Emergent Literacy: A Comparison of Formal and Informal Assessment Methods

Rebecca Harlin  
*State University of New York at Buffalo*

Sally Lipa  
*State University of New York at Geneseo*

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons

Part of the Education Commons

Recommended Citation
Emergent Literacy: A Comparison of Formal and Informal Assessment Methods

Rebecca Harlin
Sally Lipa

Young children enter school with individual and divergent literacy experiences. Considerable disagreement exists concerning how best to assess children's competence and utilize the results of instruments designed to provide this information for educators.

Since large groups of children need to be screened prior to formal reading instruction, group standardized tests are presently used to differentiate those children in need of preventive intervention from those in need of more formal reading programs. As the age level for school entry becomes younger there is a strong tendency to use the same measures for assessing four year olds as for five and six year olds, and the same measures for an identified language delayed population as for a normal population. This policy ignores differences in literacy development and the requisite program opportunities that may be essential for younger and high-risk children.

Standardized reading readiness tests, used in a diagnostic manner are known to "drive the curriculum." These tests which assess skills such as auditory discrimination, letter identification, letter-sound association, following directions and copying letters result in a readiness program designed to
master these skills. They do not include items which reflect emergent literacy research (Day and Day, 1986). Instead, the tasks provide the teacher with fragmented data such as how well the child copies geometric forms and matches initial sounds to representative pictures, but not how well the child understands the reading process. Recent editions of these tests show that no significant alterations have been made to measure children's conscious awareness of the form, use and function of written language prior to formal literacy instruction.

The predictable value of standardized reading readiness tests has long been questioned by researchers. For example, Karlin (1971) summarizes various reports of the predictive validity of reading readiness tests and reports the correlations range from .40 to .60 with later reading achievement. Coltheart (1979) and Glazzard (1977) show that variables tapped by reading readiness tests are not predictively successful nor theoretically informative. Other researchers account for the variable predictive efficiency of such tests by noting that the variables, either predictor or criterion, have been conceptualized in very general or restrictive terms (Feshbach, Adelman and Fuller, 1977). Although it has been known for some time that the best predictors of reading achievement are those tests that most closely resemble tasks involved in reading (Karlin, 1971), schools continue to use group standardized measures to identify high-risk children, establish individual and group baseline information and make curriculum decisions (Hiebert, 1986).

In recent years efforts have increased toward the construction of more efficient prediction instruments, i.e., instruments in which individual differences in acquiring emergent literacy can be observed. A better understanding of emergent
literacy has heightened awareness of the need for early identification of at-risk children as well as providing the means for fostering literacy.

Among the techniques which have contributed to our knowledge of emergent literacy behavior are the Concepts About Print Test (Clay, 1979), the Book Handling Task (Goodman and Altwerger, 1981), the Rhyme Reading Task (Morris, 1983), and the Metalinguistic Inventory (Evans, Taylor, and Blum, 1979). While none shares a common task format, each of these instruments measures a discrete aspect of literacy behavior. All of these instruments employ a concrete stimulus for the child, examine print-related situations, and measure aspects of emergent literacy behaviors found to be related to reading success. The data from such instruments provide teachers with reliable information for grouping children, planning instruction, and reporting children’s progress to parents and administrators. In spite of research results which show their effectiveness, informal assessment tasks are not commonly used as screening instruments at the preschool or primary levels.

The purpose of this investigation was to examine a number of literacy measures in light of their task demands, and their contribution to a composite picture of a child’s literacy development. Answers to the following questions were sought:

1. Does an informal measure of print awareness, the Concepts About Print Test (CAP) estimate the level of reading achievement a) for first graders, b) for high-risk primary grade students?

2. Does a standardized reading readiness battery, the Metropolitan Readiness Test (MRT) estimate the level of reading achievement a) for first graders, b) for high-risk students?
3. Does the combination of effective predictors of literacy development differ a) for first graders, b) for high-risk students?

Thus, the major focus was on comparing the effectiveness of informal and standardized readiness measures in assessing the literacy development of both normal first graders and high-risk, primary grade children.

METHOD

Sample

For the purposes of this study, 87 subjects from primary grade classrooms were chosen — four classes of first grade students and three classes of high-risk students. Selection of both groups of subjects involved intact classrooms. The 60 first grade students were from a suburban school in upstate New York. The 27 high-risk primary grade children were from three intact classes of language-delayed students — one each of six year olds, seven year olds, and eight year olds from a suburban school serving only language-delayed children.

Instruments

Three informal instruments were used to evaluate the literacy development of the subjects. These instruments included the Concepts About Print Test, the Writing Vocabulary Test, and the Sentence Dictation Test. In October, each of these instruments was administered individually to the first graders following the procedures outlined in Clay’s The Early Detection of Reading Difficulties. The high-risk children were given the Concepts About Print Test and the Writing Vocabulary Test by one of the investigators. Scoring for each item was completed following the guidelines outlined by Clay.
The Concepts About Print Test (CAP) was selected for use in this study since it allowed the researchers to obtain information about the children's understanding of print concepts in a most efficient manner. A 24-item checklist of questions was asked while the storybook, Sand, was read to the child. Among the concepts assessed were those of letter, word, print direction, and uses of punctuation. The Writing Vocabulary Test was chosen as an inventory of the words of which each subject has control, i.e., can spell correctly. This instrument consists of an open-ended task in which children are given ten minutes to write all the words they know, starting with their own name. As an evaluation of the child's ability to analyze and record the phonemes in individual words, the Sentence Dictation Test was administered. Two simple sentences were read to the subject, then repeated, one word at a time, as the child attempted to write them. Each of these instruments has been normed and used with primary grade children. It was felt that since writing ability and reading ability both result from experiences with letters, words, and stories, the three tasks provided an opportunity for children to show what they have learned about written language.

Readiness is commonly evaluated using a paper and pencil test. Thus, all subjects were given a formal assessment battery of readiness, the Metropolitan Readiness Tests. For the first grade subjects, the MRT, Level II, was group-administered by their classroom teachers in May of their kindergarten year. Subjects' scores were obtained from the school's printout. Following the guidelines for handicapped children, the MRT, Level I, was administered individually to each language-delayed subject by one of the researchers. Each high-risk subject's battery was hand-scored, following the directions in the MRT administration handbook. Local norms were established for this out-of-level test. In addition,
each subject was also given the Peabody Picture Vocabulary Test as a measure of language capacity, a frequently used predictor of reading achievement.

For the first grade subjects, the Stanford Achievement Test, Primary I, was administered by their classroom teachers in May. This battery was used as the measure of reading achievement. All test booklets were machine-scored and results for each subject were obtained from the school’s printout. For the high-risk subjects, there were no comparable scores available because no standardized reading battery was administered in their school. The Letter Identification Test from Clay’s Diagnostic Survey was administered to each language-delayed subject as a measure of reading ability. This task was administered by one of the researchers following the procedures outlined in Clay’s Early Detection of Reading Difficulties.

RESULTS
Print awareness and reading achievement
To determine the relationship between print awareness and reading achievement, the scores on the CAP were compared to the subtest scores and the total reading scores on the SAT using a Pearson product-moment correlation. For the first graders, the CAP was found to have significant correlations (p. < .001) with the Word Recognition subtest (0.494), the Reading Comprehension subtest (0.512), the Word Study subtest (0.564), and the Total Reading Score (0.531). For the high-risk subjects, their CAP scores were compared to the scores on the Letter Identification Test (0.550). Thus, for both groups, normal and high-risk, the informal measure of print awareness, CAP, was found to estimate the level of reading achievement.
Readiness Battery and Reading Achievement

Does a standardized reading readiness battery, the *Metropolitan Readiness Tests*, predict the level of reading achievement? Through two different analyses, the answer to this question was found. Using the Pearson product-moment correlation, the coefficients for the May administration of the *MRT*, Level II (for first graders), with the *SAT* subtests were 0.570 for the *Word Recognition* subtest, 0.579 for *Reading Comprehension*, 0.564 for *Word Study*, and 0.554 for the Total Reading score. All correlations were significant at the .001 level. A linear regression analysis of *MRT* scores on the Total Reading scores was computed, resulting in an R-square equal to 0.306 (F=5.148, p<.001).

