THE EFFECTS OF OBJECTIVE GUIDE QUESTIONS AND SELF-CHECKING ANSWER SHEETS UPON PERFORMANCE IN READING AND LEARNING

by

VERGIL MILLER McINTOSH

B. S., Kansas State College
of Agriculture and Applied Science, 1938

A THESIS

submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

Department of Education

KANSAS STATE COLLEGE
OF AGRICULTURE AND APPLIED SCIENCE

1939

TABLE OF CONTENTS

	Page
PURPOSE	1
RELATED LITERATURE	1
GENERAL PROCEDURE	5
RESULTS	8
CONCLUSION	18
ACKNOWLEDGMENTS	19
LITERATURE CITED	20

PURPOSE

The purpose of this research was to find the value of objective questions and self-checking answer sheets in study of assignments and also to determine the effect of the use of objective questions and self-checking answer sheets upon the reading ability of the pupil.

RELATED LITERATURE

Washburne (13) found that the inclusion of objective questions with social science material greatly improved the understanding and recall of the material by the student. He found a difference in the mean scores of more than 5 P. E. in favor of the inclusion of the objective questions.

Symonds and Chase (12) found that the value of test motivation may be estimated as the equivalent of five sheer repetitions in studying. They concluded that the most effective method that can be applied to learning is to increase the amount of drill or practice.

Noll (7) conducted an experiment to determine the effect of written tests upon achievement of Educational Psychology students. In one class, four quizzes were given at intervals of about three weeks; in the experimental

class, no quizzes other than the mid-semester and final examinations were given. He found that students given only the mid-term and final examinations made a consistent, but not a substantially higher average achievement than those in the control group. The frequent quizzes raised the achievement of only the lower one-third of the group.

In Noll's experiment, the control group was his 1935-36 class in Educational Psychology, and the experimental group, his 1936-37 class of the same course. His method of teaching may have varied enough in the two years to affect the groups decidedly. In his experimental group, the frequent quizzes were semi-objective, and the student probably never did know the exact answer expected. The papers were marked and returned to the student, but were not discussed at all unless a question was raised by the student. This probably lost one of the main objectives of a test, that is to indicate to the student the nature of his mistakes.

Peterson (9) carried on an experiment with the self-checking sheets and objective questions in guiding students in studying General Psychology. He found that students using the self-checking sheets showed definite improvement in performance as compared with the other groups. The results were statistically valid, the differences in gain being more than four times its standard error. He found

that students using the self-checking device gained from 2.4 to 3 times as much information as did those who used only the questions as a guide.

Kellogg and Payne (4) conducted an experiment on the use of true-false questions as an aid in studying. They found that of students who were given a test over the same material used in the study guide, 89 per cent made lower grades than they did on the original questions; while of those given the same questions on the examination, 46 per cent made lower scores than they made on the original questions. This suggests a lack of generalization by the students while working at the original test.

Fleenor (2) made a study with General Psychology students as subjects in home study courses using self-checking answer sheets. His study showed that the students using the self-checking answer sheets gained 21.23 points in a 400 point final examination as compared with paired students using the regular written home study methods. This difference was found to be 3.16 times its own standard error and therefore statistically significant.

Marx (6) found that pupils using questions and self-checking answer sheets in studying, gained three times as many points in a check-up test as those not having the check-sheets. His results showed a statistically signifi-

cant gain of 6.054 times its probable error in favor of the immediate check-up technique. He also found that the tester technique in no way hindered the transfer of knowledge to new situations and problems.

Parr (8) found that a work manual devised to indicate the purpose of study, and a study outline for the student, gave improvement of from 5 to 75 per cent in the amount of information gained.

Salisbury (11) found that giving students special lessons in outlining gave decided improvement in reading, in ability to solve reasoning problems, and success in content subjects to which outlining might be applied. Keys (5) found that tests given in the form of weekly rather than of monthly examinations, showed a mean performance which was 12 per cent higher. Retention by the groups tested weekly was 7 per cent superior to that of those tested monthly.

Jenkins (3) reporting a study by Book and Norvelle (1) says that students who knew the results of their work progressed faster than those who have no knowledge of their progress. The groups were then interchanged and the students who formerly had no knowledge of their progress were shown their progress and the other group was not. In this case the students shown their progress surpassed the other students in a short time. Ross (10) found that students

having information as to their progress, made higher scores. The information apparently acted as a means of motivation. Bright students were affected more by knowledge of progress than poor students.

GENERAL PROCEDURE

The present study was carried on in the seventh grade geography classes of the Manhattan Junior High School under the direction of Miss Gertrude Fulcher; and the eighth grade United States history classes of the Junction City Junior-Senior High School under the direction of Mr. V. W. Clough.

On March 8, 1939 the seventh grade Manhattan students were given an intelligence test which was constructed at this college, and found to have a high reliability and validity. On March 9, the same students were tested with the "Van Wagenen Unit Scales in Reading".

The experimental class work began on March 13. Classes were divided into control, semi-experimental, and experimental groups. The control group, first hour class, continued the class work in the usual way, and the semi-experimental group, second hour class, was given mimeographed copies of objective questions to use as a guide in studying the assignment, recording their answers to these

questions on perfo-score answer sheets. Perfo-score answer sheets are answer sheets separate from the objective questions, on which symbols are provided corresponding to the alternatives on the objective questions. The student encircles, with a pencil mark, the symbol which corresponds to his chosen answer. The experimental group, third hour class, was given the same objective questions as the semi-experimental group, but were given chemo-score, self-checking answer sheets instead of perfo-score answer sheets, so that each pupil would know immediately whether or not he had made the correct response to the question. The chemo-score answer sheets contain moisture sensitive inks, which turn blue if the student moistens the spot corresponding to the correct alternative of an objective question, or red if he chooses the incorrect response.

The students were allowed to study the lesson in the usual manner for about 20 minutes at the beginning of the hour, and then given the questions to answer, using the text if necessary. The questions were of the objective type, each having from two to four alternatives. They were constructed by the teacher and covered the important points in the lesson. There were about 20 questions for each day, which were based on from 2 to 6 different reference books.

At the end of each week, the three classes were all

given the same objective examination of from 40 to 50 questions, covering the material studied that week. The last weekly test was given April 12. On April 13, a final review test was given covering the unit on China, over which the objective questions were used. April 14, another form of the "Van Wagenen Unit Scales in Reading" was given to all the seventh grade geography classes to determine the effect of the guide questions and check-sheet upon reading ability.

The procedure used in the Junction City eighth grade United States history classes was much the same as that used in Manhattan. The first hour class being the control group, was given no objective guide questions. The second and fourth hour classes were given the objective guide questions and the perfo-score answer sheets, and the fifth and seventh hour classes were given objective guide questions and the chemo-score, self-checking, answer sheets. Objