To test the assumption that questions measuring literal comprehension and those measuring inferential comprehension are equally valid indices for both oral and silent reading tests at all skill levels, questions from the Analytic Reading Inventory were classified as either literal or inferential. Subjects, 94 children in grades two to five, read grade-appropriate test passages either orally or silently and answered the comprehension questions. Although showing no direct effects from test format (oral or silent reading) or kinds of questions asked, results showed several interaction effects at different levels of competence. These effects included the following: (1) poor readers comprehended better during oral reading than during silent reading; (2) poor readers tested better on inferential questions than on literal ones when inferential questions included items measuring main idea, cause-effect relationships, and use of implied meanings; (3) the average reader comprehended better during silent reading than during oral reading and comprehended all questions equally well, (4) the good reader was generally strong in both oral and silent reading on various measures of comprehension, and (5) the single best indicator of competency was literal comprehension. (MM)
Differences in Literal and Inferential Comprehension
After Reading Orally and Silently

Samuel D. Miller & Donald E. P. Smith
Reading & Learning Skills Center
The University of Michigan

Abstract
Responses on reading tests, both silent and oral, are measured by literal and inferential questions. If the kind of question is controlled, are there differences in comprehension due to test format (i.e., silent or oral)? Do such differences exist for both good and poor readers?

Ninety-four children in grades 2-5 were asked to read, orally and silently, grade appropriate passages from the Analytic Reading Inventory (Woods & Moe, 1977). Questions were classified as literal or inferential. A repeated measures ANOVA showed no direct effects attributable to test format (whether the child read orally or silently) or kinds of comprehension (whether the child answered literal or inferential questions) but several interaction effects at different levels of competence. Results fail to support common assumptions regarding the greater ease of silent over oral reading or of literal over inferential comprehension for poor readers, but do support contentions of deficits in automaticity and attentional focus in poor readers.
When being evaluated for reading competency, students are required to (1) read a short passage either orally or silently and (2) answer a series of passage dependent questions. Whether understanding of text is assessed by silent reading tests or by oral tests, investigators commonly use two classes of questions, those measuring literal comprehension and those measuring inferential comprehension (Clymer, 1968; MacGintie, 1973; Smith, J. M., 1978). It is usually assumed that both classes of questions are equally valid indices for both kinds of test format, silent and oral, regardless of the reader's level of skill. We tested this assumption with elementary school children.

Investigators concerned with examining oral reading-silent reading phenomena have focused on the relationship between oral reading errors and silent reading ability (Gilmore, J. Y., Gilmore, E. C., 1968; Goodman & Burke, 1968; Weber, 1970). Such studies have differed according to the emphasis investigators place on the frequency and/or the type of oral reading errors and the interpretation of how such errors relate to the development of reading competency (Leu, 1982; Spache, 1981; Weber, 1968). A few have been concerned with differences in comprehension as a function of oral and silent formats (Jules & Holmes, 1980; Guthrie & Tyler, 1976). Differences seem to be found in poor readers but not in competent readers. In accordance with these findings, writers of reading texts have agreed that oral reading is more demanding than silent reading and that expectations should be lowered accordingly when students read orally (Cheek, E. H. 1980; Spache, 1976; Otto, McMenemy, & Smith, R., 1973).

Whether such a recommendation is warranted under all conditions is moot. The aspect of comprehension being measured, whether literal or inferential, may operate as a limiting condition on the validity of the conclusion. The reason for suspecting that the two kinds of questions may make a difference is the greatly disparate demands on information processing made by each (Ausubel, D. P., 1963; Anderson, 1972). A literal question demands recognition of similarities between words in the question and words in the text whereas an inferential question demands a second level recognition, one requiring the use of implied meanings. Abbott (1972; Smith, D. E. P., Smith, J. M., 1981) used similar reasoning to generate
Differences in Literal

Instructional tasks for typical poor readers in inferential reading.

Therefore, the authors asked the following questions:

1) Are differences in reading comprehension related to the test format used (i.e., whether the student reads orally or silently), as reported earlier?

