
4-1-1988

Informal Reading Inventory Comprehension Questions: Are Classification Schemes Valid?

Rosie Webb Joels
University of Central Florida

Betty Anderson
University of Central Florida

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



Part of the Education Commons

Recommended Citation

Joels, R. W., & Anderson, B. (1988). Informal Reading Inventory Comprehension Questions: Are Classification Schemes Valid?. *Reading Horizons: A Journal of Literacy and Language Arts*, 28 (3). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol28/iss3/4

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons: A Journal of Literacy and Language Arts by an authorized editor of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



INFORMAL READING INVENTORY COMPREHENSION QUESTIONS: ARE CLASSIFICATION SCHEMES VALID?

ROSIE WEBB JOELS
BETTY ANDERSON
University of Central Florida

An area of controversy in reading diagnosis and instruction is that of the nature of comprehension. Questions have been raised about varied thinking skills employed during the comprehension of text. Is comprehension a unitary process? Or, is comprehension composed of different abilities based on the complexity of thought processes required?

This issue has been addressed by researchers, practitioners, and theorists. There appear to be two points that can be made after reading relevant literature: (1) respected professionals from a variety of related disciplines do not agree on the nature of reading comprehension and (2) in spite of this lack of agreement and the absence of data consistently substantiating differentiated comprehension skills, both instructional and testing materials include comprehension strategies based on predetermined classification schemes.

Some standardized testing materials, such as, the widely-used Comprehensive Tests of Basic Skills (1975) use different comprehension question types. Furthermore, informal diagnostic test developers suggest that comprehension skill profiles, that is, patterns of strengths and weaknesses across question types, be used in planning corrective or remedial instruction. The Informal Reading Inventory (Burns and Roe, 1985) and Classroom Reading Inventory (Silvaroli, 1982) are informal reading assessment tools that measure comprehension with items classified according to the developers' question types.

As previously stated, research on separate comprehension skills has presented conflicting findings. Davis (1944)

identified two major components of reading comprehension. In later studies (1968, 1972) Davis confirmed his earlier findings and reported that additional components of comprehension had been identified. Other researchers (Hunt, 1957; Spearritt, 1972) reported findings that, in their conclusions, substantiated Davis' results. Conflicting opinions, however, also based on research results, were presented by Harris (1948) and by Thorndike (1973-1974).

More recently, researchers have continued to disagree about reading tests' ability to measure, or even identify separate comprehension skills. While some researchers (Klein, 1979, 1980, 1981; Royer & Lynch, 1982) report that reading tests are not able to differentiate among specific comprehension skills, other investigators conclude that reading tests do have this capability (Blair and Raths, 1978; Crowell & Au, 1981).

In order to clarify the issue of comprehension question types, the authors of this report undertook an investigation to study comprehension testing in an original informal test. The JAT Reading Inventory (JAT) had been developed as part of a larger diagnostic instrument, The Progressive Reading Portfolio (Joels, Anderson, and Thompson, 1983), and consisted of graded passages across reading levels one through eight. Assessment procedures had been developed using arbitrarily classified comprehension questions and the investigators wanted to determine the extent to which these questions measure students' strengths and weaknesses in comprehension abilities.

Previous studies (Anderson and Joels, 1984-85 and in press) had provided data related to other aspects of test scoring and instructional level validity of the JAT. In addition, discriminant validity of the JAT as a diagnostic instrument appears to be established by the fact that it did yield widely varying results when administered to a heterogeneous population. The JAT established the following instructional reading levels for students in grade placements two through five: grade 2, below grade 1 to 4; grade 3, below grade 1 to 4; grade 4, below grade 1 to 6; and grade 5, below grade 1 to 8. Subjects obtained the following percentile rank ranges on a standardized test administered five months after the JAT testing: grade 2, 15-95; grade 3, 19-98; grade 4, 5-19; and

grade 5, 9-96. Thus, it appears that the JAT is sensitive to individual student's overall reading achievement.

