

# COGNITIVE ASSESSMENT OF LANGUAGE AND MATH OUTCOMES

Sue Legg  
James Algina  
Editors

## Volume XXXVI

in the series  
ADVANCES IN  
DISCOURSE PROCESSES



# **Cognitive Assessment of Language and Math Outcomes**

edited by

**Sue Legg**

and

**James Algina**

**Volume XXXVI in the Series  
ADVANCES IN DISCOURSE PROCESSES  
Roy O. Freedle, Editor**



**ABLEX PUBLISHING CORPORATION  
NORWOOD, NEW JERSEY**

Copyright © 1990 by Ablex Publishing Corporation

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without permission of the publisher.

Printed in the United States of America.

Library of Congress Cataloging-in-Publication Data

Cognitive assessment of language and math outcomes / edited by James Algina and Sue Legg.

p. cm. — (Advances in discourse processes ; v.36)

Includes bibliographical references and index.

ISBN 0-89391-541-6. — ISBN 0-89391-542-4 (pbk.)

1. Critical thinking—Ability testing. 2. Reading—Ability testing. 3. Mathematical ability—Testing. 4. Educational tests and measurements. I. Algina, James. II. Legg, Sue M. III. Series.

LB1590.3.C64 1990

370.15'2—dc20

90-36747

CIP

Ablex Publishing Corporation  
355 Chestnut Street  
Norwood, New Jersey 07648

# About the Institute

The Institute for Student Assessment and Evaluation at the University of Florida was funded by the State Legislature and provides assessment services to the Department of Education, and through the Department, to the faculty and staff of the school districts and postsecondary institutions of the State of Florida. The Institute has four major functions:

- Planning and implementing long-term research and development activities
- Solving technical problems which have policy implications
- Providing test development services to the State Assessment Program
- Expanding measurement capabilities of educational personnel within the state.

A high priority need for the State of Florida is to develop alternative ways of assessing student performance in both basic skills and higher order intellectual skills (for example, problem solving). The alternatives may include new content for items in traditional paper-and-pencil formats, new formats for paper-and-pencil tests, and non-paper-and-pencil formats.

As an initial activity, the Institute commissioned a series of papers that proposed research appropriate for the development of alternative measurement procedures. Several of these papers were presented at an invitational conference organized by the Institute in February 1986. The chapters in this book were selected from that series. Other research and development activities the Institute has completed or engaged include:

- Development of standards for the selection, use, and interpretation of standardized tests
- Development of standards for administering standardized tests
- Comparison of estimation methods in item response theory
- Use of multiple category models in achievement testing
- Development of computerized tailored testing for the College Level Academic Skills Test
- Comparison of statistical methods for comparing schools using achievement test data

**x    *About the Institute***

Case studies of students who have repeatedly failed a high school  
assessment test or the College Level Academic Skills Test  
Literature review for the differential effects of testing practices  
on various ethnic groups  
Investigation of differential effects of time limits on Hispanic  
and Caucasian examinees

# Overview

The authors have all published extensively in the cognitive approaches to measurement. A brief introduction to the authors is given below.

**Robert H. Ennis** is a Professor of the Philosophy of Education at the University of Illinois. He has been the Director of the Institute of Critical Thinking Project since 1970 and was a fellow at the Center for the Advancement of the Behavioral Sciences at Stanford University in 1983-1984. Professor Ennis has been engaged in research and writing about critical thinking for the past 30 years and has published five tests of critical thinking skills.

**Norman Frederiksen** is Distinguished Research Scientist Emeritus at the Educational Testing Service. He was the recipient of the American Psychological Association award for Distinguished Contributions to Knowledge. Dr. Frederiksen has been at the forefront of the research on the complex cognitive skills that are difficult to measure with paper and pencil tests.

**Audrey N. Grant** is Senior Lecturer, School of Education at La Trobe University in Australia. Professor Grant has taught at both the high school and university level and has authored three books on the teaching and assessment of reading. She has conducted research on and has served as a consultant for the area of adult literacy.

**Peter Johnston** is Assistant Professor of Reading at the State University of New York at Albany. His research interests include the assessment of reading and the prevention of reading problems. Professor Johnston has published a monograph entitled *Reading Comprehension Assessment: A Cognitive Approach* and has a book forthcoming on Reading Assessment.

