DOCUMENT RESUME

ED 215 315 CS 006 609

AUTHOR Raphael, Taffy E.; Pearson, P. David
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO CSR-TR-238
PUB DATE Mar 82
CONTRACT 400-76-0116
NOTE 64p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Adults; Elementary Education; *Learning Processes; *Metacognition; Prior Learning; *Questioning Techniques; *Reading Comprehension; Reading Instruction; *Reading Research
IDENTIFIERS *Reading Strategies

ABSTRACT

A study assessed the effects of metacognitive training in question answering strategies on students' ability to answer postreading comprehension questions. In the first part of the study, baseline data concerning the level of metacognitive awareness exhibited by expert readers when answering questions were gathered through the use of 44 skilled adult readers. In the second part of the study, 100 fourth, sixth, and eighth grade students of varying reading abilities were randomly assigned to experimental or control groups. Those students in the experimental condition were trained to recognize three question types (text explicit, text implicit, and script implicit) and their implied question answering strategies. Results indicated superior performance by the experimental group readers in identifying questions by type, selecting an appropriate question answering strategy, and providing a complete and accurate response to questions. Performance was better on text-based rather than knowledge-based questions, with ability and developmental levels of the students affecting their performance in predicted directions. (FL)

CENTER FOR THE STUDY OF READING

Technical Report No. 238

THE EFFECT OF METACOGNITIVE TRAINING ON CHILDREN'S QUESTION-ANSWERING BEHAVIOR

Taffy E. Raphael
University of Utah

P. David Pearson
University of Illinois at Urbana-Champaign

March 1982

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. US-NIE-C-400-76-0116.

EDITORIAL BOARD

The Effect of Metacognitive Training on Children's Question-Answering Behavior

This study represents an attempt to integrate two distinct lines of research: the role of questions as facilitators of comprehension and the role of metacognition, specifically as a means of increasing one's knowledge and flexible selection of learning strategies.

Questions, pervasive in the school environment, occur in textbooks, diagnosis and assessment procedures, and in classroom discussions. The question, often cited as a teaching technique (e.g., Socratic dialogue), is more often the means for determining what information a reader has gained from text. Labels such as skilled and unskilled reader often are based upon students' performances on these post reading comprehension questions.

Research in the area of questions reflects concerns in three related areas. First, several researchers (e.g., Barrett, 1976; Pearson & Johnson, 1978) have developed taxonomies of questions that presumably reflect some hierarchy of levels of cognitive activity. Second, researchers (e.g., Guszak, 1966; Bartolome, 1969; Chou-Hare & Pulliam, 1980) have examined the frequency of occurrence of questions from the various categories. Finally, the facilitative effect of questions on learning from or memory for textual information has been studied (e.g., Rothkopf, 1966; Frase, 1968; Anderson & Biddle, 1975). Considered together, the literature suggests (a) that a number of question categories exist, each requiring different cognitive activity or strategy use to locate correct response information; (b) that questions of a literal, or at least text-based nature, dominate the

school environment, and (c) that questions do facilitate one's learning from and memory for text.

Beyond their frequency and utility, questions can also reveal information about how the reader processes and learns from text (Kavale & Schreiner, 1979-80; Raphael, Winograd, & Pearson, 1980; Andre' & Anderson, 1978-79). By integrating the more traditional approaches to questions with a growing body of research concerning readers' awareness of and control over processes engaged in while reading (known as metacognition), researchers have been able to study the relationship between students' performances on various types of questions and their awareness of and control over those strategies used in answering questions. Thus, one question to be considered concerns our ability to teach children to use the available strategies or techniques for locating appropriate response information. That is, can we teach children to better cope with the questions they must so often face?

The second line of research, metacognition, provides the framework within which the answers to this question can be sought. The role of metacognition in research into questions stems from an increasing emphasis upon the active role of the learner or the reader. The term metacognitive knowledge has two distinct aspects: (a) knowledge of cognition--the awareness of one's own mental processes and abilities, and (b) regulation of cognition--those mechanisms that allow one to evaluate one's progress during the learning activity (Brown, 1981). Thus, in terms of question asking or question answering skills, the influence of metacognitive research has been to move question research beyond the issues of what kind of questions are asked and how questions have been asked to the issues of

what strategies students use to answer questions and when and where those strategies apply.

Within the framework of metacognition, training studies investigating the use of strategies in a number of areas have been conducted. The purpose of this work has been to evaluate whether the use of strategies facilitates the appropriate cognitive activity and whether students can be trained to employ these strategies spontaneously. Most of this research has investigated developmental differences in strategy use on memory tasks (e.g., Brown, Campione, & Murphy, 1977; Chi, 1980; Flavell, 1970). Differences that occur have been attributed to either production or to mediation deficiencies (Flavell, 1970) or to differences in prior knowledge (Chi, 1980). Production deficiencies occur when the learner either knows, or has the potential for learning when to use, a particular strategy but fails to invoke the strategy spontaneously. Mediation deficiencies occur when the learner does not invoke an appropriate strategy because of capacity limitations rather than a potentially correctable lack of strategy knowledge. The results of these training studies indicate that students representing a range in ages (K-Adult) and aptitude (retarded to normal) can be taught to use a variety of memory mnemonics successfully.

More relevant to the study of question answering strategies are the results from studies of strategy use during reading. Several researchers have suggested that the use of strategies may be one factor which differentiates the skilled from the less skilled reader (Golinkoff, 1975-76; Guthrie & Tyler, 1976; Ryan, 1981; Weinstein & Rabinovitch, 1971). As a

result of the findings concerning strategy use, researchers have begun to train readers of different ages and ability levels to increase their awareness and use of metacognitive strategies involved in question-generation (Andre' & Anderson, 1978-79; Morse, 1976; Smith, 1973) as well as question answering (Raphael, et al., 1980).

Developing awareness and training strategy use involves a number of factors. Brown, Campione, and Day (1981) have developed instructional guidelines which take several of these factors into account. The guidelines include training an instructionally relevant skill, providing feedback on performance, and giving explicit instruction regarding the applicability of the skill. The procedure is based upon the assumption that students should be informed participants in any strategy training skill program to maximize both the efficient use of a strategy and the transfer of its use to other situations. The present study uses this model in an attempt to provide an effective method for teaching students of different developmental and ability levels relevant strategies for responding differentially to situations in which questions, task demands, and available resources vary.

Specifically, we set out to evaluate the effect of providing students with a plan for directing their cognitive activity based upon their decisions about the interrelationships among a question, the available text information, and one's background knowledge. In so doing, we considered four questions:

1. Does training sensitize students to differences among the task demands and information sources associated with various kinds of questions?

2. Does training increase students' self-awareness of what they are doing when answering questions (i.e., Can they better monitor their own performance)?
3. Does training influence the quality of students' responses to questions (i.e., Are responses more accurate)?
4. Does question answering ability become more sophisticated when students (a) are aware of task demands of questions, and (b) can accurately identify the strategies they have used while answering a question?

We predicted that as a result of training, (a) students would learn to invoke question answering strategies appropriately, (b) students' awareness of their own question answering behavior would be heightened, (c) the quality of their responses (i.e., completeness and accuracy) would improve, and (d) the quality of responses would improve particularly under those circumstances in which students were able to achieve heightened awareness of their own behaviors. In addition, we expected that these predicted outcomes would vary somewhat as a function or grade or ability levels, the difficulty of the material, and the task demanded by the question.

Method

The study consisted of two experiments. The first, using adult skilled readers, attempted to provide baseline data concerning the level of metacognitive awareness exhibited by expert readers when performing tasks related to question answering. Subsumed under this goal was an attempt to determine the minimum level of instruction required for conscious

metacognitive awareness of question answering strategy use. Because the focus of interest in this study is on the effect of intervention with elementary and junior high school students, and due to the consistency of the adult data, the first study will be described only in a brief summary, followed by a detailed description of the study of primary interest. For a more complete report of the first experiment, the interested reader is directed to the original source (Raphael, 1981).

The subjects in Experiment I were 44 skilled adult readers from an introductory educational psychology course at a large midwestern university.

Subjects participated in a ten-minute instructional period which consisted of presenting definitions of three question-answer relationships (Pearson & Johnson, 1978) and the implied strategies for answering questions from each category. This was followed by a brief practice passage. After minor discussion, subjects responded to two 400- to 600-word passages and 18 comprehension questions for each passage. These materials are described in more detail in the description of Experiment II.

Scores were created based upon their ability to identify the category from which a question was created, their use of an appropriate question answering strategy, and their response quality. Results strongly suggest that skilled adult readers can (a) easily recognize, with a minimal degree of instruction, the question types and their implied question answering strategies, (b) use the most appropriate strategy in responding to the questions, and (c) exhibit consistency in the kinds of responses they provide across four expository passages. In short, Experiment I served its function of providing a baseline, corroborating the assumption that skilled

readers possessed the skills in question in their repertoire of reading and metacognitive abilities.

Experiment II

Subjects

One hundred students from a suburban school system participated in the study. Students were divided into five groups on the basis of grade level or reading ability. Thus, there were three groups of average readers at each of the fourth, sixth, and eighth grade levels and two additional groups of sixth-grade students identified as low-average and high ability. Each group was divided into ten trained and ten control (oriented, cf. p. 10) students (see Table 1). To accommodate school schedules, fourth-

---

Insert Table 1 about here.

---

and eighth-grade students were assigned randomly to treatment group (training or orientation) by class (all students had the same reading teacher); sixth-grade students were assigned randomly to the two levels of treatment individually. Three criteria were applied jointly to determine reading ability: teacher judgment, developmental reading group membership, and reading comprehension scores from the Stanford Achievement Test. As a double check on the randomization, training versus orientation comparisons were carried out with each of the five subgroups, and no significant differences were found.

Design

A 3 x 2 x 3 experimental design was used with between-subjects factors of grade (4th, 6th, and 8th) or ability (high, average, and low-average) and treatment (training and orientation). The within-subjects factor was question type (TE, TI, and SI). Two cross-age and two ability comparisons were conducted within this design. Within the age and ability comparisons, one comparison used a passage read by all students (common passage), the other used a set of passages from which students read the particular passage written at their reading level (reading level passage). That is, two comparisons involved average students at the fourth-, sixth-, and eighth-grade levels, one of which used the data from a common passage read by all students (e.g., "About Dogs") and the other used data from a reading level appropriate passage set (e.g., 4th: "Circus Clowns," 6th: "The American Cowboy," and 8th: "The Zoo Story"). Two additional comparisons involved sixth grade students of high, average, and low-average ability levels. Again, one comparison examined data from a sixth-grade level passage read by all students; a second comparison examined data from a reading level appropriate passage set.

Instructional Procedures

The instructional program was developed using the Pearson and Johnson (1978) trichotomy of question types. This categorization scheme was selected because it is unique in the method by which a question is classified. While the majority of taxonomies assume that questions can be classified in isolation, this trichotomy underscores the necessity of identifying a question type according to its relationship to both the text to which it

refers and the knowledge base of the reader. Thus, rather than speak only of question types, it becomes more appropriate to refer to question types and their implied question-answer relationships (QARs). The three question types are text explicit, text implicit, and script implicit.

A text explicit (TE) question is defined as having both the information used to create the question and that used to form an appropriate response located within a single sentence from the text. A text implicit (TI) question is defined as having the information used to create the question and that used to provide a response located in the text, but requiring the readers to integrate information across sentences, paragraphs, or pages. A script implicit (SI) question is defined as one which is based upon information in the passage, but which requires readers to search their knowledge base to provide an appropriate answer. All students in Experiment II were given simplified terms to describe the three QARs (text explicit was called Right There; text implicit, Think and Search; script implicit, On My Own).

The following brief paragraph and three questions illustrate the three QAR categories:

(1) Robbi wore a scarf to school today. It was bright red.

TE: What did Robbi wear to school today? (a scarf)

TI: What color was Robbi's scarf? (red)

SI: Why did Robbi wear a scarf to school? (it was cold)

During the instructional phases of the study, students received an explanation of the three QAR classifications and practice in identifying them. Then they received instruction concerning the implication of each

question type in terms of task demands, the amount or detail of the explanation varying with treatment group membership.

