

Valencia, S. W., Pearson, P. D., Peters, C. W., & Wixson, K. K. (1989). Theory and practice in statewide reading assessment: Closing the gap. *Educational Leadership*, 47 (7), 57-63.

SHEILA W. VALENCIA, P. DAVID PEARSON, CHARLES W. PETERS, AND KAREN K. WIXSON

Theory and Practice in Statewide Reading Assessment: Closing the Gap

Illinois and Michigan have developed tests of reading comprehension that reflect current reading theory, with an emphasis on constructing meaning.

Educators are faced with a dilemma: our knowledge of reading processes and reading instruction is at odds with our assessment instruments. As a result, we run the risk of misinterpreting assessment data. If tests do not assess what we define as skilled reading, then they cannot adequately determine progress toward that goal. Thus, if we equate high scores on existing tests with good reading, we may be led to a false sense of security. Conversely, low scores may lead us to believe that students are not reading well when, by a more valid set of criteria, they are. Furthermore, tests have a powerful impact on curriculum and instruction; they influence classroom practice. In short, tests may be insensitive to growth in the abilities we most want to foster and may be misleading instruction.

Growth of Assessment and Research

Since the mid 1970s, statewide assessment has grown exponentially; now, in 1989, it has reached monumental proportions. At last count, 46 states

If we equate high scores on existing tests with good reading, we may be led to a false sense of security.

had mandated state-regulated testing; of these, all 46 require testing in reading. About half of these states have purchased a test or set of test items from standardized test publishers to serve that purpose (Afflerbach 1987, Selden 1988). Of the half that have developed their own reading assessments, most have modeled their tests after either existing norm-referenced standardized reading tests or the specific-skills, criterion-referenced tests that do not reflect current knowledge of reading processes.

During the same period, an explosion of research about reading—from such diverse domains as cognitive psychology, linguistics, and sociology—has produced a revolution in our views of the processes of reading and reading instruction. Over the past 15 years, educators have worked ear-

nestly to gain a deeper understanding of the reading process; they are now beginning to witness the widespread translation of this knowledge into instruction. Yet we undermine this progress by continuing to use tests that are at odds with reading theory and practice. The implications of this discrepancy pervade our entire decision-making system, from the broadest policy decisions at the federal level to the most specific and immediate decisions of teachers in classrooms. As a result, many reading researchers and reading educators have called for a change in the way we assess reading (Farr and Carey 1986, Valencia and Pearson 1987).

What We Know About the Reading Process

According to current theory, good readers actively build meaning for the texts they read. Meaning is not *in* the text to be extracted by readers through a series of analyses; instead, readers *build* meaning by bringing together knowledge they already possess and information gained from the text, and filtering that blend through the purposes they bring to the task.

In sports, as in reading, one can master component skills and still not play the game very well.

The process is fluid; it varies from one reading situation to another, depending on prior knowledge, motivation, interest, culture, task, setting, and, of course, text. The strategies readers use, the meaning they construct, the results of their efforts, and the personal satisfaction they feel also vary from situation to situation (Anderson et al. 1985, Wixson and Peters 1984).

This interactive view can be illustrated with an analogy to basketball. In sports, as in reading, one can master component skills and still not play the game very well. In basketball what matters is not how well one can per-

form isolated skills—dribbling, passing, shooting, guarding, rebounding—but how well one can orchestrate all the components. Furthermore, one does not become proficient at a sport simply by practicing component skills in isolation, but rather by learning when, why, and how to apply them, and by developing a positive attitude toward oneself as an athlete. Similarly, to become proficient in reading, one must possess the key skills; but one must also learn how to integrate the skills and to adapt them to purpose, text, and context, and one must develop a positive attitude toward reading.

From this perspective, we no longer define good readers as those who are able to decode precisely all the words on the page but rather as those who can build meaning by integrating their own knowledge with information presented by an author. Good readers are not those who demonstrate mastery of a series of isolated skills, but those who can apply important skills flexibly for a variety of purposes in a variety of authentic reading situations. Good readers are not those who can read short pieces of text and answer literal comprehension questions, but those who can read longer, more complete, authentic texts about a variety of topics and respond to them thoughtfully and critically. Finally, good readers are not those who simply read on demand in school, but those who have developed a disposition for reading and a commitment to its lifelong pursuit.

