 participant questionnaires and interpretive notes $25.00 per package of 10 feedback questionnaires	Total score and four subscores: (1) Diagnostic skills (2) Coaching qualities (3) Coaching techniques (4) Coaching model Same framework, items, and scale for both the participant/ manager version and the employee version	Assess manager's ability to prepare for coaching session. Personal attitudes and beliefs supportive to the coaching process. Assess manager's ability to communicate in a meaningful way. Ability to structure coaching session so that developmental opportunities will be understood and pursued. Each individual responds to statement by answering five-point Likert-type scale.	Range of reliability estimates for participant version ware .68 to .78 and for employee version ranged from .81 to .87. Type of reliability calculated was not specified.	None reported.	Some content validity–test developers present model of coaching process and develop items rating each of the four elements of model. No data about content, criterion, and construct validity provided.	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. It could be useful to stimulate a discussion of the coaching process among employees and their managers. In higher education, it could be used in courses with simulations where students adopt different roles and discuss the coaching process. Instrument is easy to administer and score. **Limitations** Instrument has little psychometric evidence.

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
The Masterful Coaching Feedback Tool 35 items **Author** Robert Hargrove **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 2001 **Testing Time** 15 minutes **Cost** $5.00 Masterful Coaching, Self-Assessment $5.00 Masterful Coaching, Observer Assessment Instrument $15.00 Participant's Workbook $112.00 Facilitator's Package	Total score and five subscores: (1) Plans goals collaboratively (2) Provides feedback and learning (3) Invests in relationships (4) Forwards the action (5) Develops a coaching mission	Engaging in joint inquiry with individuals and groups Encourage attitude of learning Relating to others across the organization to help them be successful Generates successful action for people Develops a clear sense of what they want to accomplish with others	For participants—coefficient alphas ranged from .68 to .82. Coefficient alpha was high at .90 (Hargrove 2001). For observer ratings—reliability estimates ranged from .83 to .89 (Hargrove 2001).	None reported.	An observer measure was included to test the validity. .	None reported.	**Strengths** It could be useful to stimulate a discussion of the coaching process and key interpersonal skills that are important to be effective. In higher education, it could be used in courses with simulations where students adopt different roles and discuss the coaching process. Instrument is easy to administer and score. **Limitations** This instrument is a learning tool rather than a formal test.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Coaching Skills Inventory (CSI) 50 items **Author** Dennis C. Kinlaw **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 1999 **Testing Time** 15 minutes **Cost** $7.00 CSI, Observer $12.00CSI, Self $340.00 Facilitator's Package	Total score and five subscores: (1) Contact and core communication skills (2) Counseling skills (3) Mentoring skills (4) Tutoring skills (5) Confronting and challenging skills	Setting clear expectations, establishing objectives, probing, reflecting Changes in point of view, commitment to self-sufficiency Development of political savvy, greater proactivity in managing one's career Increased knowledge and skill, increased confidence Clarification of performance expectations, identification of performance shortfalls	Coefficients of .81 or higher were obtained for test and retest ratings on all Coaching Skills Inventory items (Kinlaw 1999).	None reported.	Builds on research studies conducted by Kinlaw (1989; 1990).	None reported.	**Strengths** It could be useful to stimulate a discussion of the coaching process and key interpersonal skills that are important to be effective. In higher education, it could be used in courses with simulations where students adopt different roles and discuss the coaching process. Instrument is easy to administer and score. **Limitations** This instrument is a learning tool rather than a formal test.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Communicative Adaptability Scale (CAS) 30 items **Author** R.L. Duran **Original Instrument Located In:** R.L. Duran. (1983). Communicative Adaptability: A Measure of Social Communicative Competence. *Communication Quarterly*, 31, 320-326. Duran, R.L. (1992). Communicative adaptability: A review of conceptualization and measurement. *Communication Quarterly*, 40, 253-268. **Testing Time** Less than 6 minutes	Total score and six subscores: (1) Social composure (2) Social experience (3) Social confirmation (4) Appropriate disclosure (5) Articulation (6) Wit	Feeling relaxed in social situations. Enjoying and participating socially. Maintaining the other's social image. Adapting one's disclosures appropriately Using appropriate syntax and grammar. Using humor to diffuse social tension.	In 10 samples, average alpha for subareas were social experience, .80; social confirmation, .84; social composure, .82; appropriate disclosure, .76; articulation, .80; and wit, .74 (Duran 1992). Overall scale alpha of .81 (Cupach and Spitzberg 1983) and overall alpha of .79 (Duran and Zakahi 1984).	None reported.	Research pertaining to construct validity found a significant difference between high and low cognitively complex persons on the social experience and wit dimensions (Duran and Kelly 1985). They also discovered that women have higher scores on social experience and appropriate disclosure. Validity studies summarized in Duran (1992).	CAS is related to Interaction Involvement Scale. Responsiveness was related to social confirmation and appropriate disclosure. Perceptiveness was related to social composure and social experience. Attentiveness was related with to social experience (Duran and Kelly 1988).	**Strengths** Scale used with primarily college students only. Instrument can be completed quickly. It has strong reliability evidence. **Limitations** Scale has been used primarily as a self-report instrument and may not indicate students' actual competencies.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Interaction Involvement Scale 18 items **Author** D.J. Cegala **Original Instrument Located In:** D.J. Cegala. (1981). Interaction Involvement: A Cognitive Dimension of Communication Competence. *Communication Education*, 30, 109-121. **Testing Time** Less than 6 minutes	Total score and three subscores: (1) Perceptiveness (2) Attentiveness (3) Responsive-ness	Interaction involvement is the degree to which people are engaged, cognitively and behaviorally, in their conversations with others. Being aware of message meanings. Hearing and observing. Person's certainty about how to respond to others during a conversation.	Test-retest reliability was .81 after 6 week delay (Cegala et al. 1982). Alphas for Responsiveness subscale ranged from .69 (Duran and Kelly 1988) to .86 (Cegala 1981.) Alphas for the perceptiveness subscale ranged from .63 (Rubin and Graham 1988) to .88 (Cegala 1981). Alphas for attentiveness subscale ranged from .64 (Duran and Kelly 1988) to .87 (Cegala 1981). Overall alphas ranged from .83 (Chen 1989) to .90 (Cegala et al. 1982).	None reported.	Factor analysis confirmed the original three dimensions (Cegala et al., 1982).	All three dimensions of Interaction Involvement Scale scores positively correlated with CCS (Cegala et al. 1982).	**Strengths** Scale used with college students only. Instrument can be completed quickly. Inclusion of both cognitive and behavioral items. It can help college students better understand their interaction involvement. Instrument has strong reliability evidence. **Limitations** Scale has been used primarily as a self-report instrument and may not indicate students' actual competencies.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Learning Tactics Inventory 32 items **Author** Maxine Dalton **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 1999 **Testing Time** 15 minutes **Cost** $15.00 Learning Tactics Inventory: Participant Workbook $30.00 Facilitator's Guide	Total score and four subscores: (1) Action (2) Thinking (3) Feeling (3) Accessing others	Briefly sketching what needs to be done and doing it Rehearsing actions before going into situation Trusting feelings; Impact feelings have on actions Talking with someone who has same experience; Emulating behavior of another person	All scales met level of internal consistency of .70 using Chronbach's alpha	None reported.	Action items drawn from work done by Revans (1980) and Bandura (1977). Thinking items drawn from work by Bandura (1977) and Mezirow (1990). Feeling items drawn from work done by Kolb (1984) and Horney (1970). Accessing others items drawn from work done by Bandura (1997).	Scores on four subscales was correlated with Schraw and Dennison (1994) scale of meta-cognitive ability (Dalton, 1999)	**Strengths** Instrument could be used with college students. Instrument can be completed quickly. It is intended to help individuals increase their self-awareness for personal development. Instrument has strong reliability evidence and information about validity. **Limitations** Inventory has been used primarily as a self-report instrument and may not indicate students' actual competencies.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Individualized Trust Scale (ITS) 15 items **Author** D.J. Cegala **Original Instrument Located In:** L.R. Wheeless and J. Grotz. (1977). The Measurement of Trust and its Relationship to Self-Disclosure. *Human Communication Research*, 3, 250-257. **Testing Time** About 1 minute	Total score	Individualized trust is "process of holding certain relevant, favorable perceptions of another person which engender certain types of dependent behaviors in a risky situation where the expected outcomes that are dependent upon that other person is not known with certainty" (Wheeless and Grotz 1977, p. 251). ITS focuses on specific person rather than trust in other people in general.	Split-half reliability of .92 for ITS (Wheeless and Grotz 1977). In research studies, an alpha of .95 reported by Snavely (1981) and .72 reported by Buller, Strzyzewski, and Comstock (1991).	None reported.	Wheeless (1977) found strong relationship between self-disclosure, individualized trust, and interpersonal solidarity.	None reported.	**Strengths** Instrument can be completed quickly. It could be used when students are working in groups to help them learn about level of trust for members in their own group. **Limitations** Scale has been used primarily as a self-report instrument and may not indicate students' actual competencies.

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Interpersonal Trust Surveys 60 items **Author** Guy L. DeFuria **Publisher** Jossey-Bass/Pfeiffer 350 Sansome Street, 5th Floor San Francisco, CA 94104-1342 (800) 274-4434 http://www.pfeiffer.com **Date** 1996 **Testing Time** 20 to 30 minutes **Cost** $9.00 Interpersonal Trust Survey, Self-Assessment $9.00 Interpersonal Trust, Observer Scoring $40.00 Facilitator Book and Instruments	Total score Ten subscales Five subscales measure the respondent's propensity to engage in trust-enhancing behaviors Five subscales measure the respondent's expectations that others will engage in trust-enhancing behaviors.	Inidividuals' behaviors of—sharing relevant information; reducing controls; allowing for mutual influence; clarifying mutual expectations; meeting others' expectations Others behaviors of—sharing relevant information; reducing controls; allowing mutual influence; clarifying mutual expectations; meeting my expectations	Coefficient alpha=.981	None reported.	Content analysis was performed using groups of subjects in a group interview process to determine that each item was uniformly interpreted.	None reported.	**Strengths** Instrument can be completed quickly. It could be used when students are working in groups to help them learn about level of trust for members in their own group. Comparisons can be made between individual's level of trust and his or her own perceptions of others' levels of trust along same dimensions. **Limitations** Instrument lacks detailed evidence of validity.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Conversational Appropriateness and Effectiveness Scale 40 items **Authors** B.H. Spitzberg and L.A. Phelps **Original Instrument Located In:** D.J. Canary and B.H. Spitzberg. (1987). Appropriateness and Effectiveness Perceptions of Conflict Strategies. *Human Communication Research*, 14, 93-118. B.H. Spitzberg and L.A. Phelps. (1982, November). Conversational Appropriateness and Effectiveness: Validation of a Criterion Measure of Relational Competence. Paper presented at the meeting of the Speech Communication Association. **Testing Time** Less than 3 minutes	Total Subscores: (1) Effectiveness (2) Specific Appropriateness (3) General Appropriateness	Goal accomplishment. Specific aspects of the conversation were appropriate. Global suitability.	Coefficient alphas for the effectiveness scale ranged from .87 (Canary and Spitzberg 1989) to .93 (Canary and Spitzberg 1987). Alphas for specific appropriateness scale ranged from .74 (Canary and Spitzberg 1990) to .85 (Canary and Spitzberg 1989) and for general appropriateness from .80 (Canary and Spitzberg 1989) to .92 (Canary and Spitzberg 1990).	None reported.	Lacks full information on construct validity.	Canary and Spitzberg (1989) found that specific appropriateness was correlated with avoidant and distributive messages. General appropriateness was correlated with integrative tactics and with trust, mutuality, and intimacy; and effectiveness correlated with trust, intimacy, and satisfaction.	**Strengths** Instrument can be completed quickly. Instrument can help college students better understand conversational appropriateness and effectiveness. These are two critical dimensions of interpersonal communication competence. Instrument has strong reliability evidence. **Limitations** Lacks information on construct validity.