**Sandra Marshall** is Associate Professor of Psychology at San Diego State University in California. Her fields of interest include cognitive psychology, information processing and problem solving, and statistics and measurement. Professor Marshall has served on the California Department of Education committee for the Model Curriculum for Mathematics for grades K-8 and the Mathematics Advisory Committee for the California Assessment Program. She is conducting research on understanding schemas in problem solving.

**James H. Royer** is Professor of Psychology at the University of Massachusetts. His research interests include the measurement of reading comprehension, the measurement of understanding and the evaluation of social programs. Professor Royer has published extensively in journals in reading and educational research and serves on the editorial boards of several major journals. His most recent book is entitled *Educational Psychology: Application and Theory*.

This book provides a discussion of the cognitive approaches to assessment from three vantage points. Two of the chapters give an overview in which critical thinking is examined as an intellectual activity to be measured by the way in which students approach problem solving. Ennis examines elements in the process of critical thinking and Frederiksen discusses strategies for students to use for problems when there is no single clearly correct solution.

The three chapters by Johnston, Royer, and Grant discuss the implications of cognitive process theory for the assessment of reading. Each of the authors criticizes the current practice of assessing only the outcomes of reading, because it emphasizes what is termed *products* as opposed to the *process* of reading. Product assessment, the authors believe, tends to compartmentalize reading skills. Reading is better understood by observing the students' ability to interact with the text; each author discusses an approach to the observation and assessment of an interactive reading process.

The final chapter applies cognitive theory in order to understand the way in which students solve arithmetic problems. Marshall illustrates how the clarification of the mental requirements involved in solving mathematics problems helps to provide assessment data that can be used to help students improve their performance.

- Vol. XXXVI. Cognitive Assessment of Language and Math Outcomes. Sue Legg & James Algina (Eds.), 1990.
- Vol. XXXVII. Pragmatics, Discourse and Text: Some Systemically-inspired Approaches. Erich H. Steiner and Robert Veltman (Eds.), 1988.
- Vol. XXXVIII. Conversational Organization and its Development. Bruce Dorval (Ed.), 1990.
- Vol. XXXIX. Developing Discourse Practices in Adolescence and Adulthood. Richard Beach & Susan Hynds (Eds.), 1990.
- Vol. XL. Text and Texture: Patterns of Cohesion. Sally Stoddard, 1990.
- Vol. XLI. Conversation Analysis of a Marital Therapy Session: The Pursuit of a Therapeutic Agenda. Jerry Edward Gale, 1990.
- Vol. XLII. Medical Discourse and Systemic Frames of Comprehension. Ronald J. Chenail, 1990.
- Vol. XLIII. What's Going On Here: Complementary Studies of Professional Talk. Allen D. Grimshaw (Ed.), 1991.



# Preface to the Series

**Roy O. Freedle**

Series Editor

This series of volumes provides a forum for the cross-fertilization of ideas from a diverse number of disciplines, all of which share a common interest in discourse—be it prose comprehension and recall, dialogue analysis, text grammar construction, computer simulation of natural language, cross-cultural comparisons of communicative competence or other related topics. The problems posed by multisentence contexts and the methods required to investigate them, while not always unique to discourse, are still sufficiently distinct as to benefit from the organized model of scientific interaction made possible by this series.

Scholars working in the discourse area from the perspective of sociolinguistics, psycholinguistics, ethnomethodology and the sociology of language, educational psychology (e.g., teacher-student interaction), the philosophy of language, computational linguistics, and related subareas are invited to submit manuscripts of monograph or book length to the series editor. Edited collections of original papers resulting from conferences will also be considered.

