

The Influence of Analysis and Evaluation Questions on Achievement in Sixth Grade Social Studies

FRANCIS P. HUNKINS
Assistant Professor of Education
University of Washington, Seattle

Introduction

QUESTIONS used by teachers in their discourse and those incorporated in instructional materials probably are significant in guiding the development of pupils' levels of knowledge and achievement. Questions reveal the operational objectives which stress, for example, the increase of pupils' knowledge of facts, of understandings, of concepts, and of pupils' skills at interpreting information and ideas.

The classroom teacher devotes a large portion of his time to asking questions, e.g., Adams (1), Aschner (2), Barr (3), Floyd (7), Stevens (9). The usefulness of questions has long been recognized as significant in the teaching-learning interaction. Yet, even with this purported awareness of the importance of questions, little has been done with regard to the effects questions have upon pupils' achievement within the class situation.

This research was concerned with determining the relative effectiveness of knowledge, analysis, and evaluation

questions in stimulating achievement in sixth grade social studies. These question types were based on three of the six hierarchical categories of Benjamin Bloom's *Taxonomy of Educational Objectives*, Bloom (4).

Knowledge questions require the recall of ideas, facts, materials, or phenomena. They call for the releasing of certain information stored in the individual's memory. Analysis questions, on the other hand, demand the arrangement and rearrangement of information into elements, relationships, and organizations. The third type of question, evaluation, requires a judgment employing criteria such as accuracy, effectiveness, economic quality, or satisfying quality. These two latter question types subsume knowledge. The evaluation type of question subsumes both knowledge and analysis.

Objectives

This study sought to determine whether a dominant use in social studies text-type materials of analysis and evaluation questions, as defined

by Bloom's *Taxonomy*, would effectively stimulate the development of sixth grade pupils' social studies achievement. The overall hypothesis tested, stated in null form was:

Use of text-type materials employing questions requiring "analysis" and "evaluation" will not result in differences in sixth grade pupils' social studies achievement when compared with the use of text-type materials incorporating questions requiring the recall of knowledge in relationship to pupils' (a) reading level, (b) sex, and (c) the interaction between these variables.

General Plan of the Study

The general plan of the study first involved constructing two sets of text-type materials and corresponding answer sheets, one set stressing questions requiring analysis and evaluation (Condition A) and the other stressing questions requiring knowledge (Condition B). Pupils in both treatment conditions were directed to read designated sections of their textbook and to respond in writing to the questions on their worksheets.

For four weeks, pupils used these materials during a thirty-five minute portion of the daily social studies period. For this study, the instructional unit dealt with Africa and Oceania and was based on chapters in the adopted social studies textbook used by the cooperating school system. The general format and directions of the two sets of special materials were identical, the only varying factor being the questions and their emphases. During the experimental period, teachers refrained from actively engaging in teaching but assisted in coordinating the pupils' use of

the materials. This lack of active teacher participation was an attempt to reduce their influence on the experimental situation.

Procedure

Subjects

Two hundred and sixty pupils served as subjects in this study. They were enrolled in eleven sixth grade classes in three elementary schools serving the same geographic area of a large suburban public school system in a northeastern Ohio community (population, 47,922, 1960 census). The eleven classes were randomly assigned to one of two experimental treatment conditions, A or B. A total of 127 pupils (67 boys and 60 girls) were assigned to Condition A, while 133 pupils (55 boys and 78 girls) were assigned to Condition B.

Background data were collected and analyzed for both pupils and teachers.

Intelligence quotients were obtained from the *California Test of Mental Maturity*, S Form. The mean IQ for boys in Treatment A was 114.84, SD 16.06 ($N = 67$). Girls in the same treatment had a mean IQ of 115.80, SD 15.10, ($N = 60$). In Treatment B, the mean IQ for boys was 112.07, SD 15.38 ($N = 55$), and the mean IQ for girls was 113.11, SD of 12.65 ($N = 78$).