All training was conducted with groups of students ranging in number from 10 to 25. While "levels of training" is used to refer to both treatment groups, a note of explanation is necessary. To reflect the fact that students in the control group received an introduction into the terminology and task used in the study, the label "orientation" group has been adopted. However, it would be misleading to consider their instruction a proposed training level. Rather, since the orientation was identical to the minimal level of introduction necessary for a skilled reader (as in Experiment 1) to perform the experimental tasks, this group should be considered a control group.

Students in the orientation group received a description of the three QARs and one brief practice passage with a sample question representing each of the three categories. Again, this reflects the minimum information needed by skilled adult readers to successfully complete the experimental task.

Those students in the training group received a four-day instructional program prior to doing the experimental task. On the first day, using identical materials with students at all ability and grade levels, question answering in general was discussed and the three question-answer relationships were described to provide a conceptual framework. The students then read passages of 30-50 words and responded to at least one question from each category per passage. Immediate feedback was given on both the selection of the category and the quality of the response. On the second

day, students were given guided practice with five grade appropriate passages (4th, 6th, or 8th grade reading level) from 75-150 words long and with corresponding questions from each of the three categories. On the third day, all students responded to a longer passage divided into four 150-word sections. Each section was followed by two questions from each QAR category. Students read and responded to questions in the first section as a group. The remaining sections and questions formed a criterion test to determine whether or not individual students had learned the three QAR categories. The fourth day consisted of individual practice for students who had scored below 75% on the criterion test. Only four students from the fourth grade did not reach criterion by day three; all students reached it by day four.

Experimental Task

Students completed the experimental task after they had been exposed to one of the two treatment levels. The orientation group received their instruction immediately prior to participation in the experimental task. For those in the training group, the experimental task was completed on day five. The task consisted of two sub-tasks after reading a passage of 400-600 words in length. The first task involved locating the correct response to a comprehension question; the second required the simultaneous classification of the question answering strategy in terms of the QAR involved in locating the response information. In other words, students were asked to both consider how they located information to be used in the response, as well as what that information should be. The following example illustrates the integrated nature of the task:

(2) What is the dog often called?

TE* (Man's Best Friend)

TI _____

SI _____

*For the schoolchildren, the mnemonics RIGHT THERE, THINK & SEARCH, and ON MY OWN replaced TE, TI, and SI respectively.

Scoring Scheme

For all responses collected in Experiments I and II, a categorization scheme (Raphael, et al., 1980) developed for scoring responses to comprehension questions was used. Basically, the scheme directs the person classifying responses to make a set of binary (yes-no) decisions to a sequence of questions concerning the answer location and appropriateness, ending up at different terminal nodes in a decision tree. The terminal nodes represent the classification of the question response. The system is quite involved; however, the reader can find a complete explanation in Raphael (1981). To aid in comprehension of this article, sample responses and their classifications will be provided.

Assume question (2) is based upon the following text:

(3) No matter what kind of dog it is, all dogs have one thing in common. They dearly love the people with whom they live. No wonder the dog is often called "Man's Best Friend."

The answer to question (2), "Man's Best Friend," would be classified as a "Total Deleted Constituent." This correct response category indicates verbatim response from the text. A response of "Mutt" or "Rover" would be classified as "Plausible Script Response" recognizing it as an acceptable

knowledge-based response. The response "Echo" would be classified as an "Implausible Script Response" indicating that it is an inappropriate knowledge-based response since it cannot be considered a typical name or label for dogs in general.

To validate the reliability of the scoring scheme, the percentage of agreement in response classification between two independent judges was calculated. A 10% sample of the test materials was systematically selected for use in this validation such that each combination of passages and treatments was represented. Though percent agreement was assessed separately for each passage, the results were so consistent that the data will be reported collapsed across the four passages. Percent agreement was 97% with no disagreements involving a change from text to script categories, and only .6% involving a distinction between appropriate/inappropriate categories.

Results

Four dependent measures were used to assess performance of the school-age students. The first, hits, identified the number of correct QAR identifications or question-answering strategy selections. The second, matches, indicated the degree of agreement between strategy identification and strategy use. The third, response quality, examined the quality of the answer provided by the reader. The fourth, hit-match-response quality, indicated the students' ability to both recognize the task demands of the question and use an appropriate strategy for generating a correct response.

Analyses will be reported by dependent measure. Within each dependent measure will be four sets of results: (a) comparison across average students in grades four, six, and eight on a common passage (hereafter,

Grade Comparison--Common Passage), (b) comparison across these same average students on a passage at their level of reading ability (hereafter, Grade Comparison--Reading Level Passage), (c) comparison across low-average, average, and high ability sixth-grade students on a common passage (hereafter, Ability Comparison--Common Passage), and (d) comparison of these sixth grade students on a passage at their level of reading (hereafter, Ability Comparison--Reading Level Passage).

Analyses of variance were conducted for each measure using the factors of ability or grade, treatment level, and question type. Throughout this experiment, post hoc analyses were computed using Newman Keuls procedures.

Means will be reported in proportions since arcsine transformations were performed on all proportions but did not affect the results. An alpha level of .05 was used throughout the study. However, for the convenience of the reader, when the F value was significant at the .01 level, it will be reported as such. Again, the reader is referred to Raphael (1980) for a complete report of the data. Due to constraints of space, the results of Experiment I will not be reported, and only minimum attention will be given to interactions which are not readily explainable or apparently due to "materials" effects.

Classifications by Question Types (Hits)

The ANOVA for the Grade Comparison--Common Passage revealed significant differences for grade, $F(2,54) = 3.32$, $p < .05$; treatment, $F(1,54) = 31.72$, $p < .01$; and question type, $F(2,108) = 9.01$, $p < .01$. The post hoc procedure revealed no commonly accepted significant differences among the three grade levels; however, differences between the fourth and sixth,

$p < .055$; and the fourth and eighth, $p < .055$, were very close to reaching standard levels of significance. Post hoc analysis of question type revealed no significant difference between text explicit and script implicit question ratings, but performance on both of these questions exceeded that of text implicit questions. Performance of students in the training condition exceeded that of students in the orientation condition (see Table 2).

Insert Table 2 about here.

The significant treatment X question type interaction, $F(2,108) = 3.76$, $p < .05$ (see Figure 1), revealed that students in the training groups exhibited relatively similar performance across question types, whereas students in the orientation condition, while lower on every question type, were particularly low on text implicit questions. Hence, it is likely that the main effect for question type was primarily due to the poor performance of the orientation group on text implicit questions. There were no other significant interactions.

Insert Figure 1 about here.

The Anova for the Grade Comparison--Reading Level Passage revealed similar patterns of results. Main effects were found for treatment, $F(1,54) = 7.97$, $p < .01$, and question type, $F(2,108) = 3.28$, $p < .05$. Post hoc analysis revealed the same pattern of significant differences found in the common passage analysis. In addition, a similar treatment X question type interaction was significant, $F(2,108) = 4.95$, $p < .01$. Given that this analysis was based on data from passages at the students' appropriate reading level, lack of a main effect for grade was not surprising; this

suggests that the treatment was operating in a manner that was independent of the students' age.

In the Ability Comparison--Common Passage, the ANOVA revealed significant main effects for ability, $F(2,53) = 10.80$, $p < .01$, and treatment, $F(1,53) = 12.59$, $p < .01$. Post hoc analyses revealed significant differences among the three ability levels, with high ability students performing at a higher level than did average who performed at a higher level than did low-average ability students. Subjects in t training group again scored at a higher level than those in the orientation group (see Table 2). These effects were complicated by a significant ability X treatment X question type interaction, $F(4,106) = 3.73$, $p < .01$, indicating that training differentially affected performance across question types (see Figure 2).

Insert Figure 2 about here.

The ANOVA for the Ability Comparison--Reading Level Passage reveals a main effect pattern similar to that found with the common passage comparison, with significant differences for ability, $F(2,53) = 10.41$, $p < .01$, treatment, $F(2,53) = 6.07$, $p < .05$, and question type, $F(2,106) = 4.98$, $p < .01$. The post hoc Newman Keuls procedure revealed significant differences between high and low-average ability students, and between high and average students, with no significant differences between the average and low-average students. This differs from the analysis in the grade-level comparison where the grade effect was no longer significant when students read passages of appropriate readability for their grade level. Instead, the effect of ability remained constant across the different passages read.

For question type, significant differences were found between text explicit and script implicit questions (see Table 2). The significant ability X question type interaction, $F(4,106) = 2.73$, $p < .05$ appeared to be caused by the performance of low-average students on script implicit questions (see Figure 3). Generally, students scored higher on text explicit than

---

Insert Figure 3 about here.

---

text implicit or script implicit questions when rating them by type. Low-average students, however, performed as well on the script implicit as on the text explicit questions, most likely explained in terms of a spurious "materials" effect. It appears that the script implicit questions for the passage appropriate to their reading level were easier than for the common passage.

Across the four ANOVAs, results suggest that children tended to rate text explicit questions with the most consistent degree of accuracy; performance on script implicit questions was generally near the level of text explicit questions. While overall, performance was lowest when rating text implicit questions, students in the training group tended to rate such questions more accurately than did those in the orientation group, often approaching performance levels on text explicit and script implicit questions (see Table 2).

Matches: Ratings and Responses

Recall that matches occur when students give responses consistent with their classification of the type of response they thought they were giving.

In the Grade Comparison--Common Passage, main effects were revealed by the ANOVA for treatment, $F(1,54) = 11.85$, $p < .01$, and question type, $F(2,108) = 7.54$, $p < .01$. Students in the training groups had a higher proportion of matches than those in the orientation group. Post hoc analyses revealed significant differences in levels of performance between text explicit and script implicit questions, and between text explicit and text implicit questions (see Table 3). However, there was no significant difference in levels of performance between text implicit and script implicit questions, though the differences were close to traditional levels of significance ($p < .10$). There were no significant interactions.

Insert Table 3 about here.

The ANOVA in the Grade Comparison--Reading Level Passage revealed a similar pattern, though only the effect for question type was significant, $F(2,108) = 9.08$, $p < .01$. The effect of treatment level, though not significant ($p = .09$) revealed a trend in the same direction as in the common passage comparison, with the training group performing at a higher level than the orientation group. Post hoc analyses revealed significant performance differences between text explicit and script implicit and between text implicit and script implicit questions, but no significant differences between the two text-based question types (see Table 3). No other significant main effects were found, nor were there any significant interactions.

The ANOVA for the Ability Comparison--Common Passage revealed significant main effects for ability, $F(2,53) = 8.25$, $p < .01$; treatment, $F(1,53) = 13.91$, $p < .01$; and question type, $F(2,106) = 19.57$, $p < .01$.

No significant interactions occurred. Post hoc computations revealed significant differences only between the low-average and average ability students, and the low-average and high ability students. The effects of treatment and question type were in the same direction as those reported in the grade comparisons. The training group performed at a higher level than the orientation group. Post hoc computations revealed significant differences between performance levels on script implicit and both text explicit and text implicit questions, but no significant differences in performance between the two text-based question types (see Table 3).

In the ANOVA performed in the Ability Comparison--Reading Level Passage, main effects were found for ability, $F(2,53) = 3.75$, $p < .05$; treatment, $F(1,53) = 5.10$, $p < .05$; and question type, $F(2,106) = 13.45$, $p < .01$. Post hoc analyses revealed similar patterns of significant differences as in the common passage comparison with the same group. There was a significant difference in performance between high ability students and both average and low-average. The differences between the low-average and average students was near the conventional level of significance ($.05 < p < .10$). Students in the training group performed at a higher level than did those in the orientation group. Significant differences occurred between script implicit questions and both text explicit and text implicit questions, with performances higher for text-based rather than script-based questions (see Table 3). There were no significant interactions.

Across the four analyses, when examining the number of matches as a function of grade or ability, treatment level, and question type, it appeared that performance on text-based questions was consistently higher than performance on script-based questions. That is, fo either text

explicit or text implicit questions, students were more likely to provide a response that was consistent with their perception of the question-answering strategy they thought was required by the question. For script implicit questions, the students were less consistent in their perception of the question and selection of appropriate question answering strategies. Though generally there were no significant differences in performance on the two text-based question types, one exception to this pattern emerged in the grade comparison on the common passage. In this instance only was the level of performance higher for text explicit than for text implicit questions.

When assessed on the same passage, performance of students in both the grade and the ability comparisons indicated the facilitative effect of training over orientation. When responding to reading level appropriate material, while training tended to be more beneficial than the orientation, the difference was not consistently at a traditional level of significance. It appeared that greater instruction in question answering strategies led to more consistency between perception and implementation of appropriate question answering strategies. While average students across the three grades did not differ significantly from one another in their performance levels on this task, sixth grade students of differing ability levels did vary in the predicted direction.