## *Volumes in the Series*

- Vol. I. Discourse Production and Comprehension. Roy O. Freedle (Ed.), 1977.
- Vol. II. New Directions in Discourse Processing. Roy O. Freedle (Ed.), 1979.
- Vol. III. The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production. Wallace L. Chafe (Ed.), 1980.
- Vol. IV. Text, Discourse, and Process: Toward a Multidisciplinary Science of Texts. Robert de Beaugrande, 1980.
- Vol. V. Ethnography and Language in Educational Settings. Judith Green & Cynthia Wallat (Eds.), 1981.
- Vol. VI. Latino Language and Communicative Behavior. Richard P. Duran (Ed.), 1981.
- Vol. VII. Narrative, Literacy and Face in Interethnic Communication. Ron Scollon & Suzanne Scollon, 1981.
- Vol. VIII. Linguistics and the Professions. Robert J. DiPietro (Ed.), 1982.
- Vol. IX. Spoken and Written Language: Exploring Orality and Literacy. Deborah Tannen (Ed.), 1982.

## vi Preface to the Series

- Vol. X. Developmental Issues in Discourse. Jonathan Fine & Roy O. Freedle (Eds.), 1983.
- Vol. XI. Test Production: Toward a Science of Composition. Robert de Beaugrande, 1984.
- Vol. XII. Coherence in Spoken and Written Discourse. Deborah Tannen (Ed.), 1984.
- Vol. XIII. The Development of Oral and Written Language in Social Contexts. Anthony D. Pellegrini & Thomas D. Yawkey (Eds.), 1984.
- Vol. XIV. What People Say They Do With Words. Jef Verschueren, 1985.
- Vol. XV. Systemic Perspectives on Discourse, Volume 1: Selected Theoretical Papers from the 9th International Systemic Workshop. James D. Benson & William S. Greaves (Eds.), 1985.
- Vol. XVI. Systemic Perspectives on Discourse, Volume 2: Selected Applied Papers from the 9th International Systemic Workshop. James D. Benson & William S. Greaves (Eds.), 1985.
- Vol. XVII. Structures and Procedures of Implicit Knowledge. Arthur C. Graesser & Leslie F. Clark, 1985.
- Vol. XVIII. Contexts of Reading. Carolyn N. Hedley & Anthony N. Baratta (Eds.), 1985.
- Vol. XIX. Discourse and Institutional Authority: Medicine, Education, and Law. Sue Fisher & Alexandra Dundas Todd (Eds.), 1986.
- Vol. XX. Evidentiality: The Linguistic Coding of Epistemology. Wallace Chafe & Johanna Nichols (Eds.), 1986.
- Vol. XXI. The Acquisition of Literacy: Ethnographic Perspectives. Bambi B. Schieffelin & Perry Gilmore (Eds.), 1986.
- Vol. XXII. Cognitive and Linguistic Analyses of Test Performance. Roy O. Freedle & Richard P. Duran (Eds.), 1987.
- Vol. XXIII. Linguistic Action: Some Empirical-Conceptual Studies. Jef Verschueren (Ed.), 1987.
- Vol. XXIV. Text and Epistemology. William Frawley, 1987.
- Vol. XXV. Second Language Discourse: A Textbook of Current Research. Jonathan Fine (Ed.), 1988.
- Vol. XXVI. Systemic Functional Approaches to Discourse. James C. Benson & William S. Greaves (Eds.), 1989.
- Vol. XXVII. Language Development: Learning Language, Learning Culture. Ruqaiya Hasan & James Martin (Eds.), 1989.
- Vol. XXVIII. Multiple Perspective Analyses of Classroom Discourse. Judith L. Green & Judith O. Harker (Eds.), 1988.
- Vol. XXIX. Linguistics in Context: Connecting Observation and Understanding. Deborah Tannen (Ed.), 1988.
- Vol. XXX. Gender and Discourse: The Power of Talk. Alexandra D. Todd & Sue Fisher (Eds.), 1988.
- Vol. XXXI. Cross-Cultural Pragmatics: Requests and Apologies. Shoshana Blum-Kulka, Julianne House, & Gabriele Kasper (Eds.), 1989.
- Vol. XXXII. Collegial Discourse. Allen D. Grimshaw, 1989.
- Vol. XXXIII. Task, Talk, and Text in the Operating Room: A Study in Medical Discourse. Catherine Johnson Pettinari, 1988.
- Vol. XXXIV. The Presence of Thought. Introspective Accounts of Reading and Writing. Marilyn S. Sternglass, 1988.
- Vol. XXXV. Japanese Conversation: Self-contextualization through Structure and Interactional Management. Senko Kumiya Maynard, 1989.

# Table of Contents

Preface to the Series	v
About the Institute	ix
Overview	xi
1 Critical Thinking Assessment: Status, Issues, Needs <i>Robert H. Ennis and Stephen P. Norris</i>	1
2 Measuring Skills in Problem Solving <i>Norman Frederiksen</i>	43
3 Steps Toward a More Naturalistic Approach to the Assessment of the Reading Process <i>Peter Johnston</i>	92
4 The Sentence Verification Technique: A New Direction in the Assessment of Reading Comprehension <i>James M. Royer</i>	144
5 Towards a Transactive Theory of the Reading Process and Research in Evaluation <i>Audrey N. Grant</i>	192
6 Assessing Knowledge Structures in Mathematics: A Cognitive Science Perspective <i>Sandra P. Marshall</i>	241
Author Index	275
Subject Index	282

# Critical Thinking Assessment: Status, Issues, Needs\*

**Robert H. Ennis**

University of Illinois  
Champaign, IL

**Stephen P. Norris**

Memorial University of Newfoundland  
St. John's, CANADA

*Critical thinking* as defined by Ennis is reasonable and reflective thinking that is focused on deciding what to believe or do, and is an important part of problem solving. It includes dispositions that a person should have in deciding what to believe or do, and abilities in clarifying questions, supporting judgments, making inferences, and selecting strategies and tactics.

This chapter reviews distinctions between or among testing and evaluation, aspect specific and comprehensive thinking tests, scoring methods, subject-matter and general-knowledge-based tests, abilities and dispositions, critical thinking and reasoning, and critical and creative thinking. Issues related to these distinctions are addressed and include problems in defining critical thinking, measuring critical thinking with machine scorable tests, accounting for the influence of different backgrounds and ideologies on critical thinking, and validating critical thinking tests.

## INTRODUCTION

Although critical thinking instruction is receiving much emphasis these days, our ways of assessing critical thinking are few and fairly primitive. This is partly because critical thinking testing and other procedures for evaluation have not received adequate attention over the past 50 years. Our education system is now paying the price, be

---

<sup>1</sup> We are indebted to James Algina, Charles Blatz, Michelle Commeyras, and Mellen Kennedy, for their scrutiny of and comments on an earlier draft. For support we are indebted to the Institute for Student Assessment and Evaluation of the College of Education of the University of Florida, to the Spencer Foundation, to the Center for Advanced Study in the Behavioral Sciences, and to the Social Sciences and Humanities Research Council of Canada.

cause people are asking many more questions than can be answered with the instruments, methods, and resources currently at our disposal.

Typical questions include the following: When are students ready to begin learning to control variables in planning investigations? How well can students infer in different subject matter areas and in problems of their everyday lives? How does ability to appraise observation statements relate to ability to make observations proficiently? How can we tell whether a given course or curriculum promotes the disposition to be well informed and to look at things from others' points of view, or, more broadly, to be open-minded? We cannot adequately answer such questions with our current instruments and procedures.

In this chapter, after offering a few distinctions, we shall present an annotated list of English-language critical thinking tests currently available. We shall then consider some main issues in the area of critical thinking evaluation, including what critical thinking is, whether machine-scorable tests can validly test for critical thinking, whether we should penalize a test taker for having and expressing an ideology different from that of the test writer, how we should deal with the fact that different beliefs about the way the world works and different levels of sophistication sometimes lead to different but justifiable answers, how we can test for critical thinking dispositions, and how we can tell for what an alleged critical thinking test is testing. (A less theoretical but more detailed consideration and application of these topics—written for practitioners—may be found in our *Evaluating Critical Thinking*, Norris & Ennis, 1989.) Finally we shall suggest some research and development work that needs to be done.

## **DISTINCTIONS FOR CLASSIFYING CRITICAL THINKING ASSESSMENT PROCEDURES**

In describing the current critical thinking assessment situation, we find it important to employ the following distinctions: testing versus other forms of evaluation; aspect-specific versus more comprehensive evaluation; machine-scorability versus scorability only by a knowledgeable person; traditional subject-matter content versus general-knowledge content; critical thinking abilities versus dispositions; critical thinking versus reasoning; and critical versus creative thinking.