Reading test scores were obtained on the subjects from the *Stanford Achievement Test*, Form W. Boys in Treatment A had a mean reading score of 38.90, SD 12.18; girls in the same treatment had a mean reading score of 41.86, SD 11.45. Boys in Treatment B had a mean reading score of 36.45, SD 10.39, while girls in the same treat-

ment had a mean reading score of 39.29, SD 9.74.

Pupils' IQ scores and reading scores were subjected to analysis of variance to determine if significant differences were present. For this analysis, scores were considered across reading levels by treatment and sex.

No significant differences in IQ were revealed between the two treatments, between boys and girls, or their interaction. Thus IQ was eliminated as a possible covariant on subsequent analyses of the criterion data. No significant differences between reading achievement were noted between the two treatments and no significant interaction either. A significant difference between the reading scores of boys and girls was observed. However, this relationship was not considered sufficient reason to use the reading scores as covariants on subsequent analyses of criterion data. Had significant differences in reading achievement existed between treatments, such scores would have been employed as covariants.

Background data on the participating teachers revealed that teachers of classes in both conditions were similar with respect to age and teaching experience.

The mean age for teachers with pupils involved in Treatment A was 35.16, while the mean age for those teachers with pupils involved in Treatment B was 33.20. Experience in teaching was similar, with means of 8.33 years for teachers in Treatment A and 8.80 years for teachers in Treatment B. The Treatment A teachers had slightly more years in teaching the sixth grade, mean 5.66 years, than did the teachers in Treatment B, mean 3.80 years. Ten

of the eleven teachers involved had Bachelor degrees. The one teacher without the degree, in Treatment A, had over 24 years of teaching experience. Of the total teacher group, only one teacher had a Master's degree, in Treatment B.

Collection of Data

A criterion test of achievement covering the selected social studies unit was constructed by the investigator. The objective of the test design was to afford a single achievement score in addition to six subscores corresponding to the six categories in Bloom's *Taxonomy*: knowledge, comprehension, application, analysis, synthesis, and evaluation. However, only the total achievement score was of concern in this phase of the investigation.

A total of 59 multiple-choice, four option items was written and submitted to two judges well-acquainted with the *Taxonomy*. With four exceptions, only items having 100 percent agreement among the judges and investigator regarding *Taxonomy* emphasis were selected for the final test which contained 42 items. The test contained seven questions in each of the six Bloom categories. Reliability of the post-test was determined to be .68 using the Kuder-Richardson formula. This reliability is quite low and the reader should bear this in mind when considering the results.

Experimental Material and Procedure

Pupils in both experimental treatments used the regularly adopted social studies text, *The Changing Old World*, by Cooper, Sorensen, and Todd (5).

The unit for study during the experiment was "Africa, Australia and New Zealand." This unit was deemed appropriate, for the investigator felt that the subjects would not bring abundant prior knowledge into the experimental situation.

Experimental pupil materials. For each experimental treatment condition, special materials and answer sheets were constructed. The sets for both conditions were identical in format and directions. Seventeen sets for each condition were developed to correspond to discrete portions of the adopted text. Both treatment groups experienced an identical introductory set to familiarize them with the materials. Condition A materials had 47.53 percent of the total questions in the analysis and evaluation categories: 28.41 percent analysis, and 19.12 percent evaluation. Condition B materials contained a question emphasis on knowledge of 87.38 percent of the total number.

These materials had been submitted to the same judges who served to classify the achievement test items. Every question of each set was categorized by the judges, and, following this, the two individuals met with the investigator to compare results. If a question emphasis could not be agreed upon after discussion, the question was either rewritten until the desired emphasis was obtained or not employed in the final draft of the materials. This procedure was followed in order to ascertain that these question emphases did exist.