Correct Response to the Questions

The scores for this dependent measure were based on the Raphael, et al. (1980) scoring scheme. By collapsing across scores in the fifteen

possible categories, accurate and inaccurate scores were determined. Response quality was categorized independently of the QAR category ratings.

An ANOVA on the Grade Comparison--Common Passage revealed significant main effects for grade, $F(1,54) = 3.60$, $p < .05$, and question type, $F(2,108) = 12.82$, $p < .01$. No significant effect for treatment was found at the conventional level of significance, though the trend favored the training group ($M = .87$), $F(1,54) = 3.36$, $p < .08$, over the orientation ($M = .82$) group. For the grade effect, post hoc computations showed significant differences between fourth- ($M = .81$) and eighth-grade ($M = .89$) students, but no differences between sixth-grade students ($M = .83$) and either fourth- or eighth-grade students. In addition, significant differences for the question type effect were found between script implicit ($M = .77$) and both text explicit ($M = .88$) and text implicit ($M = .89$) questions. In short, the number of correct responses increased proportionally with age; and students tended to perform more successfully on text-based rather than script-based questions. There were no significant interactions.

On the Grade Comparison--Reading Level Passage, the trends were not as consistent. An ANOVA revealed significant effect for grade, $F(1,54) = 4.31$, $p < .05$, and question type, $F(2,108) = 19.67$, $p < .01$. The question type results followed the same pattern as that of the common story analysis, but the grade effect differed. The only significant difference was between fourth- ($M = .83$) and sixth-grade ($M = .77$) students; the difference between sixth and eighth ($M = .84$) missed traditional statistical significance by a small amount, $p < .07$. It seemed that given a passage of appropriate

reading level, the rate of successful performance did not follow a predictable direction, perhaps indicating little more than the fact that performance depends heavily on passage constraints. The treatment effect, while in the expected direction, did not even approach conventional significance, $F(1,54) = 1.67$, $p < .21$.

These results were complicated by the two-way interaction of grade X question type, $F(4,108) = 3.69$, $p < .01$. It appeared that in the fourth and the eighth grades, more correct responses were given to text explicit than text implicit questions, and both text-based question types had a higher proportion of correct responses than did the script implicit questions. However, in the sixth grade the pattern differed. While students performed at a higher level on text than on script questions, they tended to be more successful with text implicit than text explicit questions (see Figure 4). Again, the only explanation seemed to be passage specific differences: it may be the case that for the sixth grade level readability passage, text implicit questions were inherently easier than were the text implicit questions for the other three passages, possibly as a function of degree of topic familiarity.

Insert Figure 4 about here.

The ANOVA for the Ability Comparison--Common Passage revealed significant main effects for ability, $F(2,53) = 20.01$, $p < .01$; treatment, $F(1,53) = 4.09$, $p < .05$; and question type, $F(2,106) = 37.80$, $p < .01$. Post hoc analyses reveale that all pair-wise comparisons for these three variables were significant. For the ability effect, the proportion of correct responses achieved was in the predicted direction with high ability

students ($\underline{M}$ = .91) exceeding average ($\underline{M}$ = .77), and average exceeding low-average ($\underline{M}$ = .60) students. The treatment effect again demonstrated the superior performance of the training group ($\underline{M}$ = .80) over the orientation group ($\underline{M}$ = .72). Students scored higher on the text explicit ($\underline{M}$ = .77) and text implicit ($\underline{M}$ = .85) than on the script implicit ($\underline{M}$ = .66) questions, and higher on the text implicit than on text explicit questions.

The significant ability X question type interaction, $\underline{F}(4,106) = 2.41$, $\underline{p} < .05$ (see Figure 5), revealed a similar ranking of question types across all ability levels, with the probable source of the interaction being the extremely low performance on script implicit questions by the low-average students. This was further complicated by a significant three-way interaction between ability, treatment, and question type. While the

---

Insert Figure 5 about here.

---

general pattern remained unchanged, in the training condition the magnitude of the differences in performance levels of the three ability levels was smaller than the magnitude of the differences in performance levels between the three ability levels in the orientation group. The source of the interaction was due to the fact that the general pattern of higher performance on text than script, and on text implicit over text explicit held for all groups except the trained high ability students, for whom the training led to an even level of performance across all question types (see Figure 6).

Insert Figure 6 about here.

An ANOVA for the Ability Comparison--Reading Level Passage revealed similar significant main effects: ability, $F(2,53) = 4.24$, $p < .05$; treatment, $F(1,53) = 6.29$, $p < .05$; and question type, $F(2,106) = 21.61$, $p < .01$. Post hoc analyses revealed no significant differences between the low-average ($M = .79$) and the average students ($M = .77$). However, the high ability students ($M = .87$) performed significantly higher than the other two ability groups. Post hoc analyses also revealed significant differences in performance as a function of question type only between script implicit ($M = .71$) and either the text explicit ($M = .86$) or text implicit ($M = .86$) question type. Students provided more correct responses to either type of text question than to those script-based ones. Consistent with previous findings, the training group ($M = .85$) performed at a significantly higher level than did the orientation group ($M = .77$). The significant ability X question type interaction, $F(4,106) = 2.60$, $p < .05$, revealed that while both high and average ability students exhibited a gradually decreasing trend across the TE - TI - SI trichotomy, low-average students scored higher on the text implicit than on text explicit questions (see Figure 7).

Insert Figure 7 about here.

Across the four ANOVAs for response quality, the question type effect was generally consistent. Correct responses occurred more frequently on text than on script questions, with little difference between performance on text explicit and text implicit questions. When compared on a common passage, the effect of ability was in the expected direction as was the grade effect. However, this pattern was not replicated in the comparison of reading level appropriate stories. While high ability students tended

to provide more correct responses than did average or low-average, the magnitude of the difference was less than in the common story comparison. In the comparison of fourth-, sixth-, and eighth-grade students, using the reading level passage set, fourth- and eighth-grade students provided more correct responses than did sixth-grade students. This is likely to be due to a materials effect similar to that mentioned earlier, since comparison on a common passage revealed response patterns in the expected directions.

Where the treatment level was significant, the training group performed at a higher level of success than did the orientation group, a trend that existed even when the effect did not reach a conventional level of statistical significance. For example, in the eighth grade comparison, common passage, the level of performance was higher for the training group, $.05 < p < .08$.

In summary, with the exception of those responses which involved text implicit questions in the sixth-grade passage, performances were consistent with respect to findings on the first two dependent measures. Text questions of either type elicited higher levels of performance than did script questions, and training was more facilitative than was orientation. The results for the treatment effect suggest training was particularly effective with high ability students on their responses to script questions, and with low-average and average students in their responses to text-based questions.

Correct Hit-Match-Response Quality

Recall that scores for this dependent measure were created by identifying cases in which students who received both a "hit" and a "match" on

a question also provided an answer from one of the accurate response categories. The ANOVA for the Grade Comparison--Common Passage revealed significant effects for treatment, $F(1,54) = 32.65$, $p < .01$; and question type, $F(2,108) = 13.72$, $p < .01$. Training ($M = .73$) was superior to orientation ($M = .52$). Post hoc analyses revealed no significant differences between text explicit ($M = .69$) and script implicit questions ($M = .69$), but both were significantly better than text implicit ($M = .49$). The significant treatment X question type interaction was similar to that in the Reading Level Passage comparison and will be discussed later.

The ANOVA for the Grade Comparison--Reading Level Passage revealed the same pattern. The effects for treatment, $F(1,54) = 10.15$, $p < .01$; and for question type, $F(2,108) = 6.78$, $p < .01$, were significant. Training ($M = .67$) was again superior to orientation ($M = .53$). Post hoc analyses revealed no significant differences between text explicit ($M = .62$) and script implicit questions ($M = .67$), but both differed significantly from text implicit ($M = .52$).

The significant treatment X question type interaction (see Figures 8 and 9) revealed similar patterns across comparisons. In both cases, students in the training group scored higher than those in the orientation group on all three types of questions; however, the between group difference was much larger for text implicit than for either text explicit or script implicit questions. Generally, the training appeared to greatly

---

Insert Figures 8 and 9 about here.

---

facilitate the students' awareness of the need to integrate text information, thus improving their ability to cope with this type of question. Overall, it appeared that orientation was most successful in improving performance on text explicit and script implicit questions, though even in these two categories, performance improves with the kind of additional training provided by the treatment.

An ANOVA on the Ability Comparison--Common Passage revealed significant effects for ability, $F(2,53) = 13.54$, $p < .01$, and treatment, $F(1,53) = 19.16$, $p < .01$. Post hoc analyses revealed significant differences across the three levels of ability; again, the high ability students ($M = .70$) outperformed the average ($M = .56$) who outperformed the low-average ability students ($M = .42$). The training group ($M = .65$) scored higher than the orientation group ($M = .47$).

These results were complicated by a significant ability X treatment X question type interaction $F(4,106) = 3.08$, $p < .05$. Training, as opposed to orientation, facilitated high ability students' performances on both text explicit and script implicit questions. Yet the training did not appear to improve performance on text implicit questions to any larger degree than did orientation. Within the average ability students cell, training raised the level of these students to the level of the high ability students in the orientation group on all three question types. With the low-average students, training seemed most facilitative on text-based questions, particularly text explicit ones (see Figure 10).

Insert Figure 10 about here.

The ANOVA in the Ability Comparison--Reading Level Passage revealed main effects for ability, $F(2,53) = 8.77$, $p < .01$; treatment, $F(1,53) = 11.97$, $p < .01$; and question type, $F(2,106) = 10.07$, $p < .01$. Post hoc analyses revealed significant differences between the low-average ($M = .53$) and the high ($M = .71$) ability students, and between the average ($M = .56$) and the high. Significant differences occurred between text implicit ($M = .49$) and text explicit ($M = .64$), and between text implicit and script implicit ($M = .65$) question types. There was no significant difference between the script implicit and the text explicit question types. It appeared that, in general, students with passages at their reading level performed at a lower level on text implicit questions than they did on the common sixth-grade passage. This argues for a "materials" effect explanation once again. Students in the training group ($M = .66$) performed at a higher level than those in the orientation group ($M = .53$).

The significant ability X question type interaction, $F(4,106) = 3.53$, $p < .05$, indicated that while text implicit questions were the most difficult for students of all levels of ability, the patterns of performance on text explicit and script implicit questions, and the magnitude of the pairwise differences on all question types, varied across the three levels of ability. High ability students performed highest on text explicit questions, and only slightly higher on script implicit than on text implicit questions. Average and low-average students generally followed a SI > TE > TI pattern. However, for the average students, the magnitude of the pairwise differences among the three question types was quite small, while for the low-average it was large. For the low-average students,

performance differences between text explicit and script implicit questions were smaller proportionally than performance differences between TE-TI and SI-II comparisons (see Figure 11).

Insert Figure 11 about here.

Across the four analyses for this dependent measure, the results generally favored the training over the orientation group. This was true in all cases except high ability students responding to text implicit questions. This was the only situation in which training seemed no more facilitative than orientation.

In contrast to the first three dependent measures, performance on script implicit questions was generally higher than that on text implicit questions, and often equal to that on text explicit questions. It seems to be the case that students can access an accurate answer to a script implicit question as easily as a text explicit one if they first recognize the strategy optimal for responding to the question. That is, if they realize that they should depart from the text to answer the question, they are as likely to give an appropriate response as they are when they recognize that they should consult the text.

Discussion

The first question of concern was the degree to which training would sensitize students to differences among the task demands and information sources appropriate to various kinds of questions. We predicted that students would learn the system taught them and hence would display greater sensitivity to the relationship between questions and strategy use in

locating appropriate responses. In other words, the training should have better enabled students to identify the type of QAR most readily invited by a particular question, and this is precisely what occurred, albeit with differing patterns in different analyses.

Adults who received a brief orientation performed with near perfect accuracy. While accuracy overall was lower than for the adults, average fourth-, sixth-, and eighth-grade students in the training group consistently rated the questions from each category more accurately than students in the orientation group, most notably on text implicit questions. Ability tended to affect performance in the expected direction, with the high ability students performing the most successfully, followed by average and low-average students. However, when examined in terms of training levels and QAR categories, ability differentially enhanced students' performances on each question type; yet, the different patterns appear to be most likely due to spurious materials effects. Nonetheless the overall pattern, even within interactions, was for trained students to exhibit better QAR ratings than oriented students.