Some of these distinctions are well established. Others are new. Some are unique to the field of critical thinking.

### **Testing vs. Other Forms of Evaluation**

As the term implies, *evaluation* is a process of determining the value, worth, or quality of something. Determining the quality of people's

critical thinking requires gathering information on their thinking. Testing is one way of gathering this information, but there are others, such as nonmanipulative observation of various sorts, interviews, and questionnaires. The distinction between testing and the broader category, evaluation, is important, because it reminds us that we are not limited to testing.

### **Aspect-Specific vs. More Comprehensive Critical Thinking Assessment**

Assessment is aspect-specific if it assesses only one aspect of critical thinking, such as ability to identify assumptions, to reason inductively, or to define. It is comprehensive to the extent that it assesses a number of significant aspects. We know of no test that purports to test for all aspects of critical thinking.

### **Machine-Scorability vs. Scorability Only by a Knowledgeable Person**

The most common critical thinking tests are in a multiple-choice format, which makes them easy to score by machine. Interviews, naturalistic observation, and essay tests, either with or without a standardized scoring procedure, require a knowledgeable person. They cannot really be scored by machines, although a machine can be directed to count occurrences of things that have in the past correlated with good performance. One obvious advantage of machine-scorable tests is the economy achieved by their use in large-scale testing. A disadvantage, we shall argue, lies in the fact that a student might do good thinking, employing assumptions different from those made by the test-maker, and produce an unkeyed answer which is then marked wrong.

### **Subject Specific vs. General-Knowledge-Based Assessment**

Critical thinking assessment can employ content material that is supposedly general knowledge, or can employ material from a subject area that is supposedly not general knowledge. Critical thinking assessment within some standard school subject area would be subject-specific. (Parts of the College Board's Advanced Placement ["AP"] tests would serve as examples.) Assessment employing only content that is supposedly general knowledge would be general-knowledge-based.

The categories are not mutually exclusive. Subject-specific critical thinking assessment can have content that is so widely known that it is general knowledge. This can easily happen in critical thinking tests

that use social studies content, for example *The Ennis-Weir Critical Thinking Essay Test* (Ennis & Weir, 1985).

### **Abilities and Dispositions**

*Abilities* are capacities to do things, regardless of whether or not their possessors have a tendency to do those things. *Dispositions* are what creatures are inclined to do, that is, what they have a tendency to do.

A person can have an ability without the disposition to use it. So the disposition to use one's critical thinking abilities is one very important critical thinking disposition. Two others are the disposition to be well informed and the disposition to be open minded. We shall further exemplify this distinction between dispositions and abilities when discussing the nature of critical thinking.

### **Critical Thinking and Reasoning**

College Board (1983) and some others have suggested that reasoning is a basic competence. We believe that we are talking about roughly the same sort of thing they endorse under the label *reasoning* when we use the term *critical thinking*.

### **Critical and Creative Thinking**

Sometimes, critical thinking is sharply distinguished from creative thinking. As we use the term *critical thinking* (and as we think the term is most generally used), it refers to the thinking involved in reasonably and reflectively going about deciding what to believe or do. So defined, it incorporates fruitful creative thinking activities, such as formulating hypotheses, planning investigations, and conceiving of alternatives. These things make contributions to reasonably and reflectively deciding what to believe or do. Similarly, creative thinking, that is, thinking that leads to original products of high quality, could not proceed without "directing production and editing products" (Perkins, 1981) using the judgmental aspects of critical thinking (Bailin, 1987). If these definitions are used, critical and creative thinking overlap considerably.

## **EXISTING ENGLISH-LANGUAGE CRITICAL THINKING TESTS**

The preponderance of current subject-matter testing does not call for critical thinking within the subject matter (Linn, in press). For this

reason and because we have found no clear cases of readily available critical thinking tests intended to be specific to one particular subject, we have not used the category “subject-specific” in making a list of readily available critical thinking tests. We have categorized according to comprehensiveness vs. aspect-specificity, and machine-scorability vs. non-machine-scorability.

In the following chart, the date of publication, the source of the latest edition, intended grade ranges, and a content outline are given after the name of each test. The tests listed are general-knowledge-based critical thinking tests.