A readability analysis, employing the Dale-Chall formula, Dale and Chall (6), was used to determine the reading

level of both the experimental materials and the answer sheets in both Conditions A and B. Results of this analysis revealed that the experimental materials had an average raw score of 5.64, designating a reading level well within the range of fifth and sixth grade pupils. Analysis of the answer sheets revealed an average raw score of 5.65, also designating the material as appropriate for sixth graders.

Working with the materials. Pupils in both conditions were instructed to work independently with the materials. They were given from 30 to 35 minutes each day to work with the experimental sets, reading carefully and writing their answers in the provided spaces. Pupils were provided time to check their work with the answer sheets.

Analysis of Data

The experimental design basic to this study was an analysis of covariance design. Within each treatment condition, data were analyzed according to sex and reading achievement. This resulted in a 2 x 2 x 4 (treatment x sex x reading level) classificatory scheme. Pupils were assigned reading levels according to quartile ranks as determined by their raw scores on the reading achievement test (*Stanford Achievement Test, Form W*). The four levels had the following ranges: Quartile 1, 0-31; Quartile 2, 32-38; Quartile 3, 39-47; and Quartile 4, 48-64.

Results

The post-test achievement scores were subjected to analysis of covariance adjusting for the pre-achievement

scores. These results are summarized in Table 1.

No statistically significant differences in achievement between boys and girls were observed, and there were no significant interactions. Condition A pupils achieved more than did Condition B pupils, and better readers obtained higher achievement scores than did poorer readers.

The means for between reading levels were contrasted by a series of *t* tests, all of which revealed statistically significant differences ($Q_4 > Q_3 > Q_2 > Q_1$). It appears from these data that whether one can handle high-level questions is related to how well one can read. The higher-level questions usually were more involved with regard to wording than were the knowledge questions. Also, it should be borne in mind

that both worksheets and answer sheets placed demands upon pupils' reading skills.

Conclusions

From the analyses conducted, the following conclusions are warranted:

1. The employment of high cognitive-level questions (analysis and evaluation) produced significantly greater scores in social studies achievement than did low cognitive-level questions (knowledge).

2. Better readers in both conditions achieved higher than did poorer readers.

As a result of these significant differences, the major null hypothesis relating to social studies achievement was rejected.

Source of Variation	Original		Adjusted			F
	d.f.	S.S.	d.f.	S.S.	M.S.	
Treatment	1	12.06	1	10.05	10.05	9.85**
Reading level	3	133.55	3	65.02	21.67	21.24**
Sex	1	.92	1	.92	.92	.90
Treatment x reading level	3	.49	3	.23	.07	.06
Treatment x sex	1	.75	1	.26	.26	.25
Reading level x sex	3	2.29	3	.63	.21	.20
Treatment x reading level x sex	3	8.06	3	7.21	2.40	2.35
Within groups	244	269.88	243	247.98	1.02	

** Significant at the .01 level.

Table 1. Summary of Analysis of Covariance of Post-test Scores on Achievement Test

Discussion

Analysis of the data suggests that questions requiring analysis and evaluation stimulated individuals to utilize several viewpoints regarding the information embedded in the task. It seems reasonable that pupils, using such questions, might have been forced to engage in the intellectual activity of considering various aspects of factual knowledge and evaluating the complexity, implications, and applications of such knowledge. Such mental "juggling" may have enabled pupils to know better the information with which they were dealing.

It is assumed that high-level questions will demand of individuals more intellectual activity than would be true with low-level questions. Since the data revealed a significantly greater achievement among the pupils receiving such high-level questions, one might cautiously conclude that the pupils did in fact react more actively with information presented. The evidence seems to suggest, rather than confirm, that pupils were engaged in an interaction with the materials presented. Analysis of the sub-tests of the achievement test, presently being conducted in a second phase of the research, may provide evidence to clarify somewhat exactly how the high-level questions affected various dimensions of achievement.

That the evidence is suggestive rather than confirming should not be cause for alarm. This research is a beginning, not a conclusion regarding the effects of questions. It is premature to make definite statements. These types of questions do have the potential, it seems, to make pupils uneasy, but also

to encourage them to probe their knowledge and to discover meanings.

Implications of the Study

Several implications may be drawn from this study regarding the role of questions in relation to social studies achievement in particular and overall school achievement in general. If questions at higher-cognitive levels are capable of stimulating high achievement, then teachers should be using these questions in much greater numbers than they currently do. Teachers, by improving their level of questioning, could very well make information more meaningful for their pupils. In addition, pupils in classrooms where high-level questions are used by teachers should be expected to employ such questions themselves when they engage in class discussions and other class work. Higher-level questions not only should stimulate higher levels of achievement, but also should make pupils better inquirers into the realm of knowledge.

References

1. Thomas Howard Adams. *The Development of a Method for Analysis of Questions Asked by Teachers in Classroom Discourse*. Doctor's thesis. New Brunswick, New Jersey: Rutgers, The State University, 1964. 149 pp.
2. Mary Jane McCue Aschner. "Asking Questions to Trigger Thinking." *NEA Journal* 50 (6); 44-46; September 1961.
3. A. S. Barr. *Characteristic Differences in the Teaching Performance of Good and Poor Teachers of the Social Studies*. Bloomington, Illinois: Public School Publishing Co., 1929. 127 pp.
4. Benjamin S. Bloom, editor. *Taxonomy of Educational Objectives. The Classification of Educational Goals: Handbook I Cognitive*

Domain. New York: David McKay Company, Inc., 1956. 207 pp.

5. Kenneth S. Cooper, Clarence W. Sorensen, and Lewis Paul Todd. *The Changing Old World*. Morristown, New Jersey: Silver Burdett Company, Inc., 1961. 470 pp.

6. Edgar Dale, and Jeanne S. Chall. "A Formula for Predicting Readability." *Educational Research Bulletin* 27: 11-20, 28; January 1948. 27: 37-54; February 1948.

7. William D. Floyd. *An Analysis of the Oral Questioning Activity in Selected Colorado Primary Classrooms*. Doctor's thesis. Greeley: Colorado State College, 1960. 195 pp.

8. Frank M. Gatto. *Pupils' Questions: Their Nature and Their Relationship to the Study Process*. Doctor's thesis. Pittsburgh, Pennsylvania: University of Pittsburgh, 1928. 158 pp.

9. Romiett Stevens. *The Question as a Measure of Efficiency in Instruction*. Teachers College Contribution to Education, No. 48. New York: Teachers College, Columbia University, 1912. 95 pp.

10. Soshichi Yamada. "A Study of Questioning." *Pedagogical Seminary* 20: 129-85; 1913. 5

An Evaluation of the Frostig Visual-Perceptual Training Program

JAMES N. JACOBS

Division of Program Development
Public Schools, Cincinnati, Ohio

Introduction

AUTHORITIES in the fields of education and psychology have, for a considerable period of time, recognized the critical need to identify and institute a remediation program for those children who are likely to encounter extreme difficulties in learning to read. Although there has been a general recognition of this need, productive remediation and diagnostic procedures have not as yet been adequately developed.

The traditional approach to this problem has been to obtain some global measure of a child's intellectual capacity, generally without taking notice of specific areas of strengths and weaknesses, and to determine his "reading

readiness," using instruments that rely heavily on the child's early cultural opportunities. Remediation procedures for these children have, for the most part, been absent or, if present, were based on a trial and error method.

One promising method of early prediction, diagnosis, and remediation of reading difficulties is the Marianne Frostig Developmental Test of Visual Perception, and a rather complete visual-perceptual training program to be used in conjunction with the test. The test appears to be potentially very useful inasmuch as it not only claims to predict reading success, but also gives some indications of specific areas in which the child has visual perceptual difficulties.

Copyright © 1968 by the Association for Supervision and Curriculum Development. All rights reserved.