Inadequacies of Existing Measures

Figure 1 presents the discrepancies between an interactive view of reading and most reading assessments (Valencia and Pearson 1987). Note that none of the essential attributes of an interactive view is adequately represented in existing tests.

For example, most comprehension tests present students with short, specially constructed passages followed by multiple-choice questions that focus on details and explicit information. The passages often lack sufficient elaboration or context to help readers construct meaning. Further, these passages fail to approximate those that students encounter in their classes, and

Current views of reading suggest ...

prior knowledge is an important determinant of reading comprehension.

naturally occurring texts have topical and structural integrity.

inferential and critical reading are essential for constructing meaning.

reading requires the orchestration of many reading skills.

skilled readers apply metacognitive strategies to monitor and comprehend a variety of texts for a variety of purposes.

positive habits and attitudes affect reading achievement and are important goals of reading instruction.

skilled readers are fluent.

Yet reading assessments ...

fail to assess its impact on comprehension and try to mask its effects by using many short passages about unfamiliar topics.

use short pieces of texts that do not approximate the integrity found in most authentic texts.

rely predominantly on literal and sentence-level inferential comprehension items.

often fragment reading into isolated skills for item development and reporting.

seldom assess metacognitive strategies.

rarely include measures of these literacy experiences.

fail to assess fluency.

Fig 1. Contrasts between Reading Theory and Reading Assessments

their artificiality precludes questions that encourage the complex reasoning that is the essence of comprehension. Also, existing measures do not account for the impact that prior knowledge, metacognitive strategies, and dispositions have on comprehension.

More subtle dangers pertaining to curriculum and instruction are present as well. Just by virtue of their content, format, and reporting systems, tests deeply influence schools, classrooms, teachers, and students. Most educators agree that assessment exerts great influence on curriculum—that, in fact, it often *becomes* the curriculum.

The motivation of educators who “teach to the test” may not be simply to achieve higher scores. Some teachers sincerely believe that test publishers can better define what is important to teach than they can. But regardless of their motives, teachers look to tests to help them make curricular and instructional decisions; and so the conflict between current knowledge and assessment realities hampers efforts to change the curriculum. Clearly, reading assessment must be reconceptualized.

Reconceptualizing Reading Assessment

To be valid, assessments should provide students with multiple opportunities to apply their reading skills to a variety of real-life texts and tasks. Therefore, we advocate moving assessment away from traditional quantifiable tests to a portfolio system. This portfolio system should incorporate multiple indicators of expertise (e.g., comprehension, uses of literacy, metacognitive strategies), multiple measures in various contexts (e.g., reading different genres, reading for different purposes), and ongoing measures of this evolving expertise (e.g., repeated measures over time). Assessment must also take advantage of the many classroom indicators of achievement that cannot easily be reduced to paper-and-pencil measures.

Portfolios cannot simply replace standardized tests, however. Assessment must serve many masters: policymakers, state legislatures, school boards, superintendents, teachers, parents, and students. The need for group achievement data is obvious: agencies charged with

As a result of their schooling, students will be able to read, comprehend, interpret, evaluate, and use written material.

(State Goal for Language Arts, Illinois State Board of Education, 1985)

Knowledge/Skills Related to Language Arts Goal:

- A. Recognition, recall, and summarization of information from material read
- B. Questioning and predicting, giving rationales for each before, during, and after reading
- C. Reading for various purposes and identification of text to accomplish each purpose
- D. Sensitivity to difficulties of the text, requirements of the task, abilities, and motivation
- E. Using appropriate inferences to achieve a full understanding of text
- F. Integration of information from more than one text
- G. Justification and explanation of answers to questions about material read

Fig. 2. Summary of Illinois Reading Knowledge and Skills

monitoring large numbers of students must have, as one indicator, efficient and effective means to assess progress. Therefore, we must redirect large-scale assessment tests to align reading theory and practice.

New Approaches to Large-Scale Reading Assessment

The need for change in reading assessment is sharply felt at the national, state, and local levels. For the past three years, teams of reading researchers have been working with personnel from the Illinois State Board of Education and the Michigan Department of Education to translate current reading theory into reading assessment practices (Wixson et al. 1987, Valencia and Pearson 1986). Both states began with

an interactive view of reading and a new set of state learning objectives developed by committees of teachers, school administrators, university faculty, and state board personnel that reflected concepts underlying an interactive perspective (see figs. 2 and 3). These objectives mark a clear departure from state goals of the 1970s, in which learning outcomes in reading were narrowly defined, tending to fragment and trivialize the reading process. The new focus is on larger concepts and curriculum strands, with an overriding emphasis on the process of constructing meaning. These objectives served as the framework for the construction of the assessments, thereby providing alignment among theory, curriculum goals, and assessment.