The second concern of the study was the degree to which training would sensitize students to their own question-answering behaviors. We predicted that trained students would become better monitors of their own performance, thus achieving a higher proportion of matches (providing responses which matched their QAR rating, independent of response quality).

The elementary and junior high school students again performed at a generally lower level than did the adults, with the students in the training group performing at a higher level than those in the orientation group.

This pattern was consistent across both age and ability when students were compared on a common story. When given passages at an appropriate reading level (regardless of grade level) a somewhat different pattern emerged. Both training and orientation groups performed at the same level on script implicit questions, but the training group performed at a higher level than did the orientation group on both text-based questions. All students in the training group showed greater consistency in identifying questions as text explicit and text implicit and locating the appropriate text-based information for their response than they did when the questions were script-based. It was easier for students to provide a text-based response to a QAR they had rated as text-based (TE or TI) than it was to provide a scriptal response to a QAR they had rated as script-based. This supports findings such as Guszak's (1966) and more recently, Chou-Hare and Pulliam (1980) that students have an easier time handling literal (i.e., text-based) questions. Nonetheless, the second prediction was also substantiated: training consistently enhanced performance in comparison to an orientation; however, there were variations in this pattern across question types and some minor variation as a function of grade and ability levels.

The third question of concern was the degree to which training enhanced the quality of the students' responses. We predicted that trained students would give better quality answers to the questions, quite independent of whether they also achieved hits and matches. The pattern of results suggests that the evidence for the training effect is less impressive here than for the preceding dependent measures: In only two

of the four analyses did a training effect emerge, although it approached conventional levels of significance on the other analyses.

The training effect was complicated by several interactions. Training had differential effects for each ability level. For high ability students, the effect of training over orientation was largest when they responded to script implicit questions. The average and low-average students appeared to benefit most by training in responding to both text-based questions. The data suggest that a still longer training period may be necessary to improve performance for low-average and average student on script implicit questions. There may be a threshold effect in that the high ability students were performing quite successfully on the text-based questions and the training may have merely increased their sensitivity to the value of their knowledge base. The students of lower ability improved on text-based questions, but perhaps need further training in the use of the information in their knowledge base.

In addition, when given a passage common across all grade or ability levels, the expected ability and developmental trends occur. However, when given a passage at the appropriate reading or grade level, developmental differences tended to either not reach a significant level or to decrease in the magnitude of the differences. Though the grade level differences were not stable across passage of different readability levels, ability differences remained constant and in the expected directions regardless of passage level read.

In summary, the findings suggest that overall, text-based QARs are easier for students to cope with than script-based QARs and training

appears to facilitate performances differentially across ability levels and question types.

The fourth question evaluated the degree to which correct application of strategy selection procedures that were part of the instruction would enhance the ability to respond to questions. We predicted that the response quality differences between trained and oriented students would be accentuated in precisely that subset of situations in which student judgments about the strategies they had used corresponded to the system they learned (were hits) and were consistent with the kinds of responses they actually gave (were matches).

Results again, in general, favored training over orientation with ability level affecting performance in the expected direction. It was interesting to note that while ability had an effect, there appeared to be no developmental effect across the three grade levels; yet all grade school students performed with less accuracy (had a higher proportion of incorrect responses) than did the adults.

In all analyses, performance was least successful on text implicit questions, and it is with these questions that training appeared to have its greatest effect over orientation. Across all grades, those students who received training performed at a higher level overall than did those in the orientation group; training resulted in performance levels on text implicit questions equal to those of both text explicit and script implicit questions. This occurred for passages common to the grade or ability levels as well as those of the appropriate reading level.

When examining differences as a function of ability, the same general patterns emerge with a few interesting exceptions. For example, training did not appear to enhance the performance of high ability students on text implicit questions to a greater degree than did orientation. Yet, training did significantly improve their performance on the text explicit and script implicit ones. Perhaps further training on TI questions could involve explicit directions for methods of text integration.

The pattern of performance of the low-average students is different on the fourth measure than on previous ones. Their level of accuracy for script implicit questions exceeds or differs only slightly from both text explicit and text implicit questions, suggesting that if they recognize a question as requiring a script-based response and go to their knowledge base to locate it, they perform at a higher level of accuracy than when responding to a text-based question. This appears to support the training of these students for in identifying the QAR and selecting the matching question-answering strategy, they completely alter and surpass previous performance levels.

Thus, it appears that the fourth prediction was confirmed. Performance is most enhanced when students both recognize the QAR most readily invited by the question and use an appropriate question-answering strategy; this generalization, like the others, is qualified by context-specific effects imposed by question type, passage type, and ability.

The literature concerning questions suggests that (a) there are different types of questions, each requiring different cognitive operations, (b) questions are an inherent part of the educational environment,

(c) questions are an effective tool for enhancing learning, and (d) we can learn how information is processed by studying how questions are answered. The results of this study lend support in varying degrees to each of these points.

The work with taxonomies has been largely non-empirical. A few studies (e.g., Chou-Hare & Pulliam, 1980; Washington, 1979; Wixon, 1980) provide evidence that the Pearson and Johnson (1978) taxonomy can be operationalized for both question categorization and question creation. The present study extends this work, providing data to support the idea that questions can be created within each category and that overall, adult skilled readers and students can be taught consistency in identifying QARs and selecting appropriate strategies for responding to the question. Consistent with Wixson's (1980) and Frase's (1968) findings, text implicit and script implicit questions were more difficult than text explicit questions. Consistent with Hansen's (1981) results, children generally performed better on text than script questions.

The metacognitive research, particulariy work in metamemory, suggests that students can be made aware of strategies used during a variety of cognitive tasks, and that important in iearning any skill is not only the knowledge of, but also the control over, a given process. Research (e.g., Ryan, 1981; Brown & Campione, 1977; Golinkoff, 1975-76; Olshavsky, 1976-77) has indicated that successful use of strategies is highly dependent upon the ability and age of the iearner. The results of this study are consistent with these findings, extending this work into the area of question answering strategies.

Of the sixteen analyses conducted in this study, training groups performed at a higher level than did orientation groups throughout, usually at a conventional level of significance. This provides reasonably strong support for the position that the use of question answering strategies is a trainable metacognitive skill. Generally, consistent with Day (1980) and Gordon (1980), it appeared that training in metacognitive skills is generally facilitative, and that more explicit training is superior to less explicit instruction.

Performance differed in the expected directions as a function of age and ability when students were compared on a common passage (i.e., an identical passage and question set for comparisons of grade or ability levels). An important difference in patterns emerged, however, when students were given a passage of appropriate readability. The differences between levels of ability when compared on the level-appropriate passages were smaller than had existed on the common passage, and at times did not reach the conventional level of significance. The contrast was similar when comparing across grade levels. While differences in performance existed across grade levels on the common passage, this rarely occurred when passages of appropriate reading level were used. This suggests that older or more mature readers using the same strategies and texts as the younger or less skilled readers are not necessarily required to do what appears to be the same task. The more ecologically valid task may be to use materials at appropriate reading and grade levels.

In fact, one finding of this study worth noting stems from the selection of the two passages to be used throughout the second experiment. Recall that there were two basic comparisons among levels of students. The

first involved a developmental comparison using only average students in fourth, sixth, and eighth grades. The second, an ability comparison, had as subjects only sixth grade students of low-average, average, and high reading abilities. In designing the study, major consideration was given to the material upon which the students would be compared. Some (e.g., Taylor, 1980) have argued that content must be held constant, but that the actual readability of the passage should be at the reading level of the student. Others (e.g., Tierney, Bridge, & Cera, 1978-79) have cautioned that it may be methodologically unsound to draw comparisons across different passages, that this confounds any effect with materials. Both arguments seemed to have merit; thus one common and one level appropriate passage were selected for each comparison. To maximize the comparability of the different passages to be used, extensive norming and pilot studies were conducted; in addition, the first experiment indicated that the four passages were comparable for skilled readers.

When performances were compared within ability and age levels and on treatment and question type effects across analyses of the two passages, predictable patterns emerged. The expected developmental trend occurs when all students respond to a common passage of a fourth grade readability level: 8 > 6 > 4. However, when students reacted to a reading level appropriate passage, the patterns were not as consistent. For example, there was not developmental trend for hits or matches.

However, this is not necessarily the case when ability comparisons are drawn. The effect of ability was more stable across passages, reaching significance on all four of the common and all four of the reading level

passage ANOVAs. Nonetheless, while the effect of ability remained whether one used a common or reading level appropriate passage, the discrepancy between the pairwise comparisons decreased when the students responded to passages of their reading level.

It appears that there is no simple solution to the problem of materials to be used in developmental or cross-ability studies. Reading level passages confound content with grade or ability level. Common passages place students of differing ability in different difficulty situations. Perhaps all that can be suggested at this point is a caution against interpreting significant differences in performance as being due to either ability, age, or task content until the effect is replicated with both types of passages.

Finally, this study lends support to the criteria for successful training studies suggested by Brown, Campione, and Day (1981) and listed earlier in this paper. The skill was instructionally relevant as suggested by the question literature (e.g., Anderson & Biddle, 1975; Guszak, 1966; Lehnert, 1979; Rothkopf, 1966; Wixson, 1980). Training proceeded from the simple to the complex (recall training procedures, cf. pp. 10-11). An analysis of the training and transfer tasks (Raphael, et al., 1980) provided evidence of where breakdowns could occur. Direct instruction about strategy use along with immediate feedback during the class discussions and after individual work improved performance. A variety of passages were used to facilitate transfer of the strategy use to new situations, and finally, self-checking procedures were an inherent part of operationalizing the text explicit--text implicit--script implicit distinction. Students were taught to monitor their success in selecting a strategy by using the

Right There, Think and Search, and On My Own trichotomy (cf. pp. 10-11). That the guidelines suggested by Brown, et al. were implemented and validated indicates that their instructional model has promise for future instructional design.

There were several factors which suggest limitations in interpreting the present data. For example, the sample of elementary and junior high school students was from a school system whose students typically performed above the national norms on standardized tests. Thus, each group of students was performing approximately one grade equivalent higher than "normal." A second limitation results from the absence of a "no-treatment" control group, although it would seem reasonable to expect that comparisons between treatment and no-treatment control should exceed those found between training and orientation groups. Third, all groups were taught by the experimenter rather than the classroom teachers for methodological reasons; hence there is no evidence that the strategy instruction could be incorporated easily into the classroom academic curriculum. Finally, even with the extensive norming procedures, material effects were suggested by the patterns of performance on text implicit questions in the sixth grade passages, and by the difference in performance level of the fourth grade students on the two fourth-grade passages.

A similar study could be conducted using unfamiliar materials to further examine the role of familiarity, particularly with reference to performance on script questions, and the degree of students' reliance on the text. Additional factors to be considered could include the effect of text access (being able to look back or not), increased levels of training

(as well as the use of a no-treatment control group), and differential levels of feedback. Another logical step would be to train teachers to use the techniques as part of their ongoing curricula in reading, social studies, or science classes. It is one thing to demonstrate the power of a factor; it is quite another to demonstrate that it can be easily incorporated into curricula that already suffer from an excess of demand.

References

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. Bower (Ed.), Psychology of learning and motivation (Vol. 9). New York: Academic Press, 1975.

Andre', M. D. A., & Anderson, T. H. The development of a self-questioning study technique. Reading Research Quarterly, 1978-79, 14, 605-623.

Barrett, T. C. Taxonomy of reading comprehension. In R. Smith & T. C. Barrett (Eds.), Teaching reading in the middle grades. Reading, Mass.: Addison-Wesley, 1976.

Bartolome, P. I. Teachers' objects and questions in primary reading. The Reading Teacher, 1969, 23, 27-33.

Brown, A. L. Learning how to learn from text. In J. Langer & M. Smith-Burke (Eds.), Reader meets author-bridging the gap: A psycholinguistic perspective. Newark, Del.: International Reading Association, 1981.

Brown, A. L., & Campione, J. C. Memory strategies in learning: Training children to study strategically (Tech. Rep. No. 22). Urbana: University of Illinois, Center for the Study of Reading, March 1977.

Brown, A. L., Campione, J. C., & Day, J. D. Learning to learn: On training students to learn from texts. Educational Researcher, 1981, 10, 14-21.