2) If test format differences do occur, do they occur equally with literal and inferential questions?

3) If test format differences do occur, do they occur equally for good and poor readers?

Method

Subjects

The study was conducted in six classrooms, grades two through five, in an elementary school in southeastern Michigan, one second, a first-second, two third, a fourth and a fifth grade. Caucasians and minorities (primarily Blacks) were equally represented as was sex in the ninety-four students tested.

Measure

The Analytical Reading Inventory (Woods and Moe, 1977) consists of graded narratives, administered individually, three passages (forms A, B, and C) for each grade level, grades 1 to 9. Students may read the passages either orally or silently. The test is not normed: rather, the authors analyzed passages using common readability measures (e.g., the Spache Readability Formula and the Harris-Jacobson Readability Formula) to assure proper gradation, followed by field tests. Topics presented in the passages were chosen by the test authors based on the familiarity and interest of the topic to target students. Each passage is followed by questions testing for factual information, main idea, cause-effect relationships, and an inference or conclusion. Terminology questions following each test were not included in the analysis since they are not passage dependent.

Procedure

Two categories of questions were prepared based on criteria presented by J. M. Smith (1978): 1) literal comprehension: tasks involving matching, imitation or recognition of information as it appears in the target material; 2) inferential comprehension: tasks involving recognition of
Differences in Literal information and the generation of a response based on implied meanings. The reason for categorizing questions as literal or inferential is related to the belief that, while there is no taxonomy of questions that meets the need of inclusiveness and operationality (Andre, 1979), literal and inferential questions represent the types of tasks students face typically in reading assignments and reading evaluations. Main idea, cause-effect, inference and conclusion questions were classified under inferential comprehension and factual questions appeared under literal comprehension. In grade 2, there were two literal comprehension and three inferential comprehension questions, for each of the three forms. In grades 3, 4, and 5, there were two literal comprehension and four inferential questions for each form. Thus, second grade students answered a total of 10 questions (5 for each passage) and third, fourth and fifth grade students answered a total of 12 questions (6 for each passage). A total of 69 test questions were analyzed. The senior investigator and two independent raters classified each question as "literal" or "inferential". Agreement of the two raters with the investigator's rating is represented by coefficients of .94 and .97.

Each child was asked to read one passage orally and another passage silently at his grade level (i.e., grade placement in school) and to verbalize answers to the appropriate questions. The first student read passage A silently and passage B orally. The second student read passage B silently and passage C orally. The next student read C silently and A orally. The procedure thus provided both an equal number of silent and oral passages for each child and equal representation of the three forms for each class. Each child was tested individually and all testing was completed within a two week span.

Results

Data were analyzed by a repeated measures analysis of variance, simple effects tests and F-tests (Winer, 1977; BMPD, 1981). Equivalence of question difficulty across grades and lack of order effects were confirmed. Additionally, all tests for skewness fell within the normal range. Total scores by child were ranked and divided into thirds, thus providing three levels of competency (Level Low (L), 0-66%; Level Medium (M), 67-81%; Level High (H), 82-100%).
Differences in Literal

Results appear in Tables 1-5. Table 1 presents the mean percentage correct scores and standard deviations for both test formats and for both classes of questions. Uneven N's results from tied scores. It may be noted that while the overall mean for inferential questions favors silent reading (78.2% v. 70.1%), the values are reversed for literal questions (71.8% v. 76.1%).

Table 2 presents the main analysis. The between subjects factors are levels of reading competence and the within subjects factors are test formats and question type. The absence of main effects is notable. (Means for total silent v. oral and total literal v. inferential will be found in Tables 4 and 5.) However, three of the four possible interactional effects are significant. Those are Test Format X Question Type (p=.036), Test Format X Level (p=.034) and Level X Question Type (p = .002). Significant differences were all two way interactions; there were no significant interactions involving all three predictors, question type, test format and level of competence.

Table 3 presents the results of the test of simple effects for the Test Format X Question Type interaction. Among the four comparisons, the differences between silent inferential and oral inferential favor silent inferential (p=.006). Silent inferential is also superior to silent literal at a marginal level of significance (p=.097).

Table 4 presents the results of the tests of simple effects for the Level X Question Type interaction. The Level L, low scoring students, had higher scores on inferential questions than on literal questions (p<.001), whereas Level H, the high scoring students, had higher scores on literal questions than on inferential questions (p<.001). For Level M, the average
Differences in Literal

scoring students, mean scores on the two kinds of questions were equivalent.

Insert Table 4 about here

Table 5 presents the Level X Test Format interaction. Significant differences between modalities were found at levels L and M. Oral reading scores were higher than silent reading scores for Level L students (p<.05), while silent reading scores were higher than oral for Level M students (p<.01). The mean scores for Level H were equivalent.

Insert Table 5 about here

In sum, the tests for simple effects yielded significant differences with regard to level, at levels L and H by question type and at levels L and M by test format. Contrary to what one might have expected, Level L readers had higher scores on inferential questions than they did on literal questions, and they also had higher comprehension scores when reading orally than they did when reading silently. Level M readers had higher comprehension scores after reading silently than they did when reading orally and there was no difference between their scores on literal and inferential questions. Level H readers had higher scores on literal questions than on inferential and there was no difference between their oral and silent reading scores.

Discussion

Specific reading competencies appear to differ by skill level for this sample of elementary students but the relationship is complex. Level L readers were more adept at understanding when reading orally than they were when reading silently. That is, they scored higher on questions following oral reading than they did on those following silent reading. Such a finding runs counter to the suggestion that the attentional capacities of poorer readers are overtaxed when reading orally, limiting their ability to comprehend. To the contrary, oral reading might improve the performance of poor readers by its demand for attention to individual words. Level L readers were also more successful at answering inferential than literal
Differences in Literal questions, the very tasks which many researchers believe require a deeper level of processing. Together these findings suggest that while Level L readers had difficulty reading for details, they were able to discern the text's larger meanings.

In stating that Level L readers were more successful at answering inferential questions than they were at answering literal questions, or that they were more successful at comprehending after reading orally than they were after reading silently, it should be emphasized that their performance levels were relatively low for both question types and across both test formats. The poor readers' weaknesses were not particular to either of these specific areas of expertise but rather to "general comprehension", however it may be defined.

Level M readers had developed many of the competencies that Level L readers lacked. These students were able to answer inferential and literal questions equally well (mean comprehension =75%, 76% respectively) and they read well silently (mean comprehension =80%). However, Level M readers did have greater difficulty reading orally than silently (mean comprehension =71%). It is at this level that support is found for the claim that oral reading is more demanding a task than silent reading and that expectations should be lowered accordingly. The same conclusion derives from analyzing group data without regard to levels of reading competence. Specific differences between readers at levels L and M are related to the ability of average students to answer literal questions considerably more successfully than poor readers (76% v. 49%) and to read silently more proficiently (80% v. 52%).

Level H readers, the most competent students, differed from readers at the other two levels by reading well in both test formats (oral, x=88%; silent, x=93%) and they answered literal and inferential questions with fairly high accuracy (mean percentages =96%, 85% respectively). Also, they were more successful at answering literal as contrasted with inferential questions (although, of course they answered both better than did other readers).

While Level H readers have developed the skills required to read well, readers at the two lower levels were not necessarily less able simply because they had failed to acquire the skills mastered by the most competent
Differences in Literal

readers. As reading competency scores increased, the strengths and weaknesses which characterized each level changed in a particular manner. For example, reading competency across test formats did not develop equally in small increments: Level L readers did better on comprehension after reading orally while Level M readers did better on comprehension after reading silently. The development of oral reading skills did not parallel the development of silent reading skills; nor did the ability to answer literal questions parallel the ability to answer inferential questions. Yet differences at each level were related to the development of a particular skill area, e.g., "the ability to answer inferential questions when reading well orally," or "the ability to answer literal questions when reading silently." In brief, improvement across abilities seems to occur in steps with increments in oral, silent, inferential and literal occurring at different steps.