However, one feature of the instrument that had not been investigated was the comprehension classification scheme in which the developers had categorized questions into four types: literal, inference, vocabulary, and application. In spite of a lack of validation evidence, the JAT's question types were defined in the administration procedures as assessing different thinking skills required by the reader in order to derive meaning from text. Literal question, for example, required that meaning be obtained from explicitly stated information. Inference called for the reader to infer or interpret meaning. Vocabulary questions necessitated correct definitions of words, often not contextually explained. Application questions involved the manipulation or evaluation of textual elements, for example, creatively altering or critically judging the text or solving a problem.

The Study

Students in grades two through five, selected randomly from two central Florida public elementary schools, were tested on Form A JAT passages by the authors and a trained graduate assistant. A total of 136 test protocols were analyzed for comprehension question responses.

TABLE I. Question Type Accuracy % on Each Passage

| Passage Level | N* | Literal | Inference | Vocabulary | Application |
|---------------|-----|---------|-----------|------------|-------------|
| 1 | 60 | 86 | 67 | 89 | 69 |
| 2 | 90 | 71 | 55 | 62 | 76 |
| 3 | 105 | 84 | 94 | 47 | 82 |
| 4 | 91 | 47 | 67 | 79 | 66 |
| 5 | 58 | 51 | 55 | 52 | 47 |
| 6 | 26 | 70 | 52 | 38 | 59 |
| 7 | 10 | 62 | 82 | 42 | 85 |
| 8 | 6 | 44 | 58 | 33 | 70 |

(*N = number of students tested on that passage level.)

Table 1 shows the question type accuracy means that were computed for all students tested on each passage level. Differences among question types are apparent across the passage levels. However, informal inventories are not designed from group interpretation. Results, therefore, are not intended to be used to derive group performance indicators, such as, means, ranges, or medians.

Informal reading inventories are designed for individual diagnostic testing and for interpretation relevant for subsequent instructional planning. Most important, therefore, are data on individual student performances. If meaningful differences are found across assessed constructs, corrective or remedial teaching can more effectively be planned and implemented.

There were 38 students whose instructional reading levels were determined to be two to four levels below their current grade placements. In this instrument, instructional reading level is defined as the highest passage reading level at which a student concurrently obtains a minimum of 95% word recognition accuracy and a minimum of 70% comprehension accuracy. Using these obtained scores, it appears that approximately one-third of the student group tested would be considered candidates for corrective or remedial reading instruction. When these individuals' scores on question types are studied, one does discern strengths and weaknesses across the comprehension question types.

Specifically, 31 of the 38 poor readers scored below 60% accuracy in at least one of the question type categories. Similarly, only five of these same students failed to achieve accuracy equal to or exceeding 75% in at least one of the question types. To reiterate, almost all these individual poor readers showed marked differences in their abilities to obtain meaning as measured by the classified questions.

Six representative students' performances across the four question types are shown in Table 2. Students are listed according to their instructional reading level/grade placement discrepancies. Scores for two students who had an achievement/placement discrepancy of -2 are depicted. The first of these students had a comprehension accuracy of 92% on the literal questions across all levels on which

Table 2
Individual Representative Students'
Question Type Accuracy Percentage

| Grade Placement/ Instruc- tional Level Discrepancy | Literal | Inference | Vocab ulary | Appli- cation | Differ ence |
|--|---------|-----------|----------------|------------------|----------------|
| -2 | 92 | 46 | 42 | 56 | 50 |
| -2 | 42 | 100 | 67 | 67 | 58 |
| -3 | 81 | 85 | 61 | 47 | 38 |
| -3 | 61 | 57 | 61 | 94 | 37 |
| -4 | 56 | 78 | 50 | 100 | 50 |
| -4 | 90 | 58 | 50 | 72 | 40 |

tests were administered and a 46% accuracy for the inference questions similarly assessed; these scores resulted in a difference of 50 points between this student's strongest and weakest (that is, highest and lowest) comprehension question types. The other representative student with an achievement/placement discrepancy of -2 had a difference of 58 between highest and lowest obtained scores on the comprehension questions. This student had the highest score on inference and the lowest on literal questions, however.