Brown, A. L., Campione, J. C., & Murphy, M. D. Maintenance and generalization of trained metamnemonic awareness of educable retarded children. Journal of Experimental Child Psychology, 1977, 49, 191-211.

Chi, M. T. H. Knowledge development and memory performance. In M. Friedman, J. P. Das, & N. O'Conner (Eds.), Intelligence and learning. New York: Plenum Press, 1980.

Chou-Hare, V., & Pulliam, C. A. Teacher questioning: A verification and extension. Journal of Reading Behavior, 1980, 12, 69-72.

Day, J. D. Teaching summarization skills: A comparison of training methods. Unpublished doctoral dissertation, University of Illinois, August 1980.

Flavell, J. H. Developmental studies of mediated memory. In H. W. Reese & L. P. Lipsitt (Eds.), Advances in child development and behavior (Vol. 5). New York: Academic Press, 1970.

Frase, L. T. Questions as aids to reading: Some research and theory. American Educational Research Journal, 1968, 5, 319-332.

Golinkoff, R. M. A. A comparison of reading comprehension processes in good and poor comprehenders. Reading Research Quarterly, 1975-76, 11, 623-659.

Gordon, C. J. The effects of instruction in metacomprehension and inferencing on children's comprehension abilities. Unpublished doctoral dissertation, University of Minnesota, March 1980.

Guszak, F. J. Relationship between teacher practice and knowledge of reading theory in selected classes (Project Rep. No. 5-437). Washington, D.C.: United States Department of Health, Education, & Welfare, 1966.

Guthrie, J. T., & Tyler, S. J. Psycholinguistic processing in reading and listening among good and poor readers. Journal of Reading Behavior, 1976, 8, 415-426.

Hansen, J. The effects of inference training and practice on young children's comprehension. Reading Research Quarterly, 1981, 16, 391-417.

Kavale, K., & Schreiner, R. The reading processes of above average and average readers: A comparison of the use of reasoning strategies in responding to standardized comprehension measures. Reading Research Quarterly, 1979-80, 15, 102-127.

Lehnert, W. G. The processes of question answering. Hillsdale, N.J.: Erlbaum, 1979.

Morse, J. M. Effect of reader-generated questions on learning from prose. In W. D. Miller & G. H. McNinch (Eds.), Reflections and investigations on reading. Clemons, S.C.: National Reading Conference, 1976.

Olshavsky, J. E. Reading as problem-solving: An investigation of strategies. Reading Research Quarterly, 1976-77, 12, 654-674.

Pearson, P. D., & Johnson, D. D. Teaching reading comprehension. New York: Holt, Rinehart, & Winston, 1978.

Raphael, T. E. The effect of metacognitive training on children's question answering behavior. Unpublished doctoral dissertation, University of Illinois, 1980.

Raphael, T. E., Winograd, P., & Pearson, P. D. Strategies children use when answering questions. In M. L. Kamil & A. J. Moe (Eds.), Perspectives on reading research and instruction. Washington, D.C.: National Reading Conference, 1980.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Ryan, E. B. Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), Advances in reading research, 2. New York: Academic Press, 1981.

Smith, A. E. The effectiveness of training students to generate their own questions prior to reading. In P. L. Nacke (Ed.), Diversity in mature reading: Theory and research. Boone, N.C.: National Reading Conference, 1973.

Taylor, B. M Children's memory for expository text after reading. Reading Research Quarterly, 1980, 15, 399-412.

Tierney, R. J., Bridge, C., & Cera, M. J. The discourse processing operations of children. Reading Research Quarterly, 1978-79, 14, 539-573.

Washington, E. D. The classification of reading comprehension test items and its relation to the performance of selected racial groups. Unpublished doctoral dissertation, University of Iowa, July 1979.

Weinstein, R., & Rabinovitch, M. S. Sentence structure and retention in good and poor readers. Journal of Educational Psychology, 1971, 62, 25-30.

Wixson, K. The effects of postreading questions on children's discourse comprehension and knowledge acquisition. Unpublished doctoral dissertation, Syracuse University, August 1980.

Table 1

Subject Placement, Experiment II

Ability	Grade			Totals
	4th	6th	8th	
Low-Average	--	10 Training 10 Orientation	--	20
Average	10 Training 10 Orientation	10 Training 10 Orientation	10 Training 10 Orientation	60
High	--	10 Training 10 Orientation	--	20
Totals	20	60	20	--

Table 2

Table of Means of Proportion of Hits: Experiment II

Factor/Level of Factor	Grade Comparison		Ability Comparison	
	Common Passage	Reading Level Passage	Common Passage	Reading Level Passage
Ability				
High	--	--	.71**	.74**
Average	--	--	.60	.60
Low-Average	--	--	.50	.57
Grade				
4th	.60*	.61	--	--
6th	.69	.60	--	--
8th	.69	.68	--	--
Treatment				
Training	.75**	.69**	.67**	.68*
Orientation	.57	.58	.50	.60
Question Type				
Text Explicit	.73**	.65*	.63	.69**
Text Implicit	.57	.57	.60	.56
Script Implicit	.69	.67	.59	.66

*Significant difference, $p < .05$

**Significant difference, $p < .01$

Table 3

Table of Means of Matches: Experiment II

Factor/Level of Factor	Grade Comparison		Ability Comparison	
	Common Passage	Reading Level Passage	Common Passage	Reading Level Passage
Ability				
High	--	--	.89**	.89**
Average	--	--	.84	.84
Low-Average	--	--	.71	.78
Grade				
4th	.89	.88	--	--
6th	.88	.84	--	--
8th	.90	.83	--	--
Treatment				
Training	.93**	.87	.88**	.86*
Orientation	.84	.82	.74	.80
Question Type				
Text Explicit	.93**	.87**	.85**	.88**
Text Implicit	.88	.88	.87	.86
Script Implicit	.84	.79	.71	.75

*Significant column difference, $p < .05$

**$p < .01$

Figures

Figure 1. Graph of the treatment X question type interaction for the proportion of Hits for the Grade Comparison--Common Passage.

Figure 2. Graph of the ability X treatment X question type interaction for the proportion of Hits for the Ability Comparison--Common Passage.

Figure 3. Graph of the ability X question type interaction for the proportion of Hits for the Ability Comparison--Reading Level Passage Set.

Figure 4. Graph of the grade X question type interaction for the proportion of correct responses for the Grade Comparison--Reading Level Passage Set.

Figure 5. Graph of the ability X question type interaction for the proportion of correct responses for the Ability Comparison--Common Passage.

Figure 6. Graph of the ability X treatment X question type interaction for the proportion of correct responses for the Ability Comparison--Common Passage.

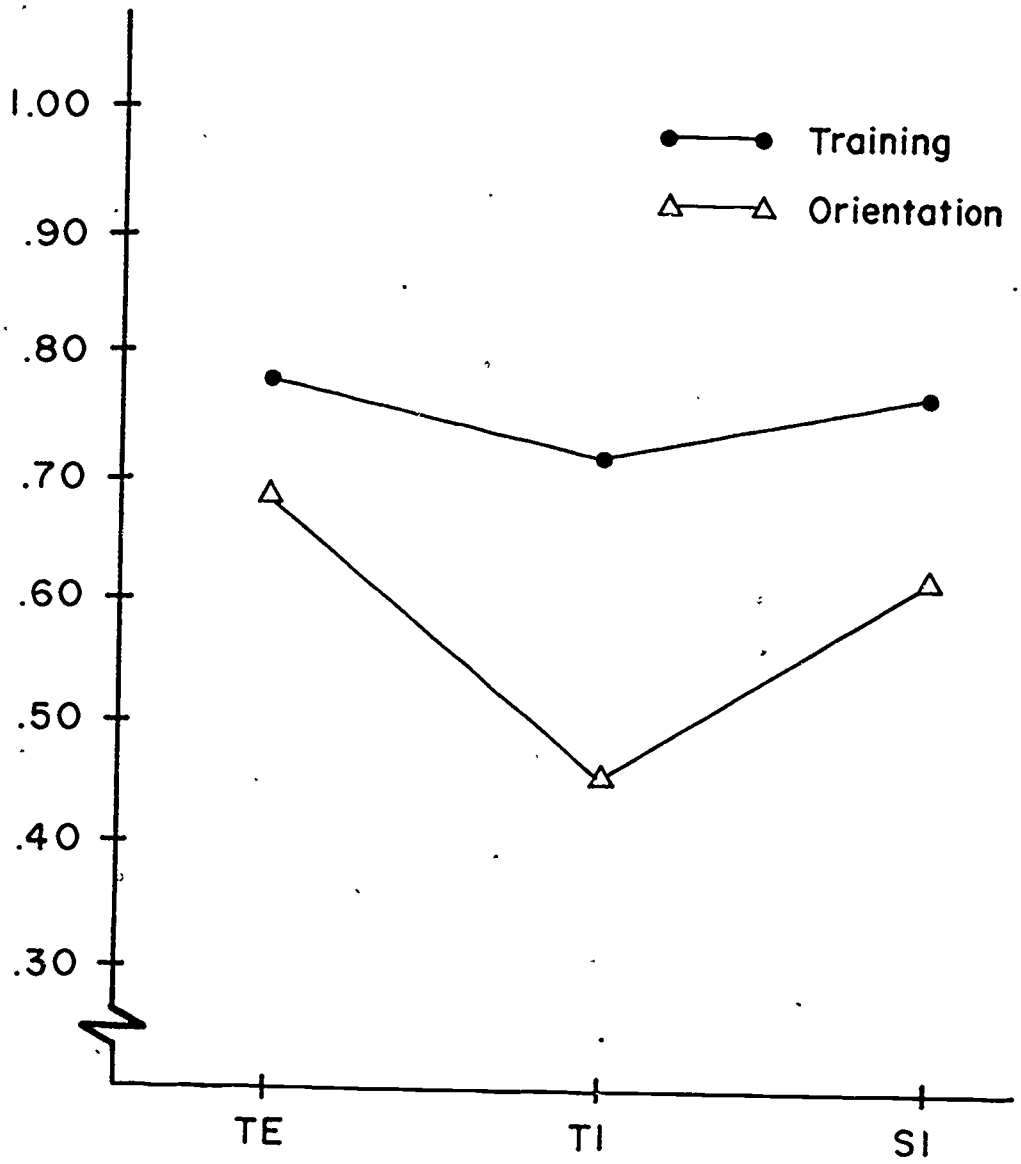
Figure 7. Graph of the ability X question type interaction for the proportion of correct responses in the Ability Comparison--Reading Level Passage Set.

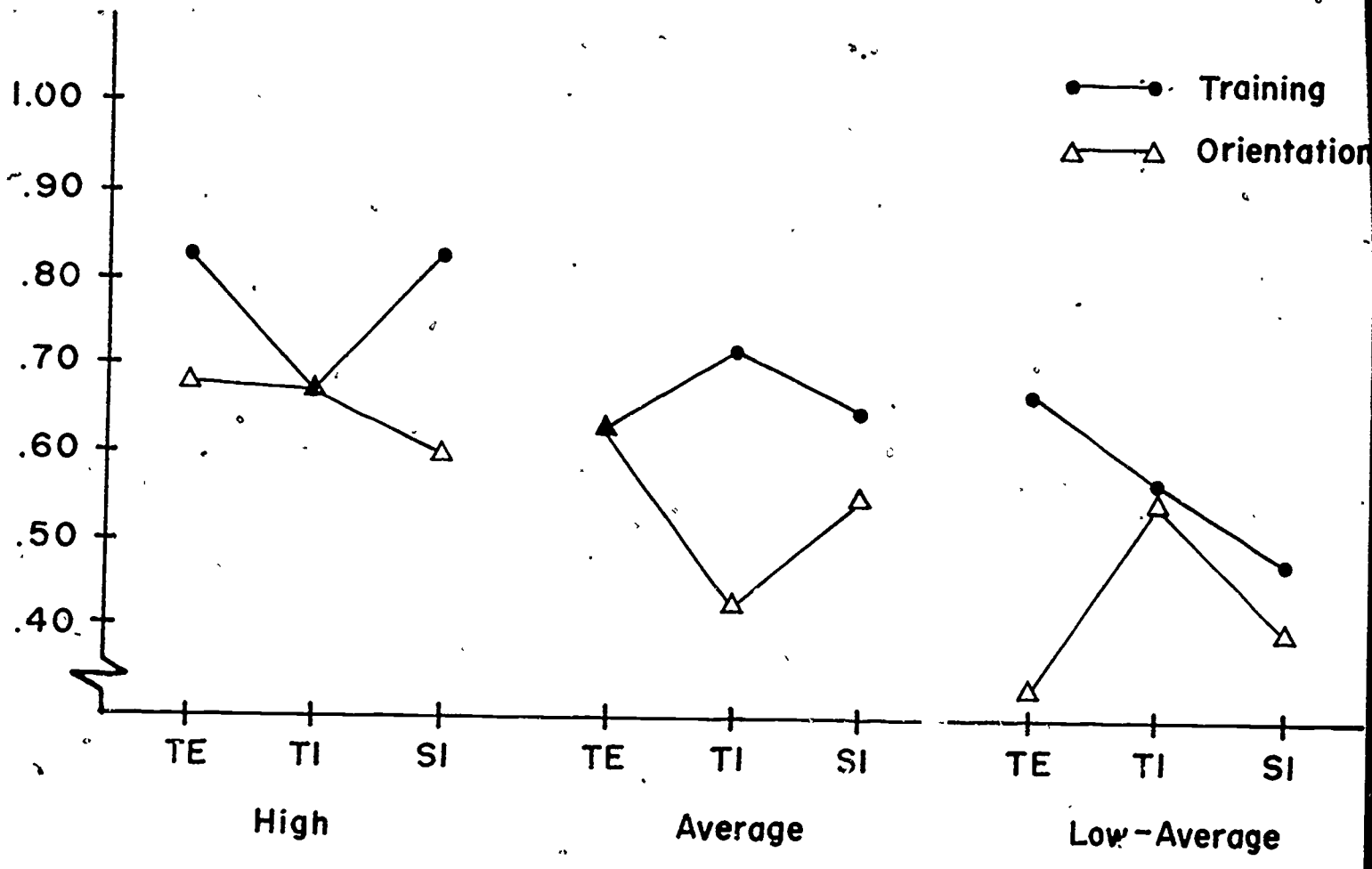
Figure 8. Graph of the treatment X question type interaction for the proportion of correct Hit-Match-Response Quality for the Grade Comparison--Common Passage.

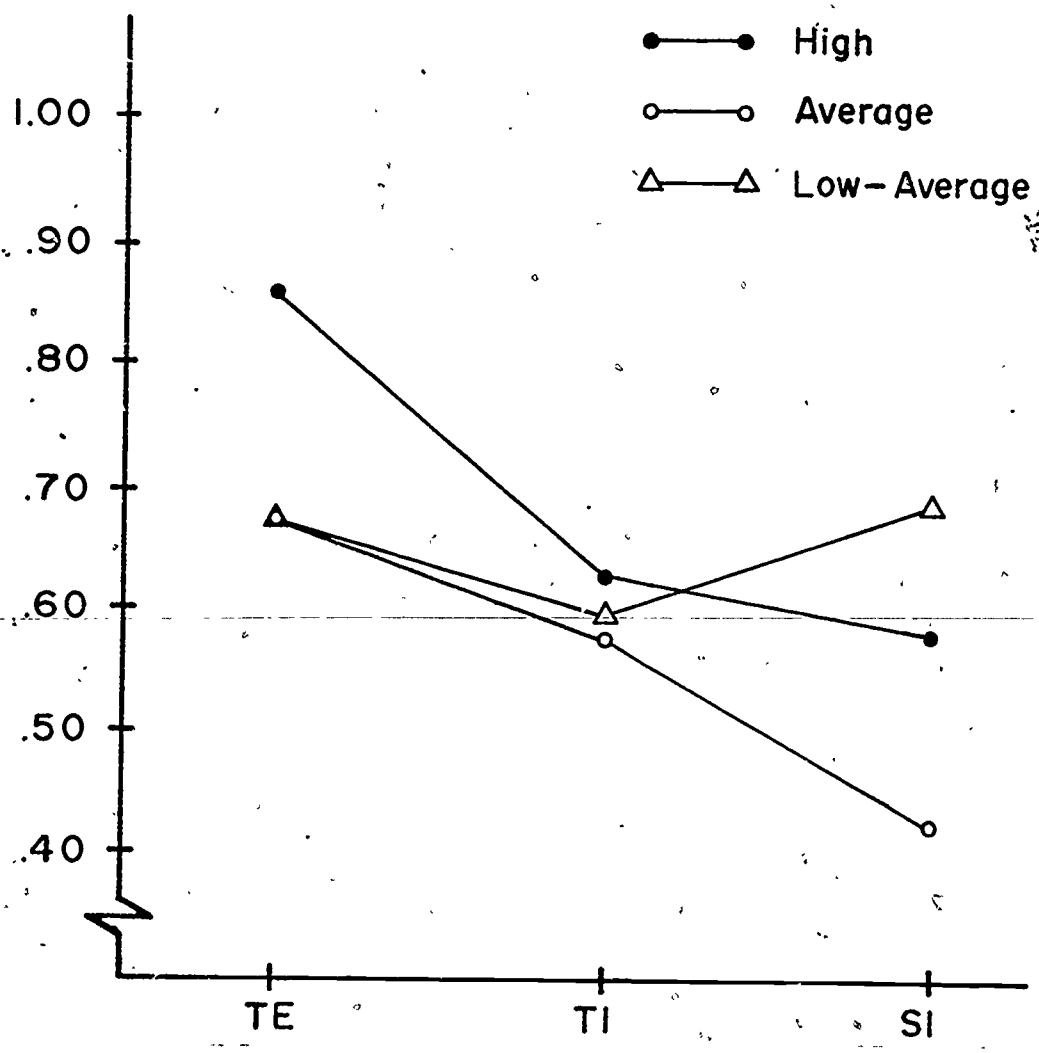
Figure 9. Graph of the treatment X question type interaction for the proportion of correct Hit-Match-Response Quality for the Grade Comparison--Reading Level Passage Set.

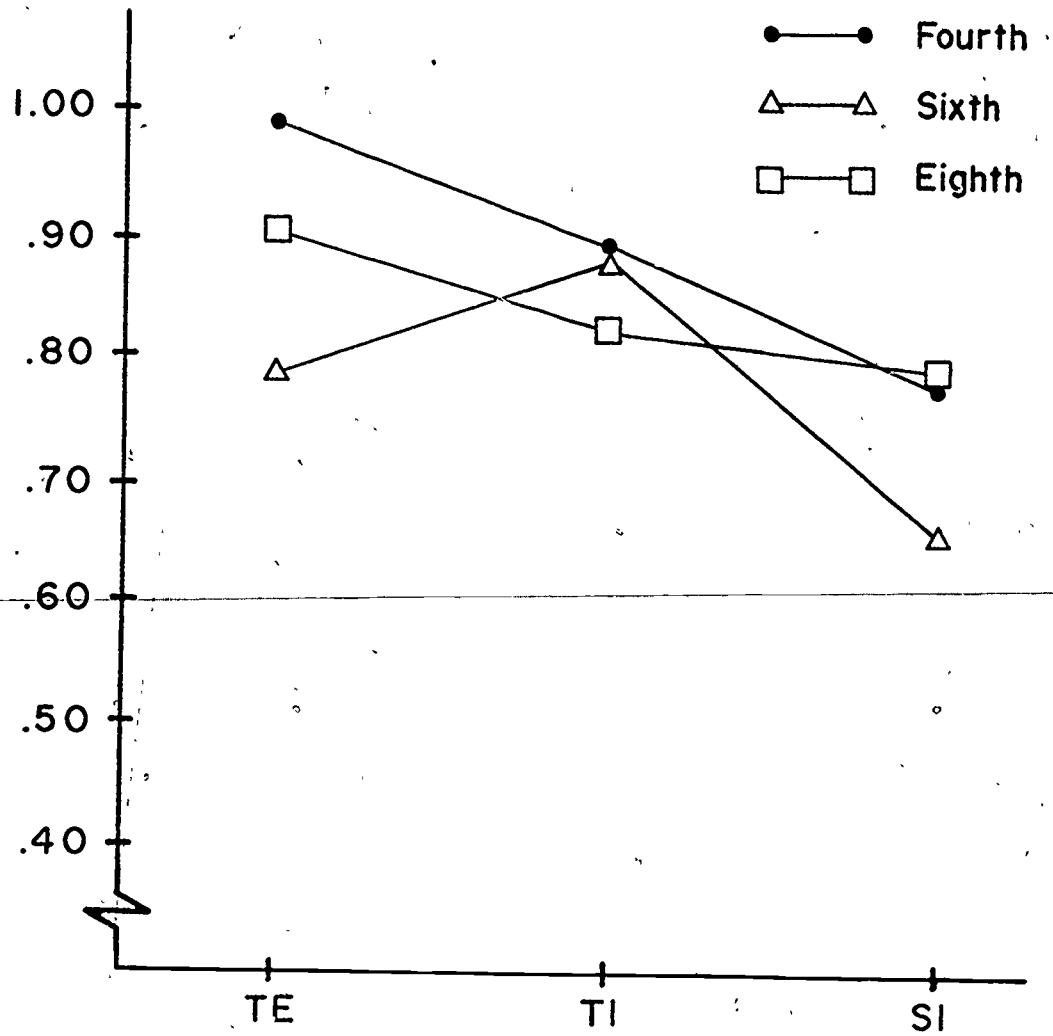
Figure 10. Graph of the ability X treatment X question type interaction for the proportion of correct Hit-Match-Response Quality for the Ability Comparison--Common Passage.

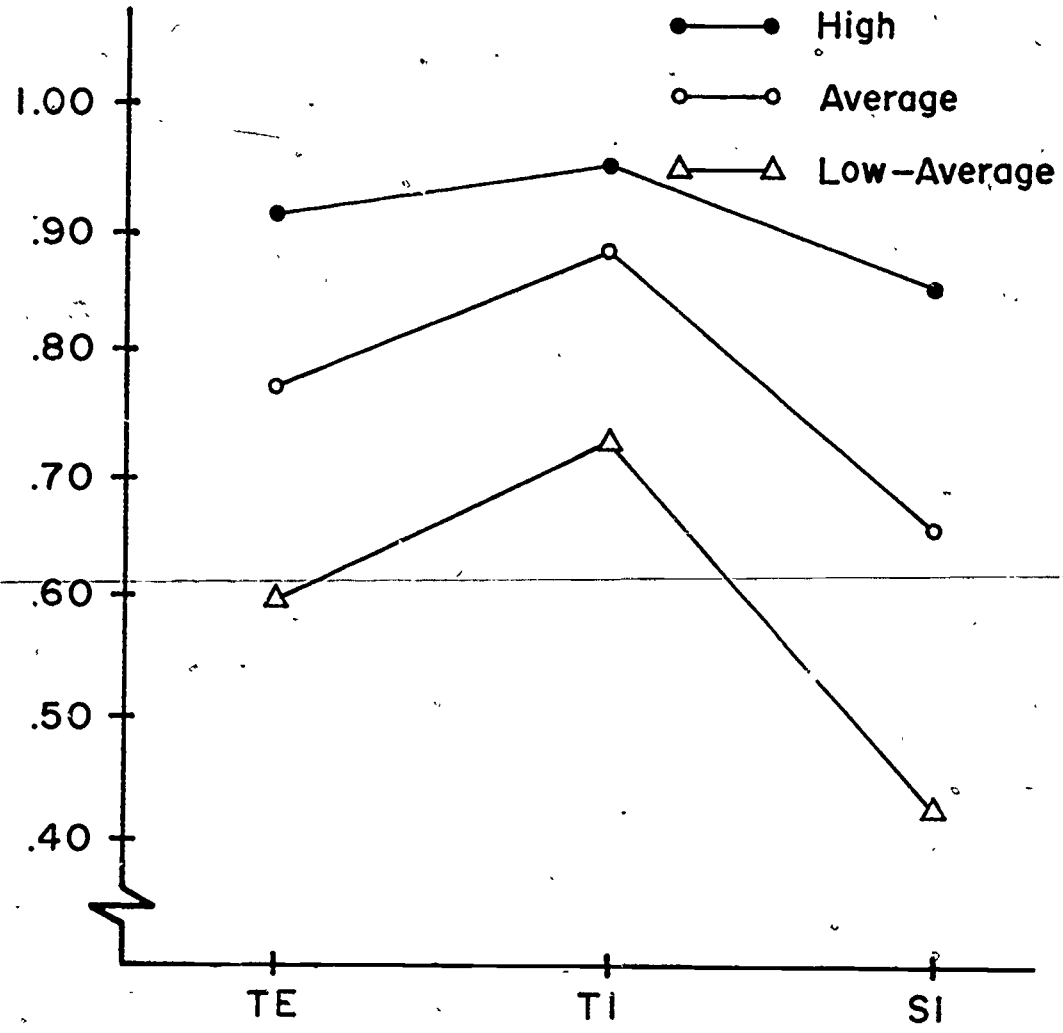
Figure 11. Graph of the ability X question type interaction for the proportion of correct Hit-Match-Response Quality for the Grade Comparison--Reading Level Passage Set.

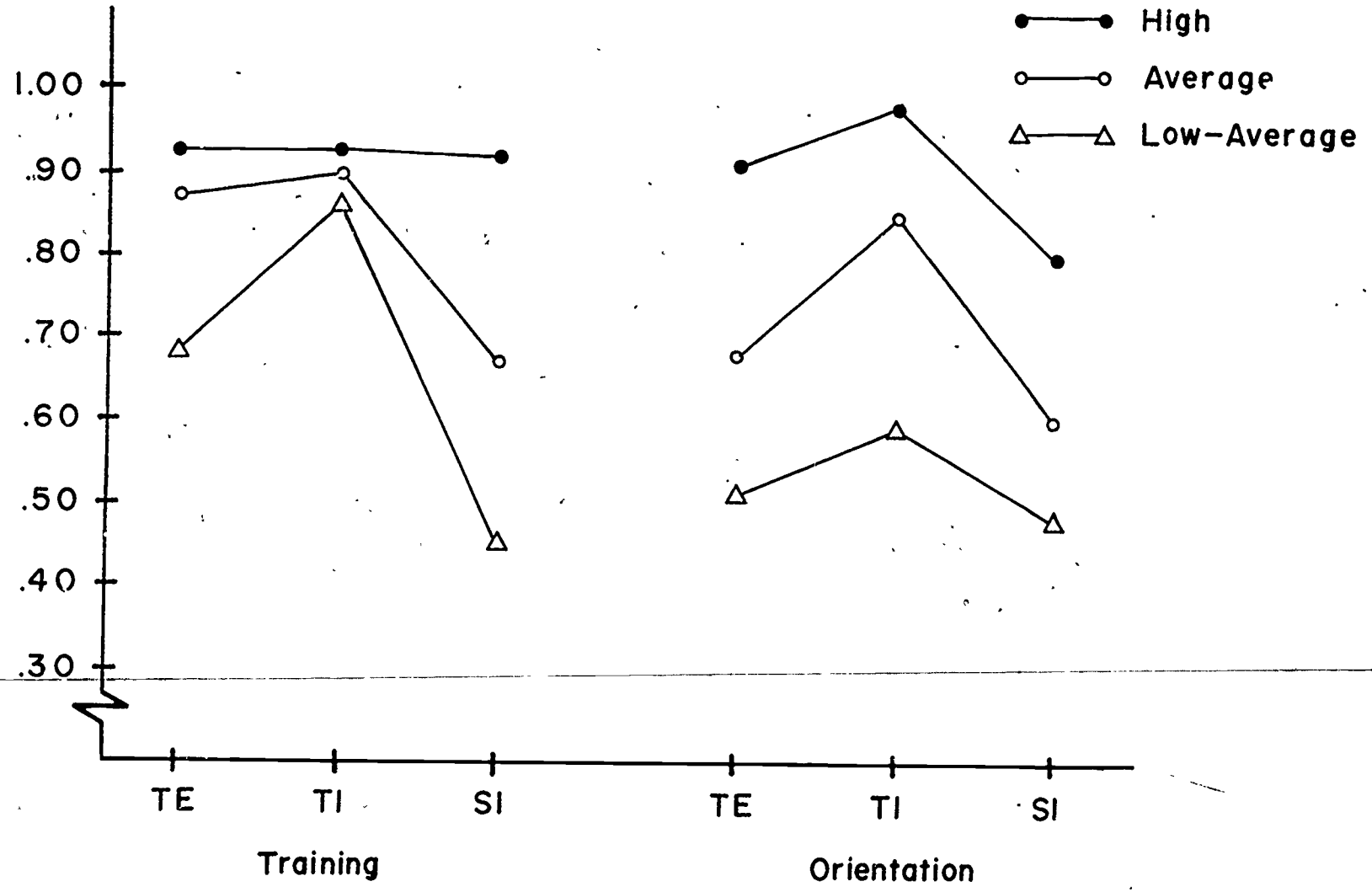

1.00
.90
.80
.70
.60
.50
.40
.30
Training
Orientation
TE
TI
SI

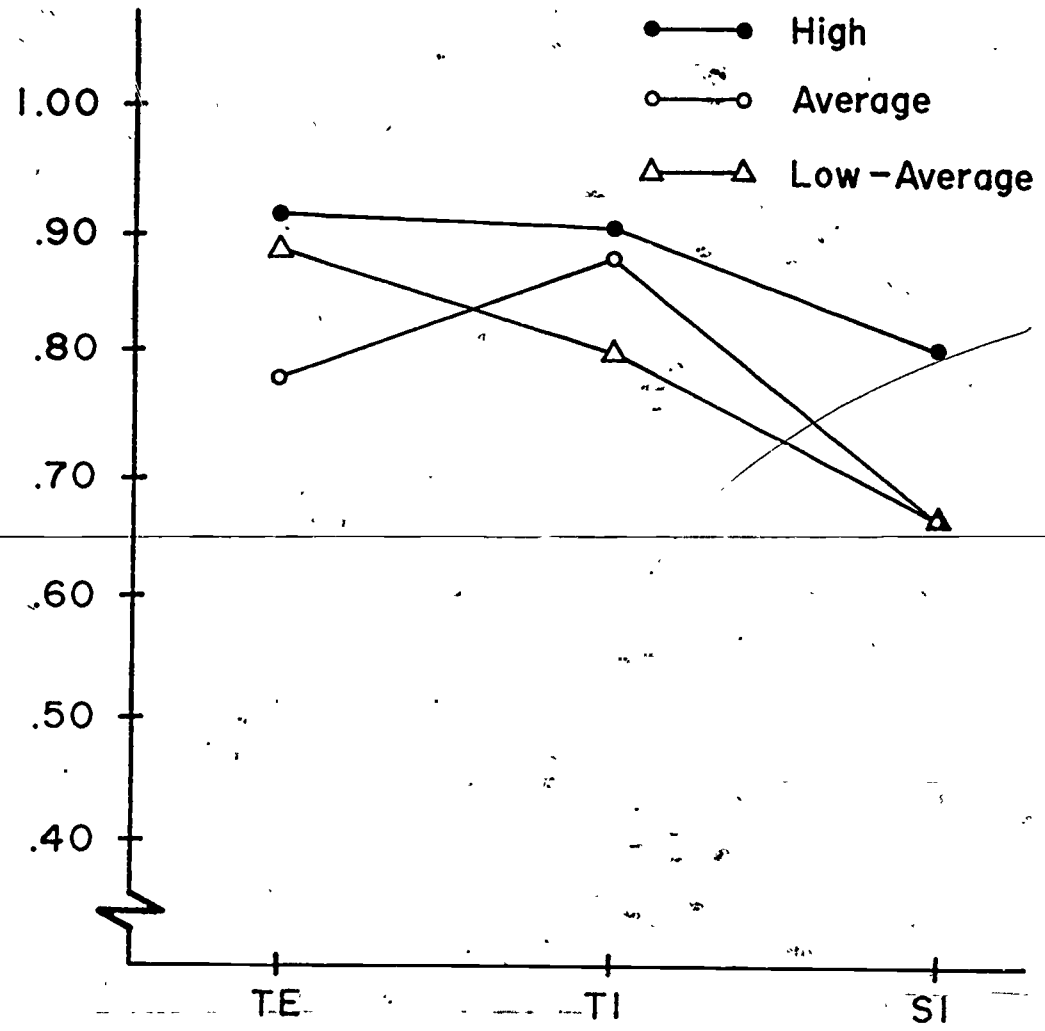

1.00
.90
.80
.70
.60
.50
.40
Training
Orientation
TE
TI
SI
TE
TI
SI
TE
TI
SI
High
Average
Low-Average

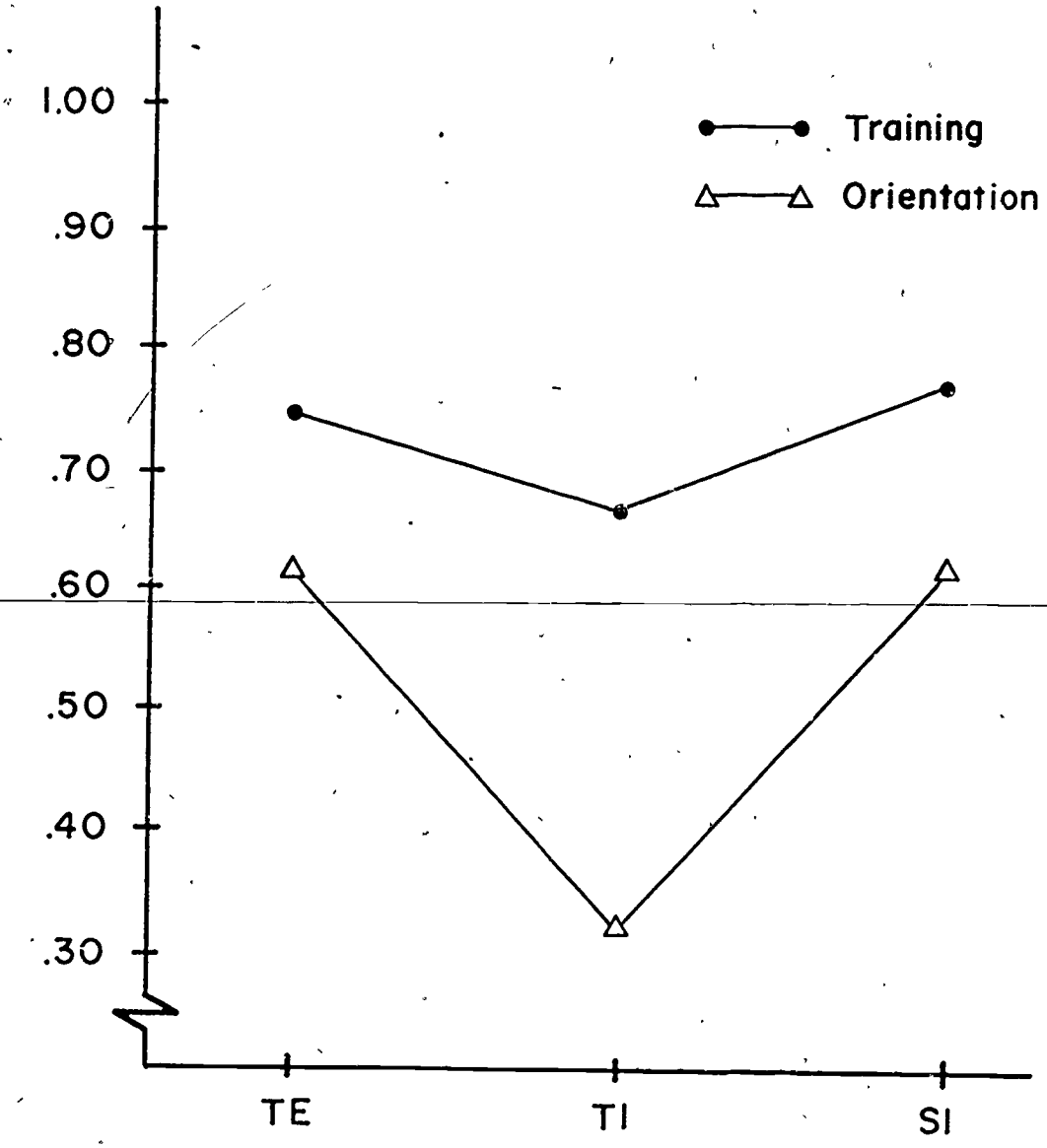

High
Average
Low-Average
1.00
.90
.80
.70
.60
.50
.40
TE
TI
SI

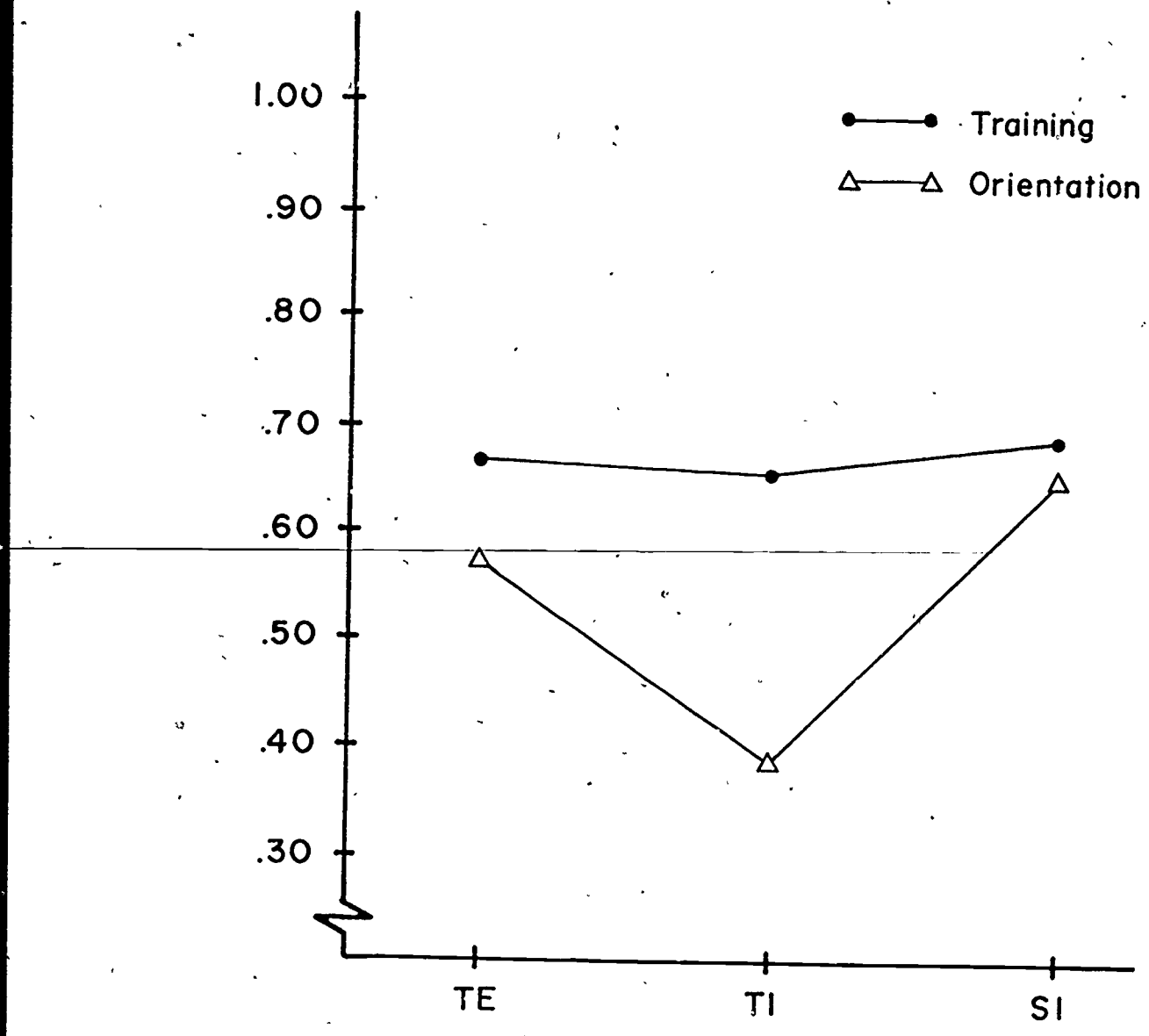

Fourth
Sixth
Eighth
1.00
.90
.80
.70
.60
.50
.40
TE
TI
SI

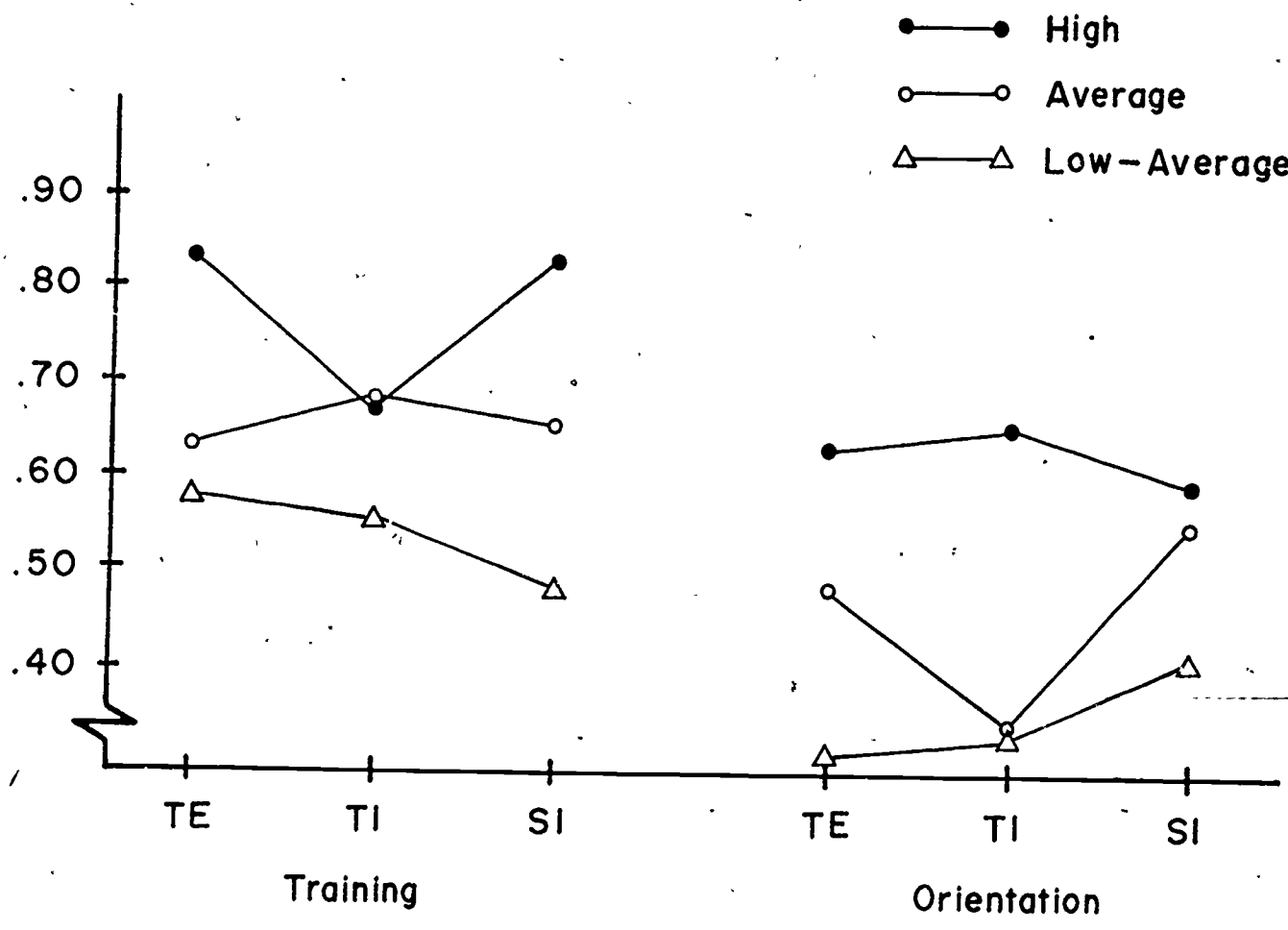

High
Average
Low-Average
1.00
.90
.80
.70
.60
.50
.40
TE
TI
SI

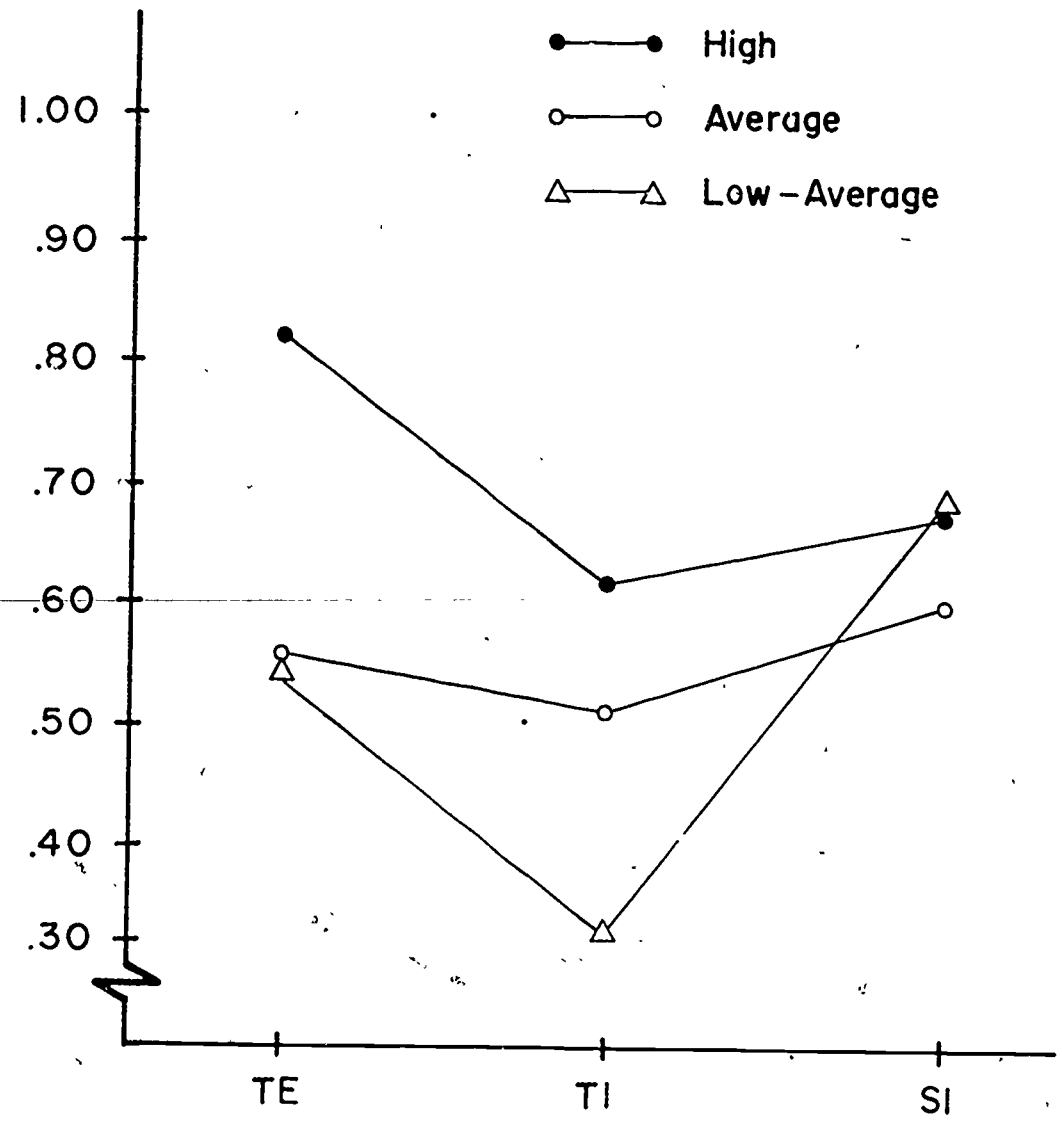

High
Average
Low-Average
1.00
.90
.80
.70
.60
.50
.40
.30
TE
TI
SI
TE
TI
SI
Training
Orientation


High
Average
Low-Average
1.00
.90
.80
.70
.60
.50
.40
TE
TI
SI


1.00
.90
.80
.70
.60
.50
.40
.30
Training
Orientation
TE
TI
SI


1.00
.90
.80
.70
.60
.50
.40
.30
Training
Orientation
TE
TI
SI


High
Average
Low-Average
.90
.80
.70
.60
.50
.40
TE
TI
SI
TE
TI
SI
Training
Orientation


High
Average
Low – Average
1.00
.90
.80
.70
.60
.50
.40
.30
TE
TI
SI