The one skill area which reflected the greatest gains as the level of reading competency scores increased was the ability to answer literal questions. These scores increased from 44% at Level L to 96% at Level H: no other skill area reflects so large an increase. Such a result may appear to support the claim that, as the number of details supporting a major idea increases, so does their recall of the major idea (Anderson & Reder, 1979; Craik & Tulving, 1975; Phiffer et al., 1983). In other words, comprehension of details is thought by some to facilitate comprehension of major ideas on which inferential questions are focused. However, the finding implies only the existence of an association between these two skill areas, not a causal relationship. The direction of causality cannot be determined based on the results of this study. It may very well be that the ability to comprehend the ideas assessed by inferential questions facilitates the recall of details. Furthermore, the inference of causality itself may be incorrect; perhaps both are related to a third (unnamed) variable operating as a cause.

In an effort to understand how a reader's performance on the inferential comprehension questions varied with level of reading competence, the data were analyzed further. As stated earlier, the inferential comprehension category consisted of questions measuring main idea, cause-effect relationships, and "pure" inference (hereafter "inference"). The reanalysis focused on how well readers at each level answered the individual items used
Differences in Literal

in the inferential comprehension composite. The results show that readers answered certain items more successfully than they answered other items. Level L readers answered correctly 79% of the main idea questions, 71% of the cause-effect questions, and only 34% of the "inference" questions. Level M readers answered correctly 85% of the main idea questions, 82% of the cause-effect questions, and only 57% of the "inference" questions. Level H readers answered correctly 87% of the main idea questions, 88% of the cause-effect questions, and 81% of the "inference" questions. All readers had greater facility with the main idea and the cause-effect questions than they had with the "inference" questions. From levels L to H (i.e., from the lowest to the highest scoring readers) the ability to answer "inference" questions increased 47% while the ability to answer main idea and cause-effect questions increased 8% and 17% respectively. Differences between the ability to answer each of the three types of questions may be related to the demands made by each on how information is processed. Examination of the questions by the authors suggests that each type of question may be ranked according to how implicit a response is required if a particular type of question is to be answered correctly. "Inference" questions, the most difficult to answer, seem to require not only a recognition component but also a generation component, not required by cause-effect or main idea questions. It should be noted that, as prior authors have stated, "inferential" questions are indeed more difficult than literal questions (about 15% in this data) but of the same magnitude regardless of competence level. A similar distinction is noted in the Pearson and Johnson question taxonomy (1976) between textual questions, both explicit and implicit, and scriptally implicit questions. Investigators using the Pearson and Johnson taxonomy have found scriptally implicit questions, "inference" questions in this study, to be the most difficult type of question for a student to answer (Johnston, 1984).

It must also be emphasized that all readers were asked to read passages appropriate to their grade level. How readers might have performed if the text had been at an easier level or at a more difficult level cannot be determined from our data. It may be that, if poorer readers had been given easier materials, they would have performed more successfully and their
Differences in Literal

competencies would be similar to the strengths and weaknesses of more competent readers, or it may be that, if good readers had been given more difficult materials, their competencies would be similar to the strengths and weaknesses of poor readers. In other words, reading proficiency appears to be a function of an interaction between (1) a student's competency, (2) the reading test format and (3) passage difficulty.