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Interpersonal Attraction Scale 15 items **Authors** J.C. McCroskey and T.A. McCain **Original Instrument Located In:** J.C. McCroskey and T.A. McCain. (1974). The Measurement of Interpersonal Attraction. *Speech Monographs*, 41, 261-266. **Testing Time** Less than 10 minutes	Total Three subscores: (1) Social attraction (2) Physical attraction (3) Task attraction	 Liking Physical appearance Respect	Internal reliabilities for 15-item scale as follows: social attraction, .84; task attraction, .81; and physical attraction, .86 (McCroskey and McCain 1974). Other researchers report similar results (Ayres 1989; Brandt, 1979; Duran and Kelly 1988; Wheeless, Frymier, and Thompson 1992). Split-half reliability was reported as .90 for social, .87 for task, and .92 for physical attraction (McCroskey et al. 1975).	Series of factor analyses revealed 15 items, three-factor solution accounting for 49% of total variation.	Instrument builds on previous research. Little evidence of concurrent or criterion-related validity.	None reported.	**Strengths** Scale used with college students. Instrument can be completed quickly. It can help college students understand their interpersonal attraction according to three relevant dimensions. Instrument has strong reliability data. **Limitations** Has been used primarily as a self-report instrument.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Affinity-Seeking Instrument (ASI) 13 items **Authors** R.A. Bell, S.W. Tremblay, and N.L. Buerkel-Rothfuss **Original Instrument Located In:** R.A. Bell, S.W. Tremblay, and N.L. Buerkel-Rothfuss. (1987). Interpersonal Attraction as a Communication Accomplishment: Development of a Measure of Affinity-Seeking Competence. *Western Journal of Speech Communication*, 51, 1-18. **Testing Time** Less than 5 minutes	Total Two subscores: (1) Affinity-seeking competence (2) Strategic performance	Ability to say and do what is necessary to be seen as interpersonally attractive. Ability to play roles to be liked by others.	Reported alphas ranged from .85 to .89 for the ASC subscale and from .80 to .87 for the SP subscale (Bell, Trembly and Buerkel-Rothfuss 1987). Later research reported ASC alpha of .81 and SP alpha of .83 (Buerkel-Rothfuss and Bell 1987).	Original authors found stability in two-factor solution.	Significant positive relationship between ASI and affinity-seeking outcomes and negative relationships between ASI and nonrelated measures (such as public communication apprehension).	None reported.	**Strengths** Instrument can be completed quickly. Can help college students understand their social-communicative competence. **Limitations** Has been used primarily as a self-report instrument and may not indicate students' actual competencies.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Relational Communication Scale (RCS) 41 items **Authors** J.K. Burgoon and J.L. Hale **Original Instrument Located In:** J.K. Burgoon and J.L. Hale. (1987). Validation and Measurement of the Fundamental Themes of Relational Communication. *Communication Monographs*, 54, 19-41. **Testing Time** 10 minutes	Total Eight subscores: (1) Immediacy/ affection (2) Similarity/ depth (3) Receptivity/ trust (4) Composure (5) Formality (6) Dominance (7) Equality (8) Task orientation	Relational communication consists of the verbal and nonverbal themes present in communication that define an interpersonal relationship. RCS captures "the relational meanings that are embedded in all communication interchanges" (Burgoon and Hale 1987, p. 40).	Burgoon and Hale (1987) reported coefficient alphas for the eight dimensions that range from .42 to .88. Other researchers report similar results (Buller et al. 1992; Kelley and Burgoon 1991).	None reported.	Criterion-related validity studies indicated that RCS is capable of discriminating for example, "immediate from nonimmediate behaviors," "pleasant from hostile voices," "high from low reward communicators" (Burgoon and Hale 1987). Walther and Burgoon (1992) reported that computer-mediated groups mirror face-to-face groups in that both experience an increase in relational communication over time. Subsequent research studies have elaborated on these findings (e.g., Buller and Aune 1988; Buller and Burgoon 1986; Burgoon, Coker, and Coker 1986; Burgoon and Hale 1988; Burgoon, Olney and Coker 1987; Burgoon, Walther, and Baesler 1992).	None reported.	**Strengths** Instrument can be completed quickly. RCS can be used as another-report (e.g., Burgoon, Olney, and Coker 1987) and observer-report measure (Burgoon and Newton 1991). Extensive research studies contribute to construct validity evidence of RCS. **Limitations** Instrument is primarily designed as a self-report measure.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Communication Flexibility Measure **Authors** M.M. Martin and R.B. Rubin **Original Instrument Located In:** M.M. Martin and R.B. Rubin. (1994). Development of a Communication Flexibility Measure. *Southern Communication Journal*, 59, 171-178. **Testing Time** 15 to 20 minutes	Total score Students respond to 14 different scenarios by indicating on five-point scale how closely the behaviors described in the scenario resemble their own. The scenarios focus on acquaintances and friends or family in formal and informal interpersonal, group, and public settings.	Designed to assess flexibility in adapting communication behaviors to different situations.	Coefficient alpha of .70 and a split-half correlation of .71 were reported.	None reported.	Content validity-focusing on different situations and circumstances.	Instrument is related to communication adaptability, rhetorical sensitivity, and social desirability.	**Strengths** The scenarios in this measure are interesting and varied. They serve as good initiators for discussion (Morreale and Backlund 1996). Students are encouraged to think about the contextual nature of communication (Morreale and Backlund 1996). **Limitations** Usefulness of instrument will be further enhanced when a relationship is established with communication competence (Morreale and Backlund 1996).

Table A-3.　Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Interpersonal Communication Satisfaction Inventory 19 items **Author** M.L. Hecht **Original Instrument Located In:** M.L. Hecht. (1978). The Conceptualization and Measurement of Interpersonal Communication Satisfaction. *Human Communication Research*, 4:253-264. **Testing Time** 15 to 20 minutes	Total score Self-perceptions about: (1)　Public speaking (2)　Interpersonal communication (3)　Interviewing (4)　Group communication (5)　Listening (6)　self-confidence	Designed to assess college students' self-perceptions of their own communication competence.	Cronbach alpha coefficients ranged form .93 to .95 on the three context subscales—class, social/family, and work.	None reported.	Students who completed the instrument on a pre- and postcourse basis demonstrated significantly higher scores after completion of the course.	None reported.	**Strengths** This instrument is a good indicator of perceived change in communication abilities. It can be used for the communication course (Rubin, Palmgreen, and Sypher 1994). **Limitations** Because instrument is self-report, the outcomes may not correspond with the actual development of these communication abilities.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Speaker's Perceptions of Situational Causes of Anxiety (CA) 14 items **Author** M.J. Beatty **Original Instrument Located In:** M.J. Beatty. (1988). Situational and Predispositional Correlates of Public Speaking Anxiety. *Communication Education*, 37, 28-39. **Testing Time** 10 minutes	Total Seven subscores: (1) Novelty (2) Formality (3) Subordinate Status (4) Conspicuous-ness (5) Unfamiliarity (6) Dissimilarity (7) Degree of Attention	Examines the influence of situational causes of anxiety or fear. Instrument administered immediately after a public speaking performance.	Alpha reliability coefficients for each situational dimension ranged from .60 to .72 (Beatty 1988; Beatty, Balfantz, and Kuwabara 1989; Beatty and Friedland 1990).	None reported.	Research studies provide some evidence of construct validity. Beatty (1988) found subjects delivering speech from front of classroom reported significantly higher scores on the Formality dimension than those giving report from their desk to audience seated in circular arrangement. Two additional studies found that novelty scores decreased with speaking experience (Beatty 1988; Beatty, Balfantz, and Kawabara 1989).	None reported.	**Strengths** Instrument can be completed quickly. It is primarily designed as self-report measure. Instrument provides in-depth review of situational causes of anxiety. **Limitations** Factor analysis could produce "more parsimonious set of constructs" resulting in higher reliability (Beatty and Friedland 1990).

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Student Motivation Scale (SMS) 12 items **Author** D.M. Christophel **Original Instrument Located In:** D.M. Christophel. (1990). The Relationships Among Teacher Immediacy Behaviors, Student Motivation, and Learning. *Communication Education*, 39, 323-340. **Testing Time** 10 minutes	Total	College students asked to complete 12-item version twice. First-time students are asked to indicate how they **feel in general** about taking classes at the university and second-time students indicate how they **feel about this specific class** (Christophel 1990, p. 327).	Christophel (1990) reported Alpha coefficients ranged from .95 to .96.	None reported.	Builds on the research work done by Beatty and Payne (1985).	Correlation between motivation and duration of speech was .54 (Beatty, Frost, Stewart 1986).	**Strengths** Instrument can be completed quickly. It examines student learning attitudes toward either the specific course or subject matter or learning more generally (the overall curriculum). **Limitations** Instrument is primarily designed as self-report measure.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Communication Satisfaction Questionnaire (CSQ) 43 items **Authors** C.W. Downs and M. Hazen **Original Instrument Located In:** C.W. Downs and M. Hazen. (1977). A Factor Analytic Study of Communication Satisfaction. *Journal of Business Communication, 14*, 63-73. **Testing Time** 20 to 30 minutes	Total score Eight subscores: (1) Communica-tion climate (2) Relationship to superiors (3) Organizational integration (4) Media quality (5) Horizontal and informal communication (6) Organizational perspective (7) Relationship with subordinates (8) Personal feedback	Examines communication at organizational and individual levels. Includes components of upward and downward communication. Examines information that employees receive about their job and related items. Helpfulness and clarity of information of information sources. Amount of activity of information networks and accuracy of information they contain. Knowledge about external events that impact the organization. Examines communi-cation overload. Superiors' under-standing of problems faced on the job.	Test-retest (2-week interval) reliability of CSQ was .94 (Downs and Hazen 1977). Coefficient alpha reliabilities ranged from .72 to .96 as reported in research studies (Potvin 1992; Downs 1991).	Principal-compon-ents factor analysis revealed eight stable factors accounting for 61% of variance (Downs and Hazen 1977).	None reported.	CSQ has been found to be highly correlated with job satisfaction (Downs and Hazen 1977) and related to turnover (Gregson 1987).	**Strengths** Instrument can be completed quickly. It is useful for analyzing communication processes in organizations. It could be used to examine student experiences in internships or cooperative education programs. Instrument has strong psychometric evidence from numerous research studies, including more than 30 dissertations. **Limitations** Instrument has little evidence about validity.

Table A-3.   Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Negotiating Style Profile 30 items **Author** Rollin Glaser and Christine Glaser **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Dates** 1982, 1996 **Testing Time** 10 minutes **Cost** $36 Participant Booklet (5 pack) $18 Feedback Booklet (5 pack) $30 Facilitator Guide	Total score Five subscores: (1) Defeat (2) Collaborate (3) Accommodate (4) Withdraw (5) Compromise	High degree of concern for substance of negotiation and low degree of concern for the relationship. High degree of concern for both substance of negotiation and for relationship. Focus on building compatible relationship in hope that negotiation will be successful. Low degree of concern for both substance of negotiation and for the relationship. Moderate concern for both dimensions of negotiating behavior.	Reliability coefficient alpha for defeat, .71; collaborate, .67; accommodate, .64; withdraw, .60; compromise, .71	Norms given for groups of respondents from several different industries. Factor analysis revealed that most items grouped well within the original areas of defeat, withdraw, and compromise. Accommodation and collaboration were not clearly separated, so items in these two areas were revised.	Based on N2 Model of Negotiating Behavior (Glaser 1994). Gordon Shea (1983) in *Creative Negotiating* describes similar relationships.	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. Items listed in instrument are relevant for college students. **Limitations** Instrument has little evidence about validity.

Table A-3. Reviews of assessments of interpersonal skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Interpersonal Influence Inventory 40 items **Authors** Rollin Glaser and Eileen Russo **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Dates** 1990, 1993, 1995, 2000 **Testing Time** 10 to 20 minutes **Cost** $32 Participant Booklet (5 pack) $30 Facilitator Guide	Total score Four subscores: (1) Openly Aggressive Behavior (2) Assertive behavior (3) Concealed aggressive behavior (4) Passive behavior	Each subscore consists of items addressing thoughts, emotions, nonverbal behavior, and verbal behavior. Believe they have rights but others do not. Thoughts of self-confidence and belief that all individuals have rights. Getting back at another person in a devious manner. Believe they should not speak their minds. Do not want to disagree. Others have rights but they do not.	Reliability coefficient for openly aggressive, .70; assertive, .76; concealed aggressive, .79; passive, .72;	Normative data from individuals in various industries. Factor analysis confirmed the original four factors.	Based in part upon previous research, including Kelley (1979) and Alberti (1977).	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. Items listed in instrument are relevant for college students. **Limitations** Instrument has little evidence about validity.