For the high-risk subjects, the results of the *MRT*, Level I, were compared to those of the *Letter Identification Test* using both a Pearson product-moment correlation and a linear regression analysis. The Pearson product-moment correlation for the *MRT* and the *Letter Identification Test* was 0.651 (p<.001). The linear regression analysis yielded an R-square equal to 0.423 (F=4.285, p<.001).

Comparison of Predictors

To compare the effectiveness of each informal and formal instrument in estimating reading achievement, Pearson product-moment correlations were calculated. For the first graders' *SAT* Total Reading Score, the strongest predictors were the *Sentence Dictation Test* (0.71) and the *Writing Vocabulary Test* (0.653). While the *Sentence Dictation Test* was also the strongest predictor for each of the three reading subtests, *Word Recognition* (0.709); *Reading Comprehension* (0.676); and *Word Study* (0.646), the *Writing Vocabulary Test* was a strong predictor of the *Word Study* subtest,
The standardized readiness battery, the *MRT*, was the second strongest predictor for the *Word Recognition* subtest (0.570) and for the *Reading Comprehension* subtest (0.579). The *Peabody Picture Vocabulary Test* was not a significant predictor of any of the reading achievement subtests. The *CAP* showed significant correlations with the Total Reading and subtest scores, but was not as strong as the other informal instruments (See Table 1).

<table>
<thead>
<tr>
<th>PPVT</th>
<th>CAP</th>
<th>DIC</th>
<th>WVC</th>
<th>SAT TOTAL</th>
<th>WORD REC</th>
<th>READ COMP</th>
<th>WORD STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRT</td>
<td>.381</td>
<td>.602</td>
<td>.748</td>
<td>.593</td>
<td>.561</td>
<td>.570</td>
<td>.579</td>
</tr>
<tr>
<td>PPVT</td>
<td>.402</td>
<td>.275</td>
<td>.281</td>
<td>.281</td>
<td>.193</td>
<td>.192</td>
<td>.226</td>
</tr>
<tr>
<td>CAP</td>
<td>.650</td>
<td>.700</td>
<td>.531</td>
<td>.494</td>
<td>.512</td>
<td>.564</td>
<td></td>
</tr>
<tr>
<td>DIC</td>
<td>.741</td>
<td>.710</td>
<td>.676</td>
<td>.646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WVC</td>
<td>.653</td>
<td>.523</td>
<td>.521</td>
<td>.584</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Next, multiple regression analyses were run to determine the effectiveness of different combinations of the informal literacy instruments in predicting the Total Reading scores for the first graders. The best combination of predictors was the *Sentence Dictation Test* and the *Writing Vocabulary Test* which resulted in an R-square of .528 (F=39.21, p<.001). The second best combination was the *CAP* and the *Writing Vocabulary* with an R-square of .513 (F=36.88, p<.001). While the weakest of the combinations was the *CAP* and the *Writing Vocabulary* with an R-square of .369 (F=20.52, p<.001), it was stronger than that of the six subtests that comprise the *MRT* battery (R-square = .306, F=5.148, p<.001). Thus, as predictors of first graders' reading achievement, the informal literacy measures were more effective than the formal readiness test battery.
TABLE 2
Intercorrelations Between Predictor Variables for High Risk Subjects