These results seem to have both practical and theoretical implications. First, they challenge two popular assumptions regarding reading instruction. As stated earlier, these assumptions imply that inferential comprehension is more difficult than literal and that comprehension after reading silently will be better than comprehension after reading orally. As the results indicate, a particular reader's deficiencies may depend upon his level of reading competence, upon whether he is reading orally or silently, and/or on whether or not the questions are literal or inferential. Most importantly, as the test format X question type interaction suggests, unless attention is directed to reading competence, one might easily assume that students should read silently if they are to comprehend the meaning of a text. For example, an evaluation of silent reading ability based upon an analysis of oral reading performance might overestimate the silent reading abilities of Level L readers whereas a similar evaluation might underestimate the silent reading abilities of Level M readers. Additionally, the accuracy of an evaluation of reading ability based upon a student's success at answering post text questions would vary as a function of both the student's level of reading competence and the type of question being asked. For example, different conclusions would be drawn concerning the abilities of Level L readers if the majority of questions were inferential than if the majority of questions were literal. For some students, i.e., levels L and M, the accuracy of the evaluation would further depend upon whether or not the student read orally or silently. The potential impact on reading evaluations of these factors--level of reading competence, the type of question, and test format--casts doubt on the validity of reading tests which fail to consider the conditions under which reading performance is evaluated.

Second, these results suggest that two serious deficits characterize poor readers, (1) lack of automaticity (LaBerge & Samuels, 1977) in the early developed reading skills (e.g., word recognition) and (2) reduced ability to
focus attention. First, to compensate for the automaticity deficit, poor
readers appear to use higher level skills as suggested by Smith (1976,
p.85). They maximize the use of the text's macro-structure (e.g., main ideas
> literal). Second, the requirement to read orally establishes constraints
which appear to increase attentional focus (e.g., oral > silent), thereby
reducing the effect of an attentional deficit. The attentional deficit
contention is supported by studies of distractibility (Hagen & Sabo, 1967;
Hallahan, Kaufman, & Ball, 1973) and of attentional deficits in the learning
disabled (Katz, 1978; Ross, 1976; Tarver, Hallahan, Kaufman & Ball, 1976)
(For reviews, see Hagan & Kail, 1975 and Hallahan & Reeve, 1980.).

Conclusion

Within the limitations of sample characteristics and the instrument
used, these conclusions appear to be justified:
1. The poor reader comprehends better during oral reading than during
silent reading.
2. The poor reader tests better on inferential questions than on literal
ones when inferential questions include items measuring main idea,
cause-effect relationships and use of implied meanings.
3. The average reader comprehends better during silent reading than
during oral reading and handles all questions equally well.
4. The good reader is generally strong in both oral and silent reading on
various measures of comprehension, but his acquisition of details is
superior.
5. The best single indicator of competency is literal comprehension,
i.e., reading for details.

Generalizations regarding the ability to read more competently
in one test format than the other (i.e., assuming that silent reading is
superior to oral reading, or vice versa) are inappropriate. Additionally,
generalizations regarding the ability to answer one type of question more
successfully than another type of question (i.e., assuming inferential
questions are more difficult to answer than literal questions, or vice versa)
are inappropriate. As expected, the ability to comprehend after reading
orally and silently, and the ability to answer successfully both literal and
inferential questions improves as the level of reading competency increases,
yet the strengths and weaknesses which characterize each level of development
change in a particular manner.


Differences in Literate


Table 1
Mean Percentage Correct of Two Kinds of Questions Following Oral and Silent Reading, by Level of Reading Competence

<table>
<thead>
<tr>
<th>Variables</th>
<th>Levels of Reading Competence</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Oral Inferential</td>
<td>59.2</td>
<td>69.2</td>
<td>81.4</td>
<td>70.1</td>
</tr>
<tr>
<td></td>
<td>(24.0)</td>
<td>(23.4)</td>
<td>(21.2)</td>
<td></td>
</tr>
<tr>
<td>Oral Literal</td>
<td>59.1</td>
<td>72.2</td>
<td>95.6</td>
<td>76.1</td>
</tr>
<tr>
<td></td>
<td>(36.3)</td>
<td>(32.0)</td>
<td>(44.0)</td>
<td></td>
</tr>
<tr>
<td>Silent Inferential</td>
<td>65.2</td>
<td>80.3</td>
<td>89.2</td>
<td>78.2</td>
</tr>
<tr>
<td></td>
<td>(27.5)</td>
<td>(17.2)</td>
<td>(14.0)</td>
<td></td>
</tr>
<tr>
<td>Silent Literal</td>
<td>39.4</td>
<td>79.6</td>
<td>97.1</td>
<td>71.8</td>
</tr>
<tr>
<td></td>
<td>(30.0)</td>
<td>(28.6)</td>
<td>(12.0)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>55.7</td>
<td>75.4</td>
<td>90.8</td>
<td>74.0</td>
</tr>
<tr>
<td>N</td>
<td>33</td>
<td>27</td>
<td>34</td>
<td>94</td>
</tr>
</tbody>
</table>