As further illustrated in Table 2, two representative students with achievement/placement discrepancies of three levels (i.e., -3) had question type performance differences of 38 and 37. Students with achievement/placement differences of -4 had comprehension type accuracy differences of 50 and 40.

These types of data are useful to classroom teachers who administer informal reading inventories in order that results can be used for instructional planning. A logical follow-up study is needed to investigate the effectiveness of instruction in the specific comprehension skill weaknesses. An experimental study would indicate the usefulness of diagnostic testing as a prelude to improving a student's reading comprehension.

Conclusion

Discriminant validity of the JAT as a diagnostic instrument appears to be established, at least in the extent to which it yields varying results for poor readers across the comprehension question types. Classification question types appear to be valid assessment techniques in determining poor achievers' specific comprehension skills' strengths and weaknesses. This conclusion is based on the apparent differential responsiveness of the question types to students who have special needs in reading instruction. The appropriateness of these patterns of comprehension skills as aids to planning effective corrective or remedial work cannot be determined in a study of this nature, requiring further investigation in an experimental design.

REFERENCES

- Anderson, B.B. & Joels, R.W. (1984-5). The informal reading inventory instructional level. In D. Read & G. Bonning (Eds.), Striving for reading excellence through research and study. Orlando, FL: Florida Reading Association.
- (in press). Informal reading inventory scoring criteria: Should repetitions count? Jrnl of Rdg Educ.
- Blair, T.R. & Raths, J.D. (1978). Reading comprehension as making 'unreasonable' inferences. Journal of Educational Research, 72, 29-31.
- Burns, P.C. & Roe, B.D. (1985). Informal reading inventory, 2nd ed. Boston: Houghton Mifflin Co.
- Comprehensive tests of basic skills (1975). Monterey, California: CTB/McGraw-Hill.
- Crowell, D.C. & Au, K.H. (1981). A scale of questions to guide comprehension instruction. The Reading Teacher, 34, 389-393.
- Davis, F.B. (1944). Fundamental factors of comprehension in reading. Psychometrika, 9, 185-197.
- (1968). Research in reading. Reading Research Quarterly, 3, 499-544.
- (1972). Psychometric research on comprehension in reading. Reading Research Quarterly, 7, 628-678.
-

- Harris, C. W. (1968). Measurement of comprehension in literature. The School Review, 56, 280-9, 332-42.
- Hunt, C. L., Jr. (1957). Can we measure specific factors associated with reading comprehension? Journal of Educational Research, 51, 161-171.
- Joels, R. W., Anderson, B. B., & Thompson, R. A. (1983). Progressive Reading Evaluation Portfolio (2nd ed.). Orlando, FL: Educational Services.
- Klein, A. E. (1979). Further evidence on the redundancy of the Stanford Achievement Test. Educational and Psychological Measurement, 39, 1061-1069.
- (1980). Redundancy in the Comprehensive Tests of Basic Skills. Educational and Psychological Measurement, 40, 1105-1110.
- (1981). Redundancy in the Iowa Tests of Basic Skills. Educational and Psychological Measurement, 41, 537-544.
- Royer, J.M. & Lynch, D.J. (1982). The misuses and appropriate uses of norm-referenced tests of reading comprehension. Reading Psychology, 3, 131-142.
- Silvaroli, N. J. (1982). Classroom reading inventory, (4th ed.). Dubuque, IA: Wm. C. Brown.
- Spearritt, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. Reading Research Quarterly, 8, 92-111.
- Thorndike, R. L. (1973-1974). Reading as reasoning. Reading Research Quarterly, 9, 135-147.
-