<table>
<thead>
<tr>
<th></th>
<th>MRT</th>
<th>WVC</th>
<th>CAP</th>
<th>LET</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT</td>
<td>.202</td>
<td>.497</td>
<td>.471</td>
<td>.472</td>
</tr>
<tr>
<td>MRT</td>
<td>.443</td>
<td>.582</td>
<td>.651</td>
<td></td>
</tr>
<tr>
<td>WVC</td>
<td></td>
<td>.631</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAP</td>
<td></td>
<td></td>
<td>.688</td>
<td></td>
</tr>
</tbody>
</table>

Similarly, in order to establish the strongest predictors of reading achievement for the high-risk students, Pearson product-moment correlations were calculated (See Table 2). For their Letter Recognition scores, the strongest predictors were the CAP (0.688) and the MRT (0.651). Next, multiple regression analyses were run to determine the effectiveness of different combinations of informal and formal instruments.

The best combination of predictors was the CAP and the MRT which resulted in an R-square of 0.549 (F=14.66, p<.001). While the second best combination was the CAP and the Writing Vocabulary Test with an R-square of 0.481 (F=11.16, p<.001), it was a stronger predictor than the six subtests of the MRT battery (R-square equal to 0.423, F=4.285, p<.001). Therefore, for both groups of students the informal literacy measures yielded the best results. The Peabody Picture Vocabulary Test was not an effective predictor of reading achievement for either group (See Tables 1 and 2).

DISCUSSION
While the effectiveness of the Concepts About Print Test for prediction of reading achievement in normal populations has been shown in past research (Day and Day, 1986; Freebody and Rust, 1985; Harlin, 1983), this is one of the first studies to support its efficiency as a predictor for high-risk
children, and to demonstrate the combined predictive qualities of the CAP, the Sentence Dictation, and the Writing Vocabulary Tests of Clay's Diagnostic Survey for first graders' reading achievement. While the standardized battery, the Metropolitan Readiness Tests, was found to be somewhat effective in identifying children at risk, its results were not sufficiently powerful to justify the time, effort, and expense of its administration. Therefore, this study's results support the use of informal, concrete tasks to assess the literacy development of both young and high-risk learners.

For teachers of primary grade children, as well as reading clinicians, the CAP has been shown to be an effective indicator of the child's knowledge and understanding of print concepts. Its ease of administration should recommend its use in the classroom as well as the reading clinic. The close correlation between the CAP and the measure of reading achievement, the SAT demonstrated the CAP's predictive qualities. As part of a preventive strategy, the CAP may be used to identify at-risk children early in the school year, thus facilitating intervention strategies. For the reading clinician, the appropriateness of the CAP as a diagnostic tool for young disabled readers has been shown.

Although the program emphasis for high-risk children was different from that of normal first graders, in that it emphasized letter name knowledge, both the CAP and the Writing Vocabulary Test were sensitive to changes in their literacy development. These children are at a stage of literacy development similar to the preschoolers studied by Mason (1982), who found that preschoolers acquired an increasing knowledge of letter names as they approached formal reading instruction. Thus, for this study, letter names were used as an indicator of print control.
The high correlations between the *Sentence Dictation Test* and the *SAT* show that not only is the informal task an accurate predictor of reading achievement, but also underscores the strong role writing plays in reading acquisition. Like the *CAP*, this instrument is easy to administer and interpret. Within a ten to fifteen minute period, a teacher can acquire powerful data about the child's ability to encode written language, a skill that is known to be related to reading success.

The high correlations between the *Writing Vocabulary Test* and the *Letter Identification Test* support the contention by many researchers (Goodman and Goodman, 1983; Springgate, 1983) that reading and writing are related tasks. Implications for including both "reading and writing" measures in pre-literacy assessments and program development for both normal and high-risk populations can be drawn from this information.