Note. Standard deviations in parentheses.
Table 2

Summary of ANOVA of Test Format, Type of Question, and Level of Reading Competence

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Format</td>
<td>1</td>
<td>.05</td>
<td>.84</td>
<td>.361</td>
</tr>
<tr>
<td>Test Format X Level</td>
<td>2</td>
<td>.43</td>
<td>3.52</td>
<td>.034</td>
</tr>
<tr>
<td>Error</td>
<td>91</td>
<td>5.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question Type</td>
<td>1</td>
<td>.00</td>
<td>.01</td>
<td>.923</td>
</tr>
<tr>
<td>Question Type X Level</td>
<td>2</td>
<td>.97</td>
<td>7.03</td>
<td>.002</td>
</tr>
<tr>
<td>Error</td>
<td>91</td>
<td>6.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Format X Question Type</td>
<td>1</td>
<td>.33</td>
<td>4.54</td>
<td>.036</td>
</tr>
<tr>
<td>Test Format X Question Type X Level</td>
<td>2</td>
<td>.23</td>
<td>1.56</td>
<td>.216</td>
</tr>
<tr>
<td>Error</td>
<td>91</td>
<td>6.59</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Level=Low, Medium, High; Test Format= Oral v. Silent; Question Type= Literal v. Inferential.
### Differences in Literal

#### Table 3
**Simple Effects Tests: Test Format X Question Type**

<table>
<thead>
<tr>
<th>Interaction</th>
<th>F-Ratio</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silent Inferential v. Oral Inferential</td>
<td>7.96</td>
<td>.006</td>
</tr>
<tr>
<td>Silent Literal v. Oral Literal</td>
<td>.86</td>
<td>.335</td>
</tr>
<tr>
<td>Silent Inferential v. Silent Literal</td>
<td>2.81</td>
<td>.097</td>
</tr>
<tr>
<td>Oral Inferential v. Oral Literal</td>
<td>2.02</td>
<td>.159</td>
</tr>
</tbody>
</table>

#### Table 4
**Simple Effects Test: Level of Reading Competence X Question Type**

<table>
<thead>
<tr>
<th>Level of Competence</th>
<th>Question Type</th>
<th>F-Ratio</th>
<th>P&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Literal</td>
<td>Inferential</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>49.2</td>
<td>32.1</td>
<td>16.04</td>
</tr>
<tr>
<td>Medium</td>
<td>75.9</td>
<td>74.6</td>
<td>.10</td>
</tr>
<tr>
<td>High</td>
<td>96.4</td>
<td>85.3</td>
<td>11.95</td>
</tr>
<tr>
<td>Total</td>
<td>73.8</td>
<td>74.0</td>
<td>.00</td>
</tr>
</tbody>
</table>

#### Table 5
**Simple Effects Tests: Level of Reading Competence X Test Format**

<table>
<thead>
<tr>
<th>Level of Competence</th>
<th>Test Format</th>
<th>F-Ratio</th>
<th>P&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Oral</td>
<td>Silent</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>59.2</td>
<td>52.3</td>
<td>5.21</td>
</tr>
<tr>
<td>Medium</td>
<td>70.7</td>
<td>80.0</td>
<td>7.72</td>
</tr>
<tr>
<td>High</td>
<td>88.5</td>
<td>93.1</td>
<td>2.25</td>
</tr>
<tr>
<td>Total</td>
<td>72.8</td>
<td>75.1</td>
<td></td>
</tr>
</tbody>
</table>