Table A-4. Reviews of assessments of listening skills in terms of selected assessment characteristics and source

Name	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Watson-Barker Listening Test 43 items **Authors** K.W. Watson and L.L. Barker **Publisher** Innolect Inc. 31012 Executive Point Drive Tega Cay, SC 29708 (803) 396-8500 http://www.innolectinc.com/wblt.html **Dates** 1986 and 1999 **Costs** Form B $249.95 Form A $229.95 **Testing Time** 30 minutes for Form A and 60 minutes for Form B	Total score Each of two different forms (A and B) includes 50 items divided into five parts. Instructions, listening material, and test questions for Form A and Form B are all on one audio cassette tape.	Part One: Evaluate test content and implications of messages. Part Two: Understand meaning in conversations. Part Three: Understand and remember information in lectures. Part Four: Evaluate emotional meanings in messages. Part Five: Follow and remember instructions and directions.	Reliability of individual parts and total score not reported.	Correlation between Form A and Form B is .42. Norms are available for both college students and professionals, including managers, supervisors, and administrators.	None reported.	Receiver apprehension test (RAT) scores are related to total listening ability and long-term memory measured by Watson-Barker Listening Test (Roberts 1986).	**Strengths** Designed for use with college students and adults in business and professional settings. It is easy to administer instrument. It focuses on communication skill that is often overlooked. Directions are clear. Normative responses are available for comparison. **Limitations** Evidence of validity and stronger reliability are needed. Test scores may be affected by reading ability.

Table A-4.   Reviews of assessments of listening skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Kentucky Comprehensive Listening Test 78 items **Authors** R.N. Bostrom and E.S. Waldhart **Publisher** Department of Communication, University of Kentucky, Lexington, KY 40506-0042 http://www.uky.edu/~bostrom **Date** 1985 **Costs** $50.00 Sample Packet Form A $150.00 Complete Kit Form A and B **Testing Time** 60 to 90 minutes	Total score Five subscores: Instrument includes audiotape and question and response sheet.	(1) Short-term listening (2) Short-term listening with rehearsal (3) Interpretative listening (4) Informative listening (5) Ability to overcome distraction while listening	Cronbach alpha coefficients for global scores on Forms A and B were .78 to .82. Subscales ranged from .46 to .76. Reliability between Forms A and B was .72. Interpretative subscale (part three) for alternative forms was .36. Test-retest coefficients ranged from .78 to .87 for various subscales.	Forms adminis- tered to more than 20,000 college students and adults.	Communication faculty from University of Kentucky helped to validate part of the test.	Kentucky Comprehensive Listening Test correlates with the Watson-Baker Listening Test	**Strengths** Students base their answers on listening to an audiotape. This allows for control over the testing environment. (Morreale and Backlund, 1996). **Limitations** The relationship between short-term listening and short-term listening/ rehearsal to effective interpersonal communication and retention of information such as that presented in lectures and classroom discussions is not clear (Morreale and Backlund, 1996).

Table A-4.   Reviews of assessments of listening skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
Learning to Listen 30 items **Author** Laurie Ribble Libove **Publisher** HRDQ 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Date** 1996 **Testing Time** 10 minutes **Cost** Starter Kit (5 participant guides and facilitator guide) = $90.00 Preview pack (participant guide and facilitator guide)=$59.00	(1) Staying focused (2) Capturing the message (3) Helping the speaker	Consciously listening and keeping one's full attention centered on the speaker. Building a complete and accurate understanding of the speaker's message. Focusing on listener's outward behaviors such as offering verbal encouragement and support or offering nonverbal encouragement and support.	No evidence reported.	None reported.	No evidence reported.	None reported.	**Strengths** Instrument is designed to help individuals identify the extent to which they practice behaviors that are associated with effective listening. **Limitations** Instrument lacks validity and reliability evidence.

Table A-4.   Reviews of assessments of listening skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Correlation with measures	Strengths and Limitations
WORKKEYS— Listening Six complex audio messages reflect different workplace settings and given by males and females of differing ages and various accents. **Publisher** ACT—WORKKEYS 2201 North Dodge Street P.O. Box 168 Iowa City, IA 52243-0168 (800) 967-5539 http://www.act.org/workkeys/contacts.html **Testing Time** 40 minutes **Cost** $12.50 per student (includes scoring of written responses and reports for individual students)	Level 5 Level 4 Level 3 Level 2 Level 1 Level 0	All primary and supportive details are present and correct, including all relation-ships among details. Response is correct in that all primary details and relation-ships among details are given and correct; may be missing supportive details or have incorrect supportive details that do not interfere with accurate communication. Response is substan-tially correct. All primary details are correct and relation-ships among them are correct. May be missing a few primary details. Some pertinent details; may have incorrect primary details, but sketch of situation is correct. Minimal pertinent information. No meaningful information, or totally inaccurate information; message is not in English	Consult ACT.	Consult ACT.	All test items submitted to external consultants for content and fairness reviews. Consult ACT.	Consult ACT.	**Strengths** Instrument focuses on skills deemed important by employers. Individual student report can be generated and used as an attachment to a resume. It provides students with recommendations about how to reach higher level of listening. **Limitations** Difficult to identify based upon limited information provided by ACT.

TABLE B

ASSESSMENT REVIEWS FOR LEADERSHIP

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Leadership Competency Inventory 46 items **Author** Stephen P. Kelner **Publisher** McBer & Company 116 Huntington Avenue Boston, MA 02116 (800) 729-8074 (617) 425-4588 http://trgmcber.haygroup.com **Date** 1993 **Testing Time** Administration time not reported. **Cost** $65 per package of 10 participant questionnaires $25 per package of 10 employee version	Total Information seeking Conceptual thinking Strategic orientation Service orientation Two identical versions: one can be completed from the perspective of the actual employee's own self-assessment and second can be completed by others' perceptions of the particular individual.	All subscores Likert scale, ranging from "0" (behavior is absent) to "4" (behavior occurring extremely frequently) Asking questions or personally investigating matters. Recognizing patterns or applying complex concepts. Aligning current actions with strategic goals or understanding external impact. Making one's self available or maintain clear communication.	No reliability evidence reported.	Not reported.	No validity evidence reported.	None reported.	**Limitations** Author provides no rationale or theoretical framework for behaviors that are used for each of the four areas assessed. No analysis on how items were developed, pretested, or selected from research on leadership. All scales composed of hierarchical levels that range from basic to more advanced. No reasons given for why some behaviors are considered to be more advanced.

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Styles of Leadership Survey (SLS) 12 items **Authors** Jay Hall, Jerry B. Harvey, and Martha S. Williams **Publisher** Teleometrics International 1755 Woodstead Court The Woodlands, TX 77380-0964 http://www.teleometrics.com **Dates** 1968-1986 **Testing Time** Untimed. **Cost** $8.95 per survey.	Total SLS is to be self-administered and self-scored.	Four subscores Twelve items with three each assigned to one of four categories: (1) Philosophy of leadership (2) Planning and goal setting (3) Implementation (4) Performance and evaluation	Manual gives median coefficient of stability (.70) with no other forms of reliability.	Four main themes of scale include profile summaries based on normalized, standardized results from about 3,000 leaders representing education, civic, business, industry, government, and service organizations.	Basis for scoring and interpretation is the managerial grid of Blake and Mouton (1964).	SLS styles correlate with personality traits in ways consistent with grid theory.	**Strengths** Instrument presents a learning opportunity for respondent to gain stronger understanding and insights into his or her leadership style. It can be useful to stimulate discussion about leadership styles. Instrument addresses topics associated with successful leadership and with decisionmaking and problem solving. It can be administered individually or in a group setting. **Limitations** Little psychometric evidence.

Table B-1. Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Management Inventory on Leadership, Motivation, and Decisionmaking 59 items **Author** Donald Kirkpatrick **Publisher** Kirkpatrick Management Training Inventories 3137 Citadel Indianapolis, IN 46268 (800) 834-02965, ext. 4937 http://buros.unl.edu/buros/jsp/clists.jsp?cateid=18&catename=Vocations **Date** 1991 **Testing Time** 15-20 minutes **Cost** $40.00 Package of 20 inventories and 20 answer booklets $10.00 Additional sample sets (inventory, answer booklet and instructor manual) $10.00 Cassette describing practical uses of inventory.	Survey designed to assess beliefs and attitudes about several aspects of leadership. Two sections— Initial section: 55 items giving total score of correct answers (agree or disagree items). Correct answers represent author's opinion (but are often backed by research findings). Second section: four items—respondents indicate frequency (described in terms of percentages) with which they have used strategy in the past and the frequency that each one should be used in the future. No formal scoring system.	Twenty items about leadership; 25 items about motivation; and 10 items about decisionmaking Describes four decisionmaking approaches	No reliability data reported.	Instructor's manual contains norms based on small samples of supervisors, human resource consultants, and human resource professionals in organizations. On average, members of these groups agreed with the author of the inventory on 83% to 89% of the items. Higher scores may indicate more well-informed perspectives about leadership, motivation, and decisionmaking, but there is no evidence to support the validity of this interpretation (Murphy 1995).	Some correct answers have a rationale that is based on research studies for items in the first section.	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. It could be used to stimulate group discussions about leadership. Strong group participation could be generated because many of the items reflect widely held but incorrect beliefs about leadership and motivation. **Limitations** Instrument has little reliability and validity evidence.

Table B-1. Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Leatherman Leadership Questionnaire (LLQ) (Revised) 339 multiple-choice questions **Author** Richard W. Leatherman **Publisher** International Training Consultants Inc. P.O. Box 35613 Richmond, VA 23235-0613 (800) 998-8764 http://assess.trainingitc.com **Dates** 1987-1992 **Testing Time** 300 to 325 minutes for battery and 150 to 165 minutes per part. **Cost** $1,500 LLQ Administrators Kit including Standard Tests for 10 participants. $95.00 LLQ Standard Test with development manual. Testing materials without charge for qualified not-for-profit college or university research.	Total score	Twenty seven subscores ranging from such areas as assigning work, coaching employees, oral communication, managing change, dealing with conflict, delegating, negotiating, conducting meetings, persuading/influencing, making presentations, problem solving, team building, and managing time. Knowledge-based paper and pencil alternative to assessment center that could be used to select leaders and define leadership needs.	KR-20 reliabilities ranged from .9054 to .9905, with an average of .9657 for the entire instrument. Deleting a few items with low item-test and item-task correlations raised the composite reliability to .9706.	Research report (99 pages) shows strong procedures used in instrument development, including detailed analysis of more than 400 related existing measurement reviews of industry studies, relevant articles, and needs assessment instruments. Provides table mapping each task to related literature. Norms available for large number of industries.	Each task of LLQ has accompanying evidence that it is content valid. High concurrent validity when LLQ is compared with overall ratings from assessment center scores and rankings and other leadership or supervision measures.	None reported.	**Strengths** Very thorough documentation is included for this instrument. Instrument has strong postadministration support. It is useful as a needs assessment or training tool. It provides strong measure of leadership knowledge. Scoring is done by publisher and results are fed back individually and in aggregate form. Total score reflects weights assigned by publisher. When evaluating total score, these weights may not reflect what is most important in a particular organization.

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Leader Behavior Analysis II (LBAII, Self-Assessment and LBAII Other) 20 items **Authors** Drea Zigarmi, Douglas Forsyth, Kenneth Blanchard, and Ronald Hambleton **Publisher** The Ken Blanchard Companies 125 State Place Escondido, CA 92029 (800) 728-6000 http://www.kenblanchard companies.com **Date** 1991 **Testing Time** 15 to 20 minutes **Cost** $8.95 Leader Behavior Analysis II, Self-Assessments and Scoring Pack $4.95 LBAII Other Form (Assessments) $42.00 Profile Package	Six scores with two primary scores Style 1 (S1) Style 2 (S2) Style 3 (S3) Style 4 (S4) Flexibility score Effectiveness score	Two types of leadership behavior, Directive and Supportive are dichotomized (high and low) to produce four LBAII styles. High direction/Low support Style 2 (S2) = High direction/High support Style 3 (S3) = Low direction/High support Style 4 (S4) = Low direction/Low support How often respondent uses a different style to solve situations. How effective respondent is in certain situations. Each item is a description of a situation requiring respondent to select one of four behavioral responses.	Internal consistency reliability of LBAII Other's instrument: alphas of S1 and S4 were typically in .80s. S2 and S3 typically in the .70s (Bernardin and Cooke, 1995). One study reported alphas for self-scale. They were .51, .45, .56, and .42 for S1, S2, S3, and S4 respectively (Bernardin and Cooke, 1995).	Designed to measure perceived leadership style from the perspective of either the leader or subordinates to the leader. LBAII self-assesses self-perceived leadership style and LBAII other, assesses perceptions of the leader's leadership style Norms provided by the distributors.	Based on situational leadership theory and application of principles expected to lead to increased satisfaction and organiza-tional effectiveness (Hersey and Blanchard 1982; Blanchard, Zigarmi, and Zigarmi 1985).	Correlated LBAII to consideration and initiating structure from the Leader Behavior Description Questionnaire. Correlation coefficients, though small (e.g., .07, .12) were all significant at .05 or better level (Bernardin and Cooke 1995). Authors present correlations demonstrating that LBAII is statistically and conceptually related to Multi-Level Management Survey by Wilson (1981).	**Strengths** Instrument may be useful to stimulate discussions about leadership styles. **Limitations** It should not be used to make decisions about respondents. Few studies compare self and other scores. Scoring done by respondent is somewhat complicated (Bernardin and Cooke, 1995).

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Management Style Inventory 15 items **Authors** J. Robert Hanson and Harvey F. Silver **Publisher** J. Robert Hanson, Sons, and Associates Pownal Offices and Conference Center 238 Hallowell Road Pownal, ME 04069 (207) 688-2265 http://www.thoughtful education.com **Date** 1981 **Testing Time** Administration time not reported. **Cost** $10.00 per inventory with minimum order of five instruments.	Total score for each of four styles of decisionmaking preferences	Four styles include: (1) sensing feeling manager (2) sensing thinking manager (3) intuitive thinking manager (4) intuitive feeling manager	No reliability evidence reported.	None reported.	Inventory based on Jung's Type Theory	None reported.	**Strengths** This instrument is a learning tool rather than a formal test. It could be used to stimulate group discussions about styles of decision making. Individual can gain better understanding about his or her own management style. It is easy to administer, easy to score, and easy to evaluate. **Limitations** Instrument lacks reliability and validity evidence.

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Student Leadership Practices Inventory 30 items **Authors** James M. Kouzes and Barry Z. Posner **Publisher** Pfeiffer: A Wiley Imprint 605 Third Avenue New York, NY 10158-0012 (212) 850-6000 www.pfeiffer.com **Date** 1998 **Testing Time** Administration time not reported. **Cost** $2.50 self-instrument $2.00 observer instrument $20.00 facilitator's guide $24.95 CD-ROM scoring software	Total Challenging the process Inspiring a shared vision Enabling others to act Modeling the way Encouraging the heart Two forms—self and observer have identical items	Searching for opportunities and experimenting and taking risks. Envisioning the future and enlisting the support of others. Fostering collaboration and strengthening others. Setting an example and planning small wins. Recognizing contributions and celebrating accomplishments.	Internal reliability scores of .66 for challenging, .79 for inspiring, .70 for enabling, .68 for modeling, and .80 for encouraging. Test-retest reliability .91.	Factor analyses reveal five factors.	Instrument based on leadership model developed by authors (Kouzes and Posner 1995).	None reported.	**Strengths** Instrument was pilot tested with college students. Instrument can be used to help individuals identify areas for improving leadership practices. Scoring software CD-ROM available and can handle up to 80 self-respondents and nearly 800 observers. Instrument provides reliability evidence. Instrument based on model developed by the authors. **Limitations** It has limited evidence about validity.

Table B-1.   Assessment reviews for leadership-individual in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Campbell Leadership Index (CLI) 30 items **Author** David Campbell **Publisher** NCS Pearson 5605 Green Circle Drive Minnetonka, MN 55343 (800) 627-7271 http://www.ncspearson.com **Date** 1998 **Testing Time** No time limit. Approximately 20 minutes **Cost** $59.00 CLI Manual $40.00 User's Guide	Overall index representing global measure of leadership effectiveness and subscores on five orientations: (1) Leadership (2) Energy (3) Affability (4) Dependability (5) Resilience	Scales for ambitious, daring, dynamic, enterprising, experienced, farsighted, original, persuasive. Recognition of physical demands required of leaders. Scales for affectionate, considerate, empowering, entertaining, friendly. Scales for credible, organized, productive, thrifty. Scales for calm, flexible, optimistic, and trusting.	Leadership coefficient alpha = .90; energy =.75; affability = .90; Dependability = .84; and resilience = .85. Test-retest reliability for overall index is .87 for self-ratings and .85 for observer ratings.	None reported.	Numerous research studies supporting validity (see user's guide).	None reported.	**Strengths** Although designed for professionals in the workplace, the items on survey could be used with college students. Instrument can be used to help individuals identify areas for improving leadership. It has reliability evidence. **Limitations** Because instrument was designed for the workplace, there is no normative data for college students.

Table B-2.   Assessment reviews for leadership-team in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Team Leadership Practices Inventory (LPI) 30 items **Author** James M. Kouzes and Barry Z. Posner **Publisher** Pfeiffer: A Wiley Imprint 605 Third Avenue New York, NY 10158-0012 (212) 850-6000 www.pfeiffer.com **Date** 1997 **Testing Time** Administration time not reported. **Cost** $35.00 Facilitator's Guide Package	Total Challenging the process Inspiring a shared vision Enabling others to act Modeling the way Encouraging the heart Two forms—self and others have identical items	Searching for opportunities and experimenting and taking risks. Envisioning the future and enlisting the support of others. Fostering collaboration and strengthening others. Setting an example and planning small wins. Recognizing contributions and celebrating accomplishments.	Reliability coefficients for—challenging the process, .86; inspiring a shared vision, .89; enabling others to act, .85; modeling the way, .83; encouraging the heart, .91. Test-retest reliability at levels greater than .90 (Kouzer and Posner 1995).	Factor analyses reveal five factors.	Instrument based on leadership model developed by authors (Kouzes and Posner 1995).	None reported.	**Strengths** The majority of items are relevant for college students who work in teams for a period of time, such as the entire semester. Instrument can be used to help team members identify areas for improving leadership practices within the team. Additionally, it identifies the leadership practices currently used within the team. It can help teams develop plans for becoming more effective. Instrument has strong reliability evidence. **Limitations** It has limited evidence about validity.

Table B-2.   Assessment reviews for leadership-team in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Method and design	Validity	Relationship with other measures	Strengths and Limitations
Team Leader Survey 36 items **Author** Ann Burress **Publisher** 2002 Renaissance Boulevard #100 King of Prussia, PA 19406-2756 (800) 633-4533 http://www.hrdq.com **Date** 1994 **Testing Time** 10 minutes **Cost** $30.00 for Facilitator Guide $18.00 Feedback Booklet (5 pack) $32.00 Participant Booklet (5 pack)	Total Six subscores (1) Change Management (2) Interpersonal Skills (3) Influence (4) Administrative (5) Thinking (6) Communication	Willingness to take risks, adjusting to the unexpected, presenting alternative ideas in changing conditions. Fosters team interaction, develops solutions that capitalize on differences among team members, works through conflicts. Performing actions that enable and empower others; provide suggestions, resources, and encouragement. Works with team to establish clear direction, helps team meet schedule requirements. Attends to nonverbal cues, identifies problems teams are avoiding. Encourages open communication, listens to team members, gives feedback to team members, communicates thoughts clearly.	Reliability coefficients for change management, .89; interpersonal, .87; influence, .85; Administrative, .83; thinking, .81; communicating, .81.	None reported	Several phases of test development with different groups of individuals. Survey based on current, relevant research such as Barry (1992), Manz and Sims (1984), Manz and Sims (1980), Likert and Araki (1986), Jessup (1990).	None reported.	**Strengths** Although tested with individuals in diverse industries, items are relevant for student team leaders. Team leaders can gain feedback from members of their team and compare scores with their own assessment. Instrument can provide useful data for leadership training and development, identification of strong skills, and those areas needing improvement. **Limitations** It has little evidence of validity.

TABLE C

ASSESSMENT INSTRUMENTS FOR INFORMATION LITERACY

TABLE C.
ASSESSMENT INSTRUMENTS FOR INFORMATION LITERACY

Before presenting the assessment review templates, it is appropriate to provide a brief caveat to the assessment professional whose job is to select and validate instruments for the assessment of information literacy. First, it is important to keep in mind that information literacy (IL) is **not** a set of skills related solely to computer literacy. While the ability to use a word processor, a spreadsheet, or an Internet browser is certainly useful in the application of information literacy skills, tests of computer skills such as the TekExam or Smart Force do not capture the breadth or depth of IL. For that reason, we include only one of these sets of tutorials/tests in the templates (see the example from SUNY Brockport).

The construct of information literacy can only be captured if it is treated as a broad set of skills for the information-intensive society that most college campuses have become. It includes skills across all psychological domains (cognitive, affective, psychomotor, and conative), and the breadth of possible outcomes touches every curriculum and discipline. The assessment professional should refer to the broad set of outcomes presented in the previous section as well as other materials on the subject from the Association of College and Research Libraries (ACRL) before designing or selecting measures of information literacy.

A class of online tutorials with embedded testing is emerging that is designed to teach and assess the breadth of IL outcomes as defined by the ACRL. This is a new approach that may have implications beyond the construct of information literacy and is one with which assessment professional should become involved at the implementation stage. Campuses must be prepared to perform psychometric tests and norm test data on the population of interest.

Assessment professionals should be aware that there is an array of interfaces and coverage of the construct IL reflected in the measures or tutorials presented in table G. The Texas Information Literacy Tutorial (TILT) is an online tutorial that is immediately available. Additionally, by spring 2004, the James Madison University (JMU) Computer-Based Testing Clusters will be completed and the Information Literacy Test will be available. Dr. Steve Wise at JMU reports that the 80-item test is based on the ACRL standards and will be a robust measure of information literacy skills. These clusters are delivered online and they also can be adapted to cover any construct one wishes. They are based on the most current theories in Item Response Theory and Computer Adaptive Testing, and they will be flexible across platforms. The clusters are just entering the testing phase at five universities, but, once that phase is complete, they will be available at no charge to any institution wishing to procure them. For more information, contact T. Dary Erwin or his staff at the Center for Assessment and Research Studies (CARS), James Madison University MSC 6806, Harrisonburg, VA 22807.

An additional assessment initiative in information literacy is being undertaken at Kent State University. The project, Standardized Assessment of Information Literacy Skills (SAILS) has as its main goal the development of a web-based tool "that is standardized, easily administered, and is valid and reliable" (National Forum on Information Literacy 2003). The instrument is based on the ACRL Information Literacy Competency Standards for Higher Education and has received endorsement from the Statistics and Measurement Committee of the Association of Research Libraries. The tool will be designed so that it can be administered by any institution that has a goal of generating data for either internal assessment or external benchmarking. Grant funds support continued development of the instrument including testing with other institutions. The SAILS development team will be soliciting participants during spring 2003 through spring 2005. Participation will include planning meetings, training workshops, and debriefings of usage as well as marketing and public relations for the project. For further information about SAILS, see http://sails.lms.kent.edu (or O'Connor, Radcliff, and Gedeon, 2002).

Table C-1 provides a detailed description of assessment instruments for information literacy including ordering information and psychometric properties of each instrument.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Computer Self-Efficacy 32-item Likert Scale **Author** C.A. Murphy, et al. **ERIC Document** ED 307 317	Self-percepts of beginning level, advanced level, and mainframe computer skills.	Self-percepts regarding computer use to include some skills related to information literacy, e.g., ability to organize information on a computer.	Factor analysis revealed a three-factor structure Cronbach's alpha reliability are .97 (Factor I), .96 (Factor II), and .92 (Factor III)	Criterion-related validity using factors as dependent variables in a regression analysis with efficacy information yielded significant betas for perception of task difficulty, prior computer use, and perception of ability to learn	Factor analysis of items reveal a three-factor structure with loadings ranging from .52 - .91 (Factor I), .35 - .99 (Factor II), and .83 - .88 (Factor III) Factor correlations range from .289 - .719	None reported	**Strengths** Ease of use/ administration, ready availability of forms as a noncommercial product, and short administration time. A measure of potential sustained motivation in computer use. **Limitations** Narrow focus (little coverage of the construct of information literacy, little evidence of correlations with true performance criteria

Table C-1. Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Cornell University Online Tutorials http://.library.corwwwnell.edu/okuref/research/tutorial.html	Tutorial modules include online multiple-choice tests that are scored and banked for future use.	Skills in information literacy and library research.	None reported	None reported	None reported	Individual studies report correlations with college GPA	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. Online programs are easily accessible and free of charge. **Limitations** The tests embedded in the modules have not been placed under psychometric scrutiny.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

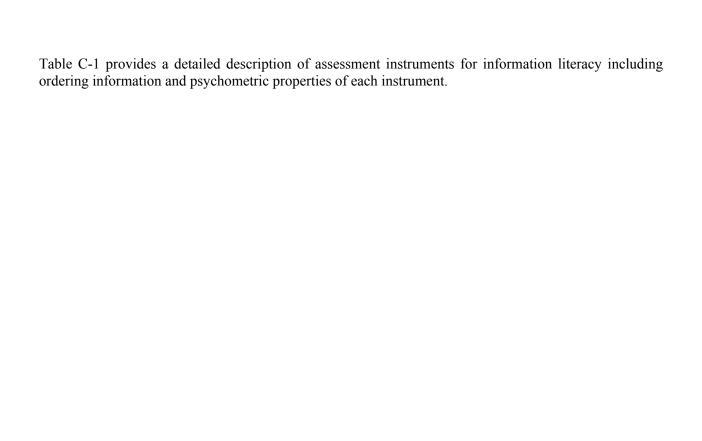| Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
JMU Go for the Gold Online Tutorials http://www.lib. jmu.edu/library/ gold/modules.htm	Tutorial modules include online multiple-choice tests that are scored and banked for future use.	Skills in information literacy and library research.	None reported	None reported	None reported	Individual studies report correlations with college GPA.	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. Online programs are easily accessible and free of charge. **Limitations** The tests embedded in the modules have not been placed under psychometric scrutiny.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Purdue University Online Tutorials (CORE) http://core.lib.purdue.edu	Tutorial modules include online multiple-choice tests that are scored and banked for future use.	Skills in information literacy and library research.	None reported	None reported	None reported	Individual studies report correlations with college GPA.	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. Online programs are easily accessible and free of charge. **Limitations** The tests embedded in the modules have not been placed under psychometric scrutiny.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Griffith University Online Tutorials http://www4.gu. edu.au/shr/lrt	Tutorial modules include online multiple-choice tests that are scored and banked for future use.	Skills in information literacy and library research.	None reported	None reported	None reported	Individual studies report correlations with college GPA.	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. Online programs are easily accessible and free of charge. **Limitations** The tests embedded in the modules have not been placed under psychometric scrutiny.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
University of Texas System Online Tutorials (TILT) http://tilt.lib. utsystem.edu	Tutorial modules include online multiple-choice tests that are scored and banked for future use.	Skills in information literacy and library research.	None reported	None reported	None reported	Individual studies report correlations with college GPA.	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. Online programs are easily accessible and free of charge. **Limitations** The tests embedded in the modules have not been placed under psychometric scrutiny.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
General Attitudes Toward Computers 30-item Likert Scale Designed for middle school, high school, and college students **Author:** M.J. Reese ***In:** Educational and Psychological Measurement, 43, 913–916*	General cognitive and affective attitude toward computers.	General approach to computer use including cognitive and affective attitude.	Cronbach's alpha = .87	Content validity assessed by a panel of experts from middle and high schools and college Evidence for construct validity based on factor analysis of items; factor loadings range from .43 - .81	Factor analysis set a criteria loading of .40 for all items included in the instrument	None reported	**Strengths** Ease of use, simplicity of scoring, and sound psychometric properties. **Limitations** No clear correlation with information literacy outcomes.

Table C-1. Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
James Madison University's Information Literacy and Computer-Based Testing Clusters **Contact** JMU Center for Assessment, JMU, Harrisonburg, VA 22807 (540) 568-6706	Each of five clusters includes a score for information searches related to the discipline-specific category. The information literacy module is an 80-item test. The complete psychometric properties will be forthcoming at release time. Those included here are based on the pilot instrumentation.	Clusters are currently being developed. The arts and humanities cluster is complete and psychometric data are being collected.	Cronbach's alpha = .65	Validation studies in progress	Correlations and descriptive comparisons between paper-and-pencil and computer-based testing Item analysis Student ratings of multimedia items	None yet reported	**Strengths** A rare combination of online tutorials/ computer-based testing and psychometric qualities. Has the best overall potential as a measure of information literacy and other constructs, e.g., critical thinking. **Limitations** Requires sophisticated computerized testing facilities, high band-width, and support services. Furthermore, assessment professionals must be knowledgeable about Item Response Theory in order to tailor tests to local needs.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Attitudes About Computers 30 Likert Scale items Designed to measure computer anxiety in college students **Author** Carol Toris **ERIC Document** ED 254 540	Three derived subscales for 1) appreciation of computers. 2) computer usage anxiety, 3) computer's negative impact on society.	Measures of anxiety in three areas that correlate with approach/ avoidance of computer usage.	Principle components analysis provided evidence of construct validity by including only items with factor loadings about .40.	High computer anxiety significantly correlated with avoidance of learning opportunities in computer-related activities.	Factor analysis with an inclusion criterion of .40 for items	Statistically significant correlations with a Computer Usage Checklist. No coefficient reported.	**Strength** Fast and easy measure of affective component of computer usage and approach/ avoidance regarding computer use. **Limitations** Is not designed as a power measure of computer skills and should be used only to supplement such measures.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
SUNY Brockport's Computer Skills Exam **Project Manager** Edwina Billings 210 Daley Hall SUNY Brockport (716) 395-2666 ebillings@brockport.edu	Computer skills in three areas: (1) Windows 95 (2) Word 2000 (3) Internet Communicator	Task-oriented skills for the three areas concerned.	To be conducted when time and resources become available.	To be conducted when time and resources become available.	None reported	None reported	**Strengths** Task-oriented measures provide scores based on actual use during the tutorials; the web-browser portion provides instruction and testing of important search skills necessary for information literacy. **Limitations** The tasks are focused more on computer skills than on the breadth of information literacy skills.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Element K http://www. elementk.com Company offering online tutorials in a range of computer-related applications including certification preparation.	While the company offers tutorials in a broad range of applications, the Internet Explorer module includes many skills necessary for information literacy.	Task-oriented skill modules require interactivity with the computer application in real time.	None reported	None reported	None reported	None reported	**Strengths** Computer interactivity gives students a real-time experience with a web browser; many of the skills required are applicable to an information search. **Limitations** The tasks are focused on computer skills rather than specifically on information literacy skills.

Table C-1.   Assessment reviews for information literacy in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Relationship with other measures	Strengths and Limitations
Penn State Library Information Literacy Program (Information Literacy and the Library) http://www. libraries.psu.edu/ instruction/infolit/ program.htm	Tutorial modules are based on four "interconnected components": knowledge of information sources; skills in finding, evaluating, using, and communicating information; generalizing knowledge and skills to applied settings; and social context for the use of information.	Skills in information literacy and library research	None reported	None reported	None reported	None reported	**Strengths** Tutorials are specifically aimed at information literacy in its broadest manifestations. The site is comprehensive and leaves few gaps based on current definitions of IL. **Limitations** There are no embedded tests, though there are interactive uses of IL skills.

TABLE D

**LEARNING MODULES AND TUTORIALS
FOR INFORMATION LITERACY**

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Standard One: The information literate student determines the nature and extent of the information needed.					
Performance Indicator: The information literate student defines and articulates the need for information.					
Outcomes					
1. Confers with instructors and participates in class discussions, peer workgroups, and electronic discussions to identify a research topic or other information need					
2. Develops a thesis statement and formulates questions based on the information need	X	X	X	X	
3. Explores general information sources to increase familiarity with the topic	X	X	X	X	X
4. Defines or modifies the information need to achieve a manageable focus	X	X	X	X	X
5. Identifies key concepts and terms that describe the information need	X	X	X	X	X

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
6. Recognizes that existing information can be combined with original thought, experimentation, and/or analysis to produce new information			X		
Performance Indicator: The information literate student identifies a variety of types and formats of potential sources for information.					
Outcomes					
1. Knows how information is formally and informally produced, organized, and disseminated.	X	X	X	X	X
2. Recognizes that knowledge can be organized into disciplines that influence the way information is accessed	X	X	X	X	X
3. Identifies the value and differences of potential resources in a variety of formats (e.g., multimedia, database, web site, data set, audiovisual, book)	X	X	X	X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
4. Identifies the purpose and audience of potential resources (e.g., popular versus scholarly, current versus historical)	X	X		X	X
5. Differentiates between primary and secondary sources, recognizing how their use and importance vary with each discipline	X				
6. Realizes that information may need to be constructed with raw data from primary sources	X				
Performance Indicator: The information literate student considers the costs and benefits of acquiring the needed information.					
Outcomes					
1. Determines the availability of needed information and makes decisions on broadening the information seeking process beyond local resources (e.g., interlibrary loan and using resources at other locations; obtaining images, videos, text, or sound)	X			X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
2. Considers the feasibility of acquiring a new language or skill (e.g., foreign or discipline-based) to gather needed information and to understand its context					X
3. Defines a realistic overall plan and timeline to acquire the needed information	X	X	X	X	X
Performance Indicator: The information literate student reevaluates the nature and extent of the information need.					
Outcomes					
1. Reviews the initial information need to clarify, revise, or refine the question		X		X	X
2. Describes criteria used to make information decisions and choices	X	X	X	X	X
Standard Two: The information literate student accesses needed information effectively and efficiently.					

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student selects the most appropriate investigative methods or information retrieval systems for accessing the needed information.					
Outcomes					
1. Identifies appropriate investigative methods (e.g., laboratory experiment, simulation, fieldwork)					
2. Investigates benefits and applicability of various investigative methods					
3. Investigates the scope, content, and organization of information retrieval systems	X	X	X	X	X
4. Selects efficient and effective approaches for accessing the information needed from the investigative method or information retrieval system	X	X	X	X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student constructs and implements effectively designed search strategies.					
Outcomes					
1. Develops a research plan appropriate to the investigative method	X	X	X	X	X
2. Identifies keywords, synonyms, and related terms for the information needed	X	X	X	X	X
3. Selects controlled vocabulary specific to the discipline or information retrieval source					X
4. Constructs a search strategy using appropriate commands for the information retrieval system selected (e.g., Boolean operators, truncation, and proximity for search engines; and internal organizers, such as indexes for books)	X	X	X	X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
5. Implements the search strategy in various information retrieval systems using different user interfaces and search engines, with different command languages, protocols, and search parameters	X	X	X	X	X
6. Implements the search using investigative protocols appropriate to the discipline					
Performance Indicator: The information literate student retrieves information online or in person using a variety of methods.					
Outcomes					
1. Uses various search systems to retrieve information in a variety of formats	X	X		X	X
2. Uses various classification schemes and other systems (e.g., call number systems or indexes) to locate information resources within the library or to identify specific sites for physical exploration		X		X	X

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
3. Uses specialized online or in-person services available at the institution to retrieve information needed (e.g., interlibrary loan or document delivery, professional associations, institutional research offices, community resources, experts and practitioners)	X	X		X	X
4. Uses surveys, letters, interviews, and other forms of inquiry to retrieve primary information					
Performance Indicator: The information literate student refines the search strategy if necessary.					
Outcomes					
1. Assesses the quantity, quality, and relevance of the search results to determine whether alternative information retrieval systems or investigative methods should be used	X	X		X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
2. Identifies gaps in the information retrieved and determines if the search strategy should be revised	X				X
3. Repeats the search using the revised strategy as necessary	X				X
Performance Indicator: The information literate student extracts, records, and manages the information and its sources.					
Outcomes					
1. Selects among various technologies the most appropriate one for the task of extracting the needed information (e.g., copy/paste software functions, photocopier, scanner, audiovisual equipment, or exploratory instruments)					
2. Creates a system for organizing the information					X
3. Differentiates between the types of sources cited and understands the elements and correct syntax of a citation for a wide range of resources	X	X		X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
4. Records all pertinent citation information for future reference					X
5. Uses various technologies to manage the information selected and organized					
Standard Three: The information literate student evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.					
Performance Indicator: The information literate student summarizes the main ideas to be extracted from the information gathered.					
Outcomes					
1. Reads the text and selects main ideas	X		X		
2. Restates textual concepts in his or her own words and selects data accurately					
3. Identifies verbatim material that can be then appropriately quoted	X	X	X	X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student articulates and applies initial criteria for evaluating both the information and its sources.					
Outcomes					
1. Examines and compares information from various sources to evaluate reliability, validity, accuracy, authority, timeliness, and point of view or bias	X	X	X	X	X
2. Analyzes the structure and logic of supporting arguments or methods					
3. Recognizes prejudice, deception, or manipulation					
4. Recognizes the cultural, physical, or other context within which the information was created and understands the impact of context on interpreting the information					X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student synthesizes main ideas to construct new concepts.					
Outcomes					
1. Recognizes interrelationships among concepts and combines them into potentially useful primary statements with supporting evidence			X		X
2. Extends initial synthesis, when possible, at a higher level of abstraction to construct new hypotheses that may require additional information					
3. Uses computer and other technologies (e.g., spreadsheets, databases, multimedia, and audio or visual equipment) for studying the interaction of ideas and other phenomena			X		

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student compares new knowledge with previous knowledge to determine the value added, contradictions, or other unique characteristics of the information.					
Outcomes					
1. Determines whether information satisfies the research or other information need	X	X	X	X	X
2. Uses consciously selected criteria to determine whether the information contradicts or verifies information used from other sources					
3. Draws conclusions based on information gathered	X		X		
4. Tests theories with discipline-appropriate techniques (e.g., simulators, experiments)					
5. Determines probable accuracy by questioning the source of the data, the Limitations of the information gathering tools or strategies, and the reasonableness of the conclusions	X	X	X	X	X

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
6. Integrates new information with previous information or knowledge					
7. Selects information that provides evidence for the topic	X		X	X	X
Performance Indicator: The information literate student determines whether the new knowledge has an impact on the individual's value system and takes steps to reconcile differences.					
Outcomes					
1. Investigates differing viewpoints encountered in the literature					
2. Determines whether to incorporate or reject viewpoints encountered					
Performance Indicator: The information literate student validates understanding and interpretation of the information through discourse with other individuals, subject area experts, and/or practitioners.					
Outcomes					
1. Participates in classroom and other discussions					

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
2. Participates in class-sponsored electronic communication forums designed to encourage discourse on the topic (e.g., e-mail, bulletin boards, chat rooms)					
3. Seeks expert opinion through a variety of mechanisms (e.g., interviews, e-mail, listservs)	X	X	X	X	X
Performance Indicator: The information literate student determines whether the initial query should be revised.					
Outcomes					
1. Determines if original information need has been satisfied or if additional information is needed	X		X		X
2. Reviews search strategy and incorporates additional concepts as necessary	X	X	X	X	X
3. Reviews information retrieval sources used and expands to include others as needed	X	X	X	X	X
Standard Four: The information literate student, individually or as a member of a group, uses information effectively to accomplish a specific purpose.					

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student applies new and previous information to the planning and creation of a particular product or performance.					
Outcomes					
1. Organizes the content in a manner that supports the purposes and format of the product or performance (e.g., outlines, drafts, storyboards)	X				
2. Articulates knowledge and skills transferred from previous experiences to planning and creating the product or performance					
3. Integrates the new and previous information, including quotations and paraphrasings, in a manner that supports the purposes of the product or performance					
4. Manipulates digital text, images, and data, as needed, transferring them from their original locations and formats to a new context					

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
Performance Indicator: The information literate student revises the development process for the product or performance.					
Outcomes					
1. Maintains a journal or log of activities related to the information seeking, evaluating, and communicating process					
2. Reflects on past successes, failures, and alternative strategies					
Performance Indicator: The information literate student communicates the product or performance effectively to others.					
Outcomes					
1. Chooses a communication medium and format that best supports the purposes of the product or performance and the intended audience					
2. Uses a range of information technology applications in creating the product or performance					

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
3. Incorporates principles of design and communication					
4. Communicates clearly and with a style that supports the purposes of the intended audience					
Standard Five: The information literate student understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally.					
Performance Indicator: The information literate student understands many of the ethical, legal and socio-economic issues surrounding information and information technology.					
Outcomes					
1. Identifies and discusses issues related to privacy and security in both the print and electronic environments					
2. Identifies and discusses issues related to free versus fee-based access to information					X
3. Identifies and discusses issues related to censorship and freedom of speech					

Table D-1.   Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
4. Demonstrates an understanding of intellectual property, copyright, and fair use of copyrighted material					X
Performance Indicator: The information literate student follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources.					
Outcomes					
1. Participates in electronic discussions following accepted practices (e.g., "netiquette")					
2. Uses approved passwords and other forms of ID for access to information resources					
3. Complies with institutional policies on access to information resources					
4. Preserves the integrity of information resources, equipment, systems, and facilities					
5. Legally obtains, stores, and disseminates text, data, images, or sounds					

Table D-1. Learning modules and tutorials for information literacy based on the Association of College and Research Libraries (ACRL) standards—Continued

ACRL standard	Library Research at Cornell http://www.library.cornell.edu/okuref/research/tutorial.html	Purdue University Library CORE+ (Quiz module available for registered users) http://www.lib.purdue.edu/undergrad/handouts.html	James Madison University Information Seeking Skills Test (Quiz module integrated into software) http://www.jmu.edu/gened/transfer2003.html#isst	Griffith University Library Tutorial (Quiz module available for registered users) http://www4.gu.edu.au/shr/lrt/	TILT: Texas Information Literacy Tutorial http://tilt.lib.utsystem.edu
6. Demonstrates an understanding of what constitutes plagiarism and does not represent work attributable to others as his or her own					X
7. Demonstrates an understanding of institutional policies related to human subjects research					
Performance Indicator: The information literate student acknowledges the use of information sources in communicating the product or performance.					
Outcomes					
1. Selects an appropriate documentation style and uses it consistently to cite sources					X
2. Posts permission granted notices, as needed, for copyrighted material					

TABLE E

ASSESSMENT INSTRUMENTS FOR
QUANTITATIVE REASONING AND QUANTITATIVE SKILLS

TABLE E. ASSESSMENT INSTRUMENTS FOR
QUANTITATIVE REASONING AND QUANTITATIVE SKILLS

In a 1994 survey by the National Center on Adult Literacy, 73 percent of all adult literacy programs surveyed used standardized tests to assess mathematical skills and mathematical reasoning **(**Gal and Schmitt 1994**)**. A survey of programs in Massachusetts revealed that 84 percent of programs in which learners were placed using assessment instruments used standardized tests, and only 10 percent used program-developed assessment procedures (Leonelli and Schwendeman 1994, p. 42). In both surveys, the Test of Adult Basic Education (TABE) was the instrument of choice (48% and 53% respectively). The TABE Applied Mathematics portion is profiled in this Sourcebook. Yet it is clear from the assessment standards of groups such as the Mathematics Association of America and the National Council of Teachers of Mathematics that other types of instruments as well as domains should be considered. The templates included in this chapter address a variety of these instruments.

Tests that purport to measure quantitative reasoning come in a variety of forms and have a variety of potential applications. It would have been easiest to stay within a single, typical test design with the usual administrative properties, but that might have misled assessment professionals into thinking that a singular approach has been taken in the development of tests of quantitative reasoning. Therefore, the templates include tests with a variety of approaches, as well as both group and individually administered tests (the type many learning centers might use). Furthermore, many of these tests are designed for a variety of purposes, and the measurement of quantitative reasoning may be only one subtest. As recommended by the Mathematics Association of America, to measure the affective components of acquiring quantitative literacy, a Mathematics Anxiety Rating Scale is included in the templates.

There is an emerging class of measurement similar to the modules developed at James Madison University (JMU), as discussed in chapter 5 and table C, for the measurement of quantitative literacy. The faculty and staff at JMU have already developed a critical thinking module and are in the process of developing instrumentation for quantitative literacy. Any individuals who are inclined to develop their own campus-based instrument in this area might find it useful to explore this avenue of computer-based testing. The JMU modules will be available for use on the dedicated server in early 2004.

The CAAP (Collegiate Assessment of Academic Proficiency), the Academic Profile, and C-Base (College Basic Academic Skills Test) are not reviewed in this Sourcebook because Dary Erwin included templates of all three (as they applied to writing) in the *NPEC Sourcebook on Assessment*: *Volume 1*. An assessment professional interested in these tests should refer to Dary's templates, which can be accessed at http://nces.ed.gov/pubs2000/2000195.pdf (Acrobat Reader is required). Each of these tests has a quantitative or mathematical skills section among the matrix of subtest offerings. The CAAP is not reviewed in the templates because substantive changes are being made in the mathematics subtest. Assessment professionals may wish to look at it when the new forms are in place and have been put through appropriate psychometric paces. The CAAP can be given as individual subtests. The Academic Profile is given as an entire matrix, so that those looking for a stand-alone mathematics test will have to judge whether or not to use the entire instrument and extract quantitative portions (an expensive way of getting a mathematics score). However, test developers of the College BASE at the University of Columbia-Missouri have recently begun marketing the subsections of the College BASE, and a template for the mathematics portion of the test is included in this Sourcebook.

One paper-and-pencil test, the Stanford Diagnostic Mathematics Test, Fourth Edition, is included in the templates for quantitative reasoning and quantitative skills. This test has sound psychometrics that might serve to measure quantitative reasoning and quantitative skills. A caveat is that it requires at least a solid

high-school mathematics background. Thus, it would probably not serve well as a test of liberal arts mathematics.

Tables E-1 and E-2 provide a detailed description of assessment instruments for quantitative reasoning, including ordering information and psychometric properties of each instrument.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
The Quant Q: A Measure of Quantitative Reasoning Skills 17 multiple-choice items **Author** Stephen Blohm **Publisher** California Academic Press 217 La Cruz Ave Millbrae, CA 94030 (650) 697-5628 http://www.calpress.com/quantq.html **Date** 1999 **Testing Time** 30 minutes **Cost** $125/25 $230/50 $400/100 includes scoring	Quantitative reasoning	Reasoning skills in relation to quantitative problems. Mathematics skills for the test are basic; the problems are designed to measure "one's ability to think outside the box" when solving quantitative problems. Appropriate for undergraduates, students seeking advanced degrees, and professionals.	KR-20 = .78 - .86	Validity studies are in progress.		California Reasoning Appraisal = .84 California Critical Thinking Skills Test = .73	**Strength** The instrument has good psychometric reliability for a short test. It takes little time to administer. **Limitations** The test has a narrow focus with little evidence of transfer to other skills within the cognitive domain that may be encompassed by the construct of quantitative reasoning.

Table **E**-1. Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
ABLE (Adult Basic Learning Examination) Level 3 **Publisher** The Psychological Corporation 555 Academic Court, San Antonio, TX 78204-2498 (800) 872-1726 http://www.psychcorp. com/catalogs/paipc/ paipc_toc.htm **Testing Time** 1 hr. 10 min. for number operations and problem solving **Cost** $84/25	Mathematics (number operations and problem solving) Also available: vocabulary, reading comprehension, language Raw to scale score conversions included in technical manual	Number operations is a 36-item subtest including interpreting fractions, factorization, equations, etc. Problem solving is a 40-item test of ability to determine an outcome, record/ retrieve information, etc.	KR-21 for Number Operations = .90 KR-21 for Problem Solving = .90	Detailed analysis of items Intercorrelations range from .44 - .71 Concurrent validity with Stanford Achievement Test = .81 for total mathematics	Sample of 4,000 adults; items reduced using difficulty values, p-values (item analysis), and biserial/point biserial correlations	Stanford Achievement Test = .81	**Strengths** The two mathematics subtests take a life-experience approach rather than reliance on algebraic achievement. This test could be useful particularly in the assessment of mathematical reasoning among liberal arts students. **Limitations** Test is designed for adult population and normed as such. Local norms would be necessary. One may have to discard subtests if not appropriate to the assessment target.