According to one theory of linguistic awareness, there is an interaction between reading acquisition and print awareness. As children learn to read, they become more sensitized to print (Ehri, 1979; Ryan, McNamara, and Kenney, 1977). While most of the children included in this study were not readers when pretested, the data revealed that they knew a great deal more than one would expect about the functions of print, and possessed a working knowledge of those functions as demonstrated by their performance on the writing tasks. This was true for the high-risk children who were not in a formal reading program, but who could write several words and name alphabet letters. Previous studies (Mason, 1980; Hiebert, 1979) revealed increasing reading readiness skills across normal preschool groups. The data from this study reveal a similar pattern for the high-risk group, but at a slower
rate and more limited progression than in the normal population. While Mason (1982) found emergent literacy behaviors occurring naturally among normally developing preschoolers, VanKleeck and Schuele (1987) suggest that emergent literacy behaviors do not develop naturally among language-delayed, high-risk children. Instead, they need active teaching both at home and school to foster the development of these concepts.

Analysis of the data obtained from the informal measures, CAP and Writing Vocabulary, and the Metropolitan Readiness Tests (MRT) reveal that while both are good predictors of letter name knowledge, the informal tests have several advantages.

The nature of the CAP measure allows the examiner to obtain individual profiles of children's print awareness. These profiles provide the teacher with usable information for instruction. For example, if a child does not know the left to right progression for reading, direct modeling and specific teaching can be incorporated in the child's program.

The Writing Vocabulary Test provides a measure of how children approach writing, their use of the alphabet and invented spelling patterns, and their formation of letters and letter sequences. Handicapped children should be offered the opportunity to develop a writing/reading relationship. Too often, these children are provided with oral instruction requiring oral feedback. Writing as a form of communication is often neglected because of predetermined notions that oral language and reading are precursors to writing. This relationship was not supported by the correlations of the Peabody Picture Vocabulary Test and the reading achievement measures.
For teachers, another advantage of informal measures is the opportunity to observe early emergence of oral/written language behaviors. In contrast, while administering group standardized reading readiness tests, which probe for the mastery of a skill, teachers have no indication of the strategies children are using to respond to those items. Informal measures provide a description of the emergent behavior and are more suitable indices for intervention needs. For example, the data from the CAP includes book handling tasks; basic concepts, e.g., front of book, first, last, etc.; identification of print containing the message; and reading terms such as letter and word. These data cannot be obtained from traditional standardized tests.

Rather than testing high-risk and normal children to determine if they are "ready" for formal reading, informal assessments should be periodically conducted to determine the extent to which emergent literacy behaviors are developing. This diagnostic information should not be used to "sort" children, but rather to enable meaningful intervention activities in which children have many experiences with print. No child should be deprived of print experiences. On the contrary, rich experiences with literature, shared reading, language experience stories, writing, and reading simple messages should be the program emphasis.

References


Rebecca Harlin is a faculty member at the State University of New York at Buffalo. Sally Lipa is a faculty member at the State University of New York at Geneseo. Requests for further information about the research described in this article should be accompanied by a SASE, and sent to Dr. Sally Lipa, Department of Elementary, Secondary, and Reading Education, State University College at Geneseo, Geneseo, NY 14454.

---

**Professional news**

A recent themed issue of the *Ohio Media Spectrum*, published by the Ohio Educational Library/Media Association has as its topic, "Reading: Key to the Past, Present, Future." Copies of the issue are available for $6 from OELMA, 40 South Third Street, Suite 230, Columbus, Ohio 43215.

The fifth World Conference on Computers in Education (WCCE/90), will be held in Sydney, Australia, July 9-13, 1990. Sponsors of WCCE/90 note that it will be "a conference for all aspects of computer-related education in all education environments." Those interested in receiving further information should write to: WCCE/90, PO Box 319, Darlinghurst, NSW, Australia 2010.

The thirty-fifth annual convention of the International Reading Association will be held in Atlanta, Georgia from May 6-11, 1990. The conference theme is "International Literacy Year: Celebration, Inspiration, Dedication," and the featured speaker at the opening general session will be Coretta Scott King.