### *Available Critical Thinking Tests.*

#### 1. *Tests that are More or Less Comprehensive*

##### *Machine-Scorable*

*Basic Skills for Critical Thinking* (1979) (five forms) by Gary E. McCuen. Greenhaven Press, Inc., 577 Shoreview Park Rd., St. Paul, MN 55112. (Aimed at high school students.) Sections on source of information, primary and secondary sources, fact and opinion, prejudice and reasons, stereotypes, ethnocentrism, library card catalogue, and *Reader's Guide to Periodical Literature*.

*Cornell Critical Thinking Test, Level X* (1985) by Robert H. Ennis and Jason Millman. Midwest Publications, PO Box 448, Pacific Grove, CA 93950. (Aimed at grades 4–14.) Sections on induction, credibility, observation, deduction, and assumption identification.

*Cornell Critical Thinking Test, Level Z* (1985) by Robert H. Ennis and Jason Millman. Midwest Publications, PO Box 448, Pacific Grove, CA 93950. (Aimed at advanced or gifted high school students, college students, and other adults.) Sections on induction, credibility, prediction and experimental planning, fallacies (especially equivocation), deduction, definition, and assumption identification.

*Judgment: Deductive Logic and Assumption Recognition* (1971) by Edith Shaffer and Joann Steiger. Instructional Objectives Exchange, P.O. Box 24095, Los Angeles, CA 90024. (Aimed at grades 7–12.) Sections on deduction, assumption recognition, credibility, and dealing with emotionally loaded content.

*New Jersey Test of Reasoning Skills* (1983) developed by Virginia Shipman. IAPC, Test Division, Montclair State College, Upper Montclair, NJ 08043. (Aimed at grades 4–college.) A large number of items (about half) dealing with the classical syllogism and the meaning of categorical statements and smaller numbers of items



dealing with assumption identification, induction, good reasons, and distinguishing differences of kind and degree.

*Ross Test of Higher Cognitive Processes* (1976) by John D. Ross and Catherine M. Ross. Academic Therapy Publications, 20 Commercial Blvd., Novato, CA 94947. (Aimed at grades 4–college.) Sections on verbal analogies, deduction, assumption identification, word relationships, sentence sequencing, interpreting answers to questions, information sufficiency and relevance in mathematics problems, and analysis of attributes of complex stick figures.

*Test of Enquiry Skills* (1979) by Barry J. Fraser. Australian Council for Educational Research Limited, Frederick St., Hawthorn, Victoria 3122, Australia. (Aimed at grades 7–10.) Sections on using reference materials, interpreting and processing information, comprehension of science reading, experimental design, conclusions, and generalizations.

*Watson-Glaser Critical Thinking Appraisal* (1980) (two forms) by Goodwin Watson and Edward Maynard Glaser. The Psychological Corporation, 555 Academic Court, San Antonio, TX 78204. (Aimed at grade 9 through adulthood.) Sections on induction, assumption identification, deduction, following beyond a reasonable doubt, and argument evaluation.

#### *Not Machine-Scorable*

*The Ennis-Weir Critical Thinking Essay Test* (1985) by Robert H. Ennis and Eric Weir. Midwest Publications, PO Box 448, Pacific Grove, CA 93950. (Aimed at grades 7 through college). Also intended to be used as teaching material. Incorporates getting the point, seeing the reasons and assumptions, stating one's point, offering good reasons, seeing other possibilities (including other possible explanations), and responding to/avoiding equivocation, irrelevance, circularity, reversal of an if-then (or other conditional) relationship, overgeneralization, credibility problems, and the use of emotive language to persuade.

#### 2. *Aspect-Specific Critical Thinking Tests.* (All machine scorable)

##### *Aspect:-Deduction*

*The Cornell Class-Reasoning Test, Form X* (1964) by Robert H. Ennis, William L. Gardiner, Richard Morrow, Dieter Paulus, and Lucille Ringel. Illinois Critical Thinking Project, University of Illinois, 1310 S. 6th Street, Champaign, IL 61820. (Aimed at grades 4–14.) Seventy-two items, each containing a premise asserting a class relationship, such as “No A's are B's.” Each of twelve logical forms is tested by six items of varying types of content.