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
Differential Aptitude Tests, 5th Ed.; Form C Level 2 **Publisher** The Psychological Corporation (address in previous template) **Testing Time** Numerical reasoning subtest for personnel and career assessment contains 25 items taking 20 minutes to administer. **Cost** $121 per 25 booklets $63.50 per 25 scan answer sheets	Numerical ability Also available: Verbal reasoning, abstract reasoning, perceptual speed and accuracy, mechanical reasoning, and space relations	The 25-item, numerical ability subtest is designed to measure reasoning not computation, therefore, the mathematics skill level is below grade level to focus on task rather than computation.	KRa0 combined grade 12 sample = .91 Alternate-forms reliability for level 2 grade 11 = .74; level 1 grade 8 = .83	Form C combined sample intercorrelations range from .30 (perceptual speed) to .92 (scholastic aptitude)	Items selected through a National Item Tryout Program; items reduced using item analysis (biserial correlations within a range of .35 - .70)	Technical manual reports a variety of correlations across many samples for ACT, Armed Forces Vocational Aptitude Test, Otis-Lennon School Ability, etc.	**Strengths** Test has psychometric properties, provides; solid effort to test mathematics reasoning rather than computation, and provides a vast database on relationships with other instruments. **Limitations** Some collegiate environments may see a ceiling effect especially among science students. The standardization population was high school students for level 2. The spread of item difficulty should make the test appropriate for most colleges especially for students in the liberal arts or social sciences.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
K-Fast (Kaufman Functional Academic Skills Test) Individually administered test designed to measure functional mathematical skills and reading **Publisher** American Guidance Service, Inc. 4201 Woodland Rd. Circle Pines, MN 55014-1796 http://www.agsnet.com/templates/productview.asp?Group=a3540 **Cost** $144.95/for complete kit with 25 scoring records	Functional mathematical skills and reading Raw to standard score conversion tables are included in the manual.	Designed to supplement intelligence, achievement, and adaptive behavior tests. Tests how well an individual applies mental ability to life tasks demanding mathematical reasoning.	Mean split-half reliability across all age groups = .88 Test-retest reliability = .87 Mean SEMs across all age groups = 5.1	Factor analysis showed a g-factor loading of .79 with general cognitive tests. Also showed loadings of .59 for crystallized intelligence and .33 for fluid intelligence.	A broad bank of items was reduced using item analysis, reliability analysis, correlations analysis, and factor analysis.	WISC-R full scale = .77 WAIS-R full scale (ages 16 – 34) = .69 Stanford-Binet IV Composite = .86	**Strengths** The test takes a functional approach to the measurement of mathematical reasoning. Items include interpretation of everyday data that disconnects it from conventional mathematics or algebra tests. It has psychometric properties. **Limitations** It is designed as an individual test though this reviewer sees no reason why it could not be adapted for group use. It may have a ceiling effect in mathematically oriented student populations. The reading section would have to be discarded for those not in need of this portion.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
Scholastic Abilities Test for Adults **Publisher** Pro.ed 8700 Shoal Creek Blvd. Austin, TX 78758 (512) 451-3246 http://www.proedinc.com/ store/index.php?mode= product_detail&id=0931 **Testing Time** 15 minutes, scoring sheets included **Cost** $159/ complete kit (10 test books, 25 response books, & 25 profile records	Quantitative reasoning (QR) Also available: Verbal reasoning, nonverbal reasoning, reading vocabulary, reading comprehension, mathematics calculation, mathematics application, writing mechanics, and writing comprehension	QR subtest measures one's ability to recognize and apply quantitative concepts. The examinee is presented a visual stimulus series of numbers that has a missing number. The problem is then computed.	Cronbach's alpha for age composites = .87 Test-retest reliability (measure of stability only by parsing out alpha) = .95	Item analysis using point biserial; mean r's across age groups = .49 Correlations with established test used for construct validity (see next column) Intercorrelations across subtests reported	Item bank reduction carried out using point biserial with a .3 minimum score for item discrimination.	The technical manual includes correlations with a variety of established instruments, including the ACT composite = .45, WAIS-R full scale = .60, WRAT arithmetic = .87, etc.	**Strengths** The test is brief but has psychometric properties. It has a scale aimed specifically at quantitative reasoning with little computational dependence. It is a useful test for liberal arts programs, though others could benefit. **Limitations** The test, while normed across a variety of age groups, is aimed at an adult population rather than a college population. There is the possibility of a ceiling effect in mathematically oriented, college student populations.

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
Mathematics Anxiety Rating Scale (MARS) **Author** Richard M. Suinn, Ph.D. 808 Cheyenne Drive, Fort Collins, CO 80525 (970) 491-1351 http://www.colostate.edu/Depts/NatSci/html/Suinn.html **Testing Time** A 95-item Likert-scale test taking 15 to 20 minutes **Cost** $60/100 tests	Mathematics anxiety	A total raw score is compared to a percentile scale. The author suggests that a score above the 75th percentile is a candidate for intervention by a learning center. Author further suggests that local norms be developed.	Test-retest r = .78 Coefficient alpha = .97	R with the Differential Aptitude Test (DAT) (see above) = -.35 R with grades in a mathematics course = -.29; with number of yrs. in mathematics = -.44; with number of yrs. in calculus = .21		R with DAT = -.35	**Strengths** Instrument provides a psychometrically sound measure of the affective response to all things mathematical. Assessment professionals may find this a useful way to broaden measures of math-related competence. **Limitations** As with any self-reported, affective scale, it should be paired with reliable direct tests and not be substituted for a cognitive measure of quantitative reasoning or skills.

Table E-1. Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
Employee Aptitude Survey **Publisher** Psychological Services, Inc., 100 West Broadway, Suite 1100, Glendale, CA 91210 (818) 244-0033 http://www.psionline.com/skills_aptitude.htm **Testing Time** A 20-item speed test to be taken in 5 minutes. **Cost** $50.75 for 1 – 19 $43.50 for 20 – 100; $10.50 for scoring key; $21.25 for examiner's manual	Numerical reasoning Also available: Verbal comprehension, numerical ability, visual pursuit, visual speed and accuracy, space visualization, word fluency, verbal reasoning, symbolic reasoning, manual speed	Designed to measure the ability to analyze logical relationships and to discover underlying principles. All items are number series extrapolation whereby the examinee selects the next number in a series of five numbers.	Alternate forms r = .81; SEM = 1.84 Test intercorrelations for form 1 (EAS 1) with different occupational groups range of r = .06 (Engineering students) to .57 (leadworkers); slightly higher r's with EAS 2	Factorial validity with factor loading of .57 for reasoning (technical manual reports all factor loadings). Meta-analysis of 725 validity coefficients from 160 studies reported in technical manual.	Factor analysis on items selected by experts to reduce to appropriate items in each subtest area.	Dozens of correlations and factor loadings reported in the technical manual, including WAIS (loading of .81 with quantitative reasoning), admissions test for graduate study in business, and more.	**Strengths** Instrument is a short, easy-to-administer test of numerical reasoning with great psychometric properties. **Limitations** Theoretical sticklers may wonder if a simple number series test can adequately measure a complex construct like numerical reasoning. One must trust completely the psychometric properties as evidence of the tests construct validity.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitations
Test of Adult Basic Education (TABE) Form 7 & 8, level A (advanced; for grade levels to 14.9) **Publisher** CTB/McGraw-Hill 20 Ryan Ranch Rd., Monterey, CA 93940-5703 (800) 538-9547 http://www.ctb.com/ products_services/tabe/ index.html **Testing Time** A 50-item test with a 50-minute time limit Machine scoring done by McGraw-Hill or a software package available; hand-scoring also an option **Cost** $79 per 25 tests	Applied mathematics Also available: Reading, mathematics computation, language, and spelling	Applied mathematics is designed to assess the examinee's "ability to apply a wide range of basic mathematical skills, methods, and concepts to tasks such as budgeting, planning, predicting results, and interpreting data."	Complete battery KR-20 for level 7, college students = .93 KR-20 for level 8, college students = .92. Inter-correlations, item parameters, and item difficulties, and error curves included.	Content validity based on expert selection (see next column).	Items developed by content specialists then items from a sample administration were subjected to IRT using a three-parameter logistic model taking into account item discrimination, difficulty, and guessing for each item.	None reported	**Strengths** The TABE is the most frequently used of tests for the measurement of applied mathematics (quantitative reasoning) among adults. Items lean heavily toward real-life problems, though some computation is required. **Limitationss** A ceiling effect may occur in mathematically oriented, college student populations. Test appropriate for nontechnical programs.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
The RBH Arithmetic Reasoning Test, Forms I & II (Richardson, Bellows, Henry & Co., Inc.) A problem-solving test also available as a measure of "mathematical and analytical reasoning." **Marketed by** EPredix 3500 US Bank Place, 601 Second Ave. South, Minneapolis, MN 55402 (202) 659-3755 http://www.epredix.com/ **Testing Time** A 25-item test; examinees given 15 minutes for completion **Cost** $55/25 tests Price drops to $50 for 25 packs of 25	Arithmetic reasoning	The manual describes the test as "a short, steeply scaled test…of the basic arithmetic operations…whose content involves: determination of selling price; distribution of costs; discounting; production rates; wage and salary rates; overtime procedures; tax operations, dividend and profit determination and the like."	Split-half reliability studies on various occupational populations (petroleum products salespersons, etc.) range form Spearman-Brown r = .83 - .91	Criterion-related validity studies using supervisors rankings range of r = .29 - .38 Using on-the-job performance tests, range of r = .41 Manual reports many other types of studies based on performance ratings with a wide range of r values.		Learning ability Form S, range of r = .62 - .75 Form T, range of r = .49 - .67 Watson-Glaser Critical Thinking Test, r = .56 Shop Arithmetic Form I, r = .65 Manual reports correlations with several other instruments with range of r = -.10 - .59	**Strengths** The test is a problem-based test of arithmetic reasoning, though some calculation is necessary. The test is an example of an industrial or business-based use of tests with criterion-related studies focused clearly on occupational task. This test might provide a clear counterpoint to typical college-based tests of quantitatively based skills. **Limitationss** There may be a ceiling effect with mathematically oriented, college student populations, though the test is clearly focused on problem solving so that reading ability and logic play a key role in deriving answers. Revisions would have to be made for machine scoring of mass testing because work is shown in the left margin and the answer is placed in a box at right.

Table E-1.   Assessment reviews for quantitative reasoning in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
The Test of Everyday Reasoning 35 multiple-choice items **Author** Peter Facione **Publisher** California Academic Press 217 La Cruz Ave Millbrae, CA 94030 (650) 697-5628 http://www.calpress.com/ quantq.html **Date** 1998 **Testing Time** 30 Minutes **Cost** $125/25 $230/50 $400/100 Add 7% shipping	Analysis, inference, evaluation, deduction, and induction A test of critical thinking, but closely related to quantitative reasoning in much of its content.	Reasoning skills in relation to quantitative (graphing, flow charts, geometric shapes, etc.) and word problems. Mathematics and verbal skills needed for the test are basic. The Delphi definition of critical thinking is used for the subscores. Designed for middle and high school as well as adult populations.	KR-20 across four sample populations range from .72 - .89	Correlation with the Cornell Critical Thinking Test = .766 Correlation with GPA (level not defined) = .445		The following rho values are reported: PSAT mathematics = .527, verbal = .505; Iowa Educational Development advanced quantitative thinking = .522, Total quantitative thinking = .521, literary materials = .533; ACT mathematics = .413, English = .388, science reasoning = .524.	**Strengths** Instrument has good psychometric reliability for a short test; takes little time to administer. **Limitationss** Instrument has a narrow focus with little evidence of transfer to other skills within the cognitive domain that may be encompassed by the construct of quantitative reasoning.

Table E-2.   Assessment reviews for quantitative skills in terms of selected assessment characteristics and source

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
The Mathematics Association of America Test Series: Mathematics Test CR/1C, Algebra Test Form 4G, and Calculus Readiness Test Form 1E **Publisher** Mathematics Association of America 1529 18th St., NW Washington, DC 20036 **Testing Time** Each test consists of 25 – 32 multiple-choice items taking 30 – 40 minutes to administer. NOTE: The MAA discontinued publication of the series after 1995 but allow continued use of these tests at no charge to colleges and universities.	Total correct for each test: Algebra, mathematics, and calculus readiness Scoring norms have been provided by MAA though since discontinuance, MAA and this author suggest local norms.	The test was designed for use in placement and value-added assessment.	KR – 20 range .8 - .85	Criterion-related validity .45 - .5 with beginning mathematics grades; the strongest criterion-related correlations are the Calculus Readiness Test and beginning science/ mathematics calculus classes. Content validity: a board of mathematicians selected items based on mathematics needs in college-level courses.			**Strengths** These are straightforward mathematics tests designed specifically to measure mathematics skills. The psychometrics were solid and local administrations at several institutions bear out the psychometric desirability of these instruments for placement and value-added assessment. **Limitationss** These tests would be beyond the level of so-called "liberal arts" mathematics courses and would probably result in a floor effect.

Table E-2.   Assessment reviews for quantitative skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
Stanford Diagnostic Mathematics Test, 4th Ed., Forms J/K This test is one of the few covering both quantitative reasoning and skills. Multiple-choice and free-response tests available; hand-scoring sheets included in package; raw to scaled score conversion tables included. **Publisher** Harcourt Educational Measurement P.O. Box 839954 San Antonio, TX 78283-3954 (800) 872-1726 http://www.hemweb.com/trophy/mathtest/sdmt4.htm **Cost** Complete battery and all summaries, both MC and open-ended format: $4.50 first package, $2.25 each additional one	Concepts and applications; computation Subtests of C & A include number systems and numeration, problem solving, graphs and tables, statistics and probability, and geometry and measurement	The test is intended to diagnose students' strengths and weaknesses in the major componental areas of mathematics in grades K - 14. The manual asserts that the uses can be both summative and formative in nature.	KR-21 for combined scores for first-year college students = .91 (Form J) and .86 (Form K). Combined subtest for first-year college students KR-21 range from r = .27 - .82 (Form J) and r= .40 - .86 (Form K). Total alternate forms r = .91	Intercorrelations with corresponding subtests, range of r = .58 - .91 Correlation of complete third edition with complete fourth edition r = .78	Expert item writers submitted a bank of items in each area then reviewed by content and measurement experts. After initial sampling data received, items were tested by traditional methods as well as Rasch modeling; final item selection based on p-value distributions; bias panels also reviewed items.	Correlation between SDMT 4 total scores and Otis-Lennon School Abilities Test r = .63 Subtests between SDMT 4 and OLSAT range r = .43 - .63	**Strengths** The test has first-year collegiate norms making it a good candidate for assessing mathematics skills in the general education distribution. Item difficulty is sufficient to avoid a potential ceiling effect in mathematically oriented, college student populations. Measures an array of skills. **Limitationss** There are computation requirements in much of the battery making the test as much a measure of mathematics skills as quantitative reasoning. This may be outweighed by the lack of ceiling effect in some populations. Hand scoring could be cumbersome in large populations, so machine conversions must be made.

Table E-2. Assessment reviews for quantitative skills in terms of selected assessment characteristics and source —Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
New Standards Mathematics Portfolio: High School Portfolio guidelines and sample folder with spaces for exhibits, entry slip masters, and scoring profile for all categories of skills recommended by NCTM Though this tool is designed for high school, any assessment professional considering or working with an academic department considering mathematical portfolios either for quantitative skills or quantitative reasoning should have a look at this package. **Publisher** National Center on Education and the Economy 700 11th St. NW Washington, DC 20001 (202) 783-3668	Recommended exhibits include the following: conceptual understanding (number and operation, geometry and measurement, function and algebra, and statistics and probability), problem solving, project exhibits (data study, mathematical modeling, design of a physical structure, management and planning, pure mathematical investigation, and history of mathematical idea), skills and communication.	This package represents a thorough set of guidelines including mathematical domains, goals and objectives, and logistical materials for a complete portfolio for mathematical skills in the areas listed. Scoring profiles provide four discrete steps for evaluating each portfolio entry.	N/A	N/A	N/A	N/A	**Strengths** Though designed for the high school level, this system could be easily modified for the collegiate level. It provides a powerful shortcut to anyone looking into mathematics portfolios especially as a general education requirement. **Limitationss** There would have to be slight modifications made for use at the collegiate level. Additionally, sticklers for direct tests may wish to administer a standardized paper-and-pencil test and run correlations with portfolio scores as a test of reliability.

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
BRIGANCE Employability Skills Inventory (ESI) Mathematics skills and concepts may be given separately. Individually or group-administered multiple-choice format with individual learner record supplied **Publisher** Curriculum Associates, Inc. PO Box 2001 North Billerica, MA 01862 (800) 225-0248 **Author** Albert Brigance **Cost** ESI Battery $161.90	Mathematics skills and concepts in an array of scores across a range of skill levels, including computation of whole numbers, conversion of fractions and decimals, geometric concepts, conversion of measurements, mathematics abbreviation and symbols, quantitative concepts, problem solving, time concepts, calendar, estimation and rounding.	The instrument is primarily a diagnostic tool and a prescriptive for curricular placement and adjustment. It is included here as a tool for the most basic skills and is most appropriate for community college/vocational education and adult education. Could be used in an undergraduate institution where expected mathematics skills are a minimal requirement.	Author states that the purpose of the test is for value-added assessment No normative data are supplied and though the author further states that the instrument is correlated with the Comprehensive Adult Student Assessment System (CASAS) and the Secretary's Commission on Necessary Skills (SCANS), but no correlations are supplied.	Author states that the purpose of the test is "criterion-referenced assessments" based on individual curricular program needs.		Author mentions correlations with CASAS and SCANS, but no correlations are provided.	**Strengths** The instrument provides a broad array of mathematics and other skills that can be used to create an individual learner profile. It would be very useful in adult programs, vocational programs, community college programs, and some programs relying solely on a liberal arts mathematics orientation. **Limitationss** A ceiling effect in any science or engineering program at the undergraduate level could occur. The individualized learning profile would not be an advantage in programs seeking only large assessment data sets. This instrument is best used in programs requiring basic mathematics skills where diagnostics and formative evaluation are important.

Table E-2.　Assessment reviews for quantitative skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
Major Field Achievement Test: Mathematics II Test 50 multiple-choice items spread over all domains of the mathematics curriculum at the college level. Test booklets and answer sheets must be returned to ETS for scoring. **Publisher** Educational Testing Service Princeton, NJ 08541 (609) 683-2272 **Cost** $24.50 per test	The items are divided into the following categories: calculus, 35%; linear and abstract algebra, 35%; additional topics (probability, statistics, logic, etc.) 30%.	The instrument is primarily designed for math-intensive curricula. It is mainly administered by mathematics departments, but is also used by engineering and science departments to provide summative evaluation of curricula.	Content validity by panel of experts	New studies being currently conducted	Items are designed and selected by faculty representing the discipline from varying backgrounds.		**Strengths** Test has solid design and psychometrics. A broad array of topics covered with normed subscores is provided. Test security is also provided. **Limitationss** Test would not be appropriate for a general education approach seeking to measure a floor level of skill. A floor effect would result if given to liberal arts students to measure quantitative reasoning or skills.

Table E-2.   Assessment reviews for quantitative skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
University of Wisconsin System Mathematics Test See http://outreach.math.wisc.edu/local/ Courses/mathprep/ placemnt.html **Publisher** University of Wisconsin Center for Placement Testing, University of Wisconsin, Madison, WI 53706 (608) 263-4291 An array of multiple-choice items in all major areas of mathematics **Cost** $5.00/battery. Pricing is currently being discussed but NPEC has been informed that the price will be close to the language test series costing	The tests are arranged in three sections, A, B, and C. These sections combined contain the following scores and percentage weights: Elementary algebra: arithmetic – 25, basic algebra – 47.5, intuitive geometry – 27.5 Intermediate algebra: basic algebra – 50, advanced algebra – 40, geometry – 10 College algebra: advanced algebra – 70, analytic geometry – 30 Trigonometry: Trigonometry – 90, geometry – 10.	The battery is so designed that students of all backgrounds may be tested with varying elements of the battery. The primary purpose at UW is for placement into college mathematics courses. For example, entering liberal arts students may take only Section A containing items in elementary and intermediate algebra for placement. However, as students progress through a program, other test sections may be added.	Below are the reliability coefficients of each form within each section of the battery. Section A: Form 821 = .865; Form 83X = .867; Form 87X = .865; Form 89X = .848 Section B: Form 821 = .941; Form 83X = .937; Form 87X = .941; Form 89X = .933 Section C: Form 821 = .875; Form 83X = .863; Form 87X = .865; Form 89X = .862	New validity studies currently underway.	A panel of experts in the field from inside and outside the University gather to write items. The items are piloted often on more than one occasion to distinguish between students at various levels of mathematical skill.		**Strengths** The battery covers all major areas of mathematics skill. It is a proven placement instrument with strong psychometric properties that could be adapted to an array of diagnostic and value-added assessment models. **Limitationss** It has no calculus section, though the instrument will suffice for precalculus.

Table E-1.   Assessment reviews for quantitative skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
College BASE Mathematics Test An array of multiple-choice items in general mathematics proficiency, algebra, and geometry *NOTE:* the mathematics clusters are a subset of a broader general education assessment. **Publisher** Assessment Resource Center University of Missouri-Columbia, College of Education 2800 Maguire Blvd., Columbia, MO 65211 (573) 882-4694 e-mail contact: HumphreysP@ missouri.edu **Cost** $1.00 – 2.00/student plus $5.70 for scoring each mathematics subtest	Tests are arranged in clusters. An individual student score is provided for each cluster and for the overall mathematics subject. The institutional report summarizes cluster and subject scores. The mathematics portion contains three clusters: General mathematics proficiency: Use mathematics techniques in the solution of real-life problems, use language, notation and deduction to express quantitative ideas, etc. (mean of test = 301, s.d. = 49.1).	The college base is, according to the descriptive brochure, "a carefully constructed assessment tool supported by advanced technical expertise and capabilities." The mathematics (and each) portion of the test is criterion-referenced and assesses students' knowledge and mastery of specific skills.	Below are the reliability coefficients using KR20 as a measure of internal consistency for each cluster of the mathematics test. General mathematics proficiency: KR20 = .76 (practical applications = .64, properties & notations = .56, using statistics = .55) Algebra: KR20 = .82 (evaluating expressions = .68, equations & inequalities = .73)	Content validity: Evidence based on a systematic and careful construction of all facets of the test (see technical manual, pp. 108 – 109). Criterion-related validity: based on correlations between C-BASE and ACT, SAT, and GPA; r-values the overall mathematics test and ACT = .64, SATV = .23, SATQ = .58, GPA = .43 (canonical correlations also reported and significant) Construct validity: a thorough factor analysis of each item within clusters and subtests (see manual, tables 60 – 65 for factor loadings)	The College BASE has undergone a thorough, iterative development process. It is necessary for assessment professionals to consult the *College BASE Technical Manual* by Osterlind and Merz for a full discussion of a process that clearly took several years.	See criterion-related validity	**Strengths** The clusters cover the essential areas of mathematics skill. It is a proven assessment instrument with very strong psychometric properties that could be adapted to an array of diagnostic and value-added assessment models at both the individual and institutional level. **Limitationss** The focus seems to be on basic skills, so institutions requiring calculus-level assessments may have to add a further instrument.

Table E-2. Assessment reviews for quantitative skills in terms of selected assessment characteristics and source—Continued

Name/Description	Scores	Definition	Reliability	Validity	Method and design	Correlation with other measures	Strengths and Limitationss
	Algebra: Evaluate algebraic and numerical expressions and solve equations and inequalities (mean of test = 282, s.d. = 39.9) Geometry: Recognize two- and three-dimensional figures and use properties of two- and three-dimensional figures for geometric calculations (mean of test = 292, s.d. = 46.4. The overall mathematics test mean = 290, s.d. = 54.2)		Geometry: KR20 = .72 (two- and three-dimensional figures = .46, geometrical calculations = .67)				