Although there are differences between the Illinois and the Michigan assessments in format and item specifications, their common theoretical basis has produced striking similarities. Both assessments devote special attention to the selection of narrative and expository texts that are representative of what students read in school. Both assessments consist of a primary test component, *constructing meaning*, and three supporting components: *topic familiarity*; *metacognitive knowledge and strategies*; and *reading attitudes, habits, and self-perceptions*. This model places the construction of meaning at the center of reading while at the same time recognizing the need to assess

For narrative texts, the text map resembles a story grammar: setting, problem, key events, resolution, characterization, and theme.

other factors to adequately interpret comprehension performance.

Text selection. Passages for the assessments are full-length narrative and expository texts drawn from children's magazines, trade books, reference books, and textbooks. The use of such texts provides several advantages over traditional assessment texts: (1) they are representative of the types, content, and structure of materials stu-

dents usually encounter in their classrooms; (2) they are more interesting and motivating for students to read; (3) they engage students in more complex reasoning and thinking about what they are reading; and (4) they permit the construction of more inferential and critical reading questions.

Item developers construct a text map for each selection. The map serves as a blueprint of the important

facts, ideas, and concepts in the passages to guide the development of the constructing-meaning questions. It also helps to highlight important organizational features of a selection (Wixson and Peters 1987). For narrative texts, the map resembles a story grammar: it includes the setting, problem, key events, resolution, characterization, and theme (see fig. 4). Other critical facets of narrative texts—author's craft, mood, and tone, for example—are included in the questioning if they are pertinent to the selection. For expository texts, a conceptual map reflects relationships among the main ideas, supporting ideas, and details, again with a focus on the key information needed to understand the passage (see fig. 5). Questions requiring the student to interpret information presented in adjunct aids (e.g., maps, charts, graphs), to detect the author's bias, or to distinguish fact from opinion may also be asked.

Constructing meaning. A deep understanding of the text requires the reader to understand the information explicitly stated in the text, to integrate information presented throughout the text, and to use that information or those concepts in ways that move beyond the immediate text. The purpose of the constructing-meaning items is to ensure that students are able to construct a holistic representation of the text, rather than to ensure that some predetermined set of skills is represented. The text drives the focus of the questioning, rather than the

Reading is the process of constructing meaning through the dynamic interaction among the reader's existing knowledge, the information suggested by the written language, and the context of the reading situation.

(Michigan Reading Association, 1984)

I. Constructive Meaning

A. Interactive Reading

1. ability to construct meaning under a variety of different reader, text, and contextual factors

B. Skills for Constructing Meaning

1. ability to use a variety of strategies to recognize words (e.g., predictions, context clues, phonics, and structural analysis)
2. ability to use contextual clues to aid vocabulary and concept development
3. ability to recall/recognize text-based information
4. ability to integrate information within a text
5. ability to integrate information from more than one text
6. ability to evaluate and react critically to what has been read
7. ability to construct a statement of a central purpose or theme
8. ability to identify major ideas/events and supporting information within and across texts

II. Knowledge About Reading

A. Goals and Purposes

1. knowing that the goal of reading is constructing meaning
2. knowing that reading is communication

B. READER-TEXT-CONTEXTUAL Factors that Influence Reading

1. knowing about READER characteristics that influence reading (e.g., prior knowledge, purpose, interest, attitudes, word recognition and comprehension strategies)
2. knowing about TEXT factors
3. knowing about CONTEXTUAL factors

C. Strategies

1. knowing a variety of strategies for identifying words (e.g., predictions, context clues, phonics, and structural analysis)
2. knowing a variety of strategies to aid comprehension (e.g., outlining, skimming, detailed reading)
3. knowing when and why to use certain word recognition and comprehension strategies
4. knowing that it is important to monitor and regulate comprehension
5. knowing that strategies are employed flexibly (i.e., they are differentiated by reader, text, contextual factors)

III. Attitudes and Self-Perceptions

- A. developing a positive attitude toward reading
- B. choosing to read often during free time both at home and in school
- C. choosing to read a variety of materials for a variety of purposes
- D. developing an understanding of one's competencies and limitations in reading
- E. developing a positive attitude (image) toward oneself as a reader

Fig. 3. Summary of the Michigan Reading Definition and Reading Objectives

For expository texts, a conceptual map reflects relationships among the main ideas, supporting ideas, and details.

The Dip

Themes:

Main idea: Tick learned it was good to share the dip with another kid.
 Abstract: Trust and sharing can lead to personal relationships.

Plot:

Problem: Tick wanted the dip to himself, but the other kid came.
 Resolution: The dip has new meaning for Tick, because he is sharing it with the other kid.

Setting:

Location: The dip, which is on a secluded spot near a river
 Relation to theme: Provides an isolated location Tick believes is his

Major Characters:

Name	Traits	Function
Tick	Tough, combative, sensitive, caring	To examine his defense of a possession
The other kid	Tough, firm, combative, caring	To challenge Tick's beliefs about his possessions

Major Events:

- Tick has a possession, the dip, which he believes is his own until the other kid appears.
- Tick tells the other kid to leave, but she refuses.
- Tick and the other kid fight.
- The dip is spoiled for Tick and becomes a battleground.
- A truce is declared, but Tick is still unhappy.
- An injured duck appears, and both try to catch it.
- Tick ignores the "boundary" in order to work with the other kid to save the duck.
- The duck dies.
- The other kid offers to leave the dip.
- Tick says she can stay.
- The dip becomes both of theirs.

Vocabulary:

Dip, pummeling, truce, feuding

From: J. Andrews, (1984), "The Dip," *Cricket Magazine*.

Fig. 4. Sample Story Map, Grade 7

questions driving the construction or the selection of the text. For example, given a passage about different methods of conserving energy, students might answer questions about one particular method of conservation, a comparison of the benefits of two different methods, or a projection of which method might be the most effective in a new environment. This in-depth thinking about a text distinguishes these assessments from others.

Topic familiarity. Before reading the test passages, students answer questions about their prior knowledge of the important concepts that underlie the central ideas of each passage. The purpose of this section is not to assess students' knowledge of obscure

or unfamiliar terms or to assess general knowledge of many topics; it is to estimate the breadth and depth of students' knowledge about passage-specific concepts. For example, given the passage on conserving energy, students might be asked to identify key attributes and examples of the concept of conservation or to identify important ideas that might be found in a text on this topic. Students are not penalized or compensated for their lack of prior knowledge. Instead, a student's level of prior knowledge for a particular passage is used to help interpret the constructing meaning score for that passage.

Metacognitive knowledge. The metacognitive components of these tests measure students' knowledge of how to

Kay Hachten Educational Videotapes

PRESENTS



PRACTICE IN PEER COACHING

A series designed to adapt to your staff development needs

Five VHS videotape programs feature classroom lessons and conferences conducted by peer coaches. The programs capture actual teaching episodes and unrehearsed conferences. Educational specialist Kay Hachten provides expert analysis of lessons and conferences. In each program, a different aspect of observation, lesson analysis, conferencing, or communication is discussed. Each tape is accompanied by a training manual which may be duplicated.

- Program 1:**
Observing and Script-Taping Lessons, 6 lesson segments; 52 min.
- Program 2:**
Language Lesson and Conference, 5th grade; 65 min.
- Program 3:**
Writing Lesson and Conference, 3rd grade; 50 min.
- Program 4:**
Reading Lesson and Conference, 2nd grade; 71 min.
- Program 5:**
Music Lesson and Conference, 6th grade; 62 min.

All tapes are \$210 each plus \$3.00 shipping and handling. **SAVE** by ordering the complete series of 5 tapes for \$200 each plus \$5.00 shipping and handling.
 20 minute preview tape available (\$10 - 30 day return)

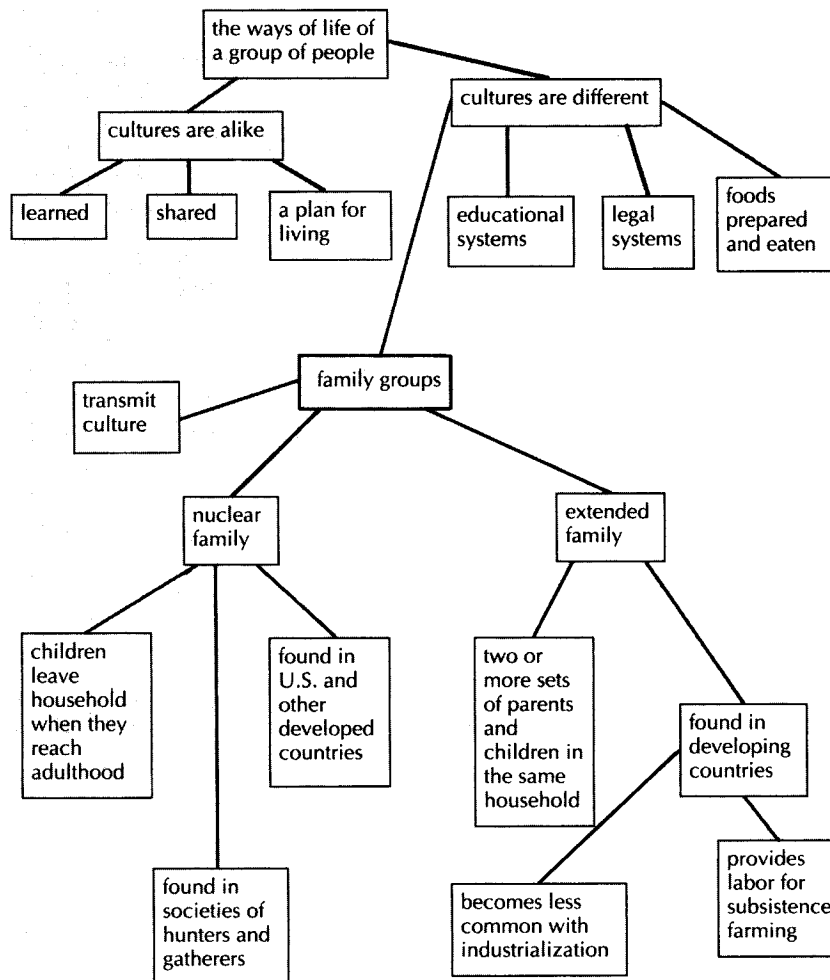
TECHNOLOGY AND TRAINING AT ITS FINEST

Be watching for our upcoming series which features Peer Coaching at the secondary level.

Kay Hachten Educational Videotapes, Inc.

P.O. Box 26724, Santa Ana, CA 92799
(714) 557-7770

Cultures and Families



Theme: The culture and the family influence each other and are shaped by many factors.

Main idea: Families in all cultures are similar in some ways, but are also very different in other ways.

From: Philip Bacon et al., (1983), *Our World Today* (Newton, Mass.: Allyn & Bacon, Inc.).

Fig. 5. Sample Expository Map, Grade 6

manage their reading strategies to meet the demands of various texts and tasks. These items are constructed to accompany the actual text passages included on the assessment. So rather than assessing students' generic reading strategies, the items are tied to a direct application in an actual text. For example, readers may be asked how the text is organized; what the purposes of charts, graphs, and illustrations are; or what strategies

would be helpful when answering a specific question or encountering an unknown phrase. As with the topic familiarity component, the purpose here is to help interpret comprehension performance rather than to create an independent subtest score.

Habits and attitudes. The final components of these tests assess literacy habits, attitudes, and self-perceptions related to reading performance. Some

Teachers should not have to set aside good instruction to prepare students to take a test; instead, good instruction itself should be the best preparation.

items ask students about their interest in the selection and their perceptions of the difficulty of the text and the questions. Others ask students about their reading and writing behaviors in school and the various ways in which they use reading and writing in their lives.

A Step Forward

Changing statewide reading assessments has required patience and courage. State board staff have worked diligently to develop these assessment tools and to convey to other educators the implications of an interactive view of reading. Those of us who helped design the assessments have had to make difficult decisions to balance our desire to operationalize complex concepts in valid ways against the need for efficiency. But we have taken a step forward. We have begun trying to narrow the gap between what we know about how people read and how we assess reading. Rather than providing a definitive solution to the reading assessment dilemma, we have demonstrated that theoretically and psychometrically sound alternatives are being developed.

Assessment based upon our best knowledge about learning to read sends a message in support of sound up-to-date instructional practice. Teachers should not have to set aside good instruction to prepare students to take a test; instead, good instruction itself should be the best preparation. □

References

Afflerbach, P. (1987). "The Statewide Assessment of Reading." Paper presented at the annual meeting of the American Education Research Association, Washington, D.C.

Anderson, R.C., E.H. Hiebert, J.A. Scott, and I.A. Wilkenson. (1985). *Becoming A Nation of Readers*. Urbana, Ill.: University of Illinois Center for the Study of Reading.

Farr, R., and R.F. Carey. (1986). *Reading: What Can Be Measured?* Newark, Del.: International Reading Association.

Selden, R. (1988). Paper presented at the annual meeting of the International Reading Association Conference, Toronto.

Valencia, S.W., and P.D. Pearson. (1986). "Reading Assessment Initiative in the State of Illinois—1985-1986." Springfield, Ill: Illinois State Board of Education.

Valencia, S.W., and P.D. Pearson. (1987). "Reading Assessment: Time for a Change."

The Reading Teacher 40: 726-732.

Wixson, K.K., and C.W. Peters. (1984). "Reading Redefined: A Michigan Reading Association Position Paper." *Michigan Reading Journal* 17: 4-7.

Wixson, K.K., and C.W. Peters. (1987). "Comprehension Assessment: Implementing an Interactive View of Reading." *Educational Psychologist* 22, 3 and 4: 333-356.

Wixson, K.K., C.W. Peters, E.M. Weber, and E.D. Roeber. (1987). "New Directions in Statewide Reading Assessment." *The Reading Teacher* 40: 726-732.

Recommended Readings

Cohen, M. (1988). "Designing State Assessment Systems." *Educational Leadership* 69, 8: 583-588.

Duckett, W. (1988). "Peer, An Interview with Stuart Rankin." *Pbi Delta Kappan* 69, 8: 605-608.

Peters, C.W., K.K. Wixson, S.W. Valencia, and P.D. Pearson. (In press). "Changing

Statewide Reading Assessment: A Case Study of the Illinois and the Michigan State-wide Assessment Projects." Berkeley, Calif.: Commission on Testing and Public Policy, University of California at Berkeley.

Valencia, S.W., and P.D. Pearson. (1988). "Principle for Classroom Comprehension Assessment." *Remedial and Special Education* 9, 1: 26-35.

Sheila W. Valencia is Assistant Professor, University of Washington, 122 Miller Hall DQ-12, Seattle, WA 98195. **P. David Pearson** is Professor and Co-Director of the Center for the Study of Reading, 51 Gerty Dr., Champaign, IL 61820. **Charles W. Peters** is Reading Consultant, Oakland Schools, 2100 Pontiac Lake Rd., Pontiac, MI 48054. **Karen K. Wixson** is Associate Professor of Education, University of Michigan, 1008 School of Education, Ann Arbor, MI 48109.

Critical and Creative Thinking in the Classroom

1989 Summer Institute

This institute gives teachers and other educational practitioners an opportunity to work with leaders in the field of teaching for critical and creative thinking. Participants will gain an understanding of the nature of critical and creative thinking, learn classroom techniques for teaching and assessing it, and develop curriculum materials and teaching strategies for infusing it into mainstream instruction. Most activities carry graduate credit. The institute offers one-day practitioner seminars, week-long mini-courses, and three-week curriculum-building courses, as well as a special seminar for administrators. Participants may enroll for one, two, or three weeks. Residential accommodations are available.

The staff will consist of active researchers, classroom teachers, and school administrators who have made major contributions to the field. Among them are: John Barell, Montclair State College; Stephen Norris, Memorial University of Newfoundland; David Perkins, Harvard University; Vincent Ruggiero, State University of New York; Robert Swartz, University of Massachusetts at Boston; Karen Warren, Macalester College; Mary Anne Wolff, North Reading (MA) Public Schools.

July 10-28, 1989
Boston, Massachusetts

Fourth National Summer Institute for
Teachers, School Administrators, and
other Educational Practitioners

Sponsored by the Critical and
Creative Thinking Program,
University of Massachusetts
at Boston

UMass Boston



For further information, contact:
Kathleen FitzPatrick
University of Massachusetts at Boston
Division of Continuing Education
Harbor Campus
Boston, MA 02125-3393
(617) 929-8814

This advertisement appears at no expense to the
Commonwealth of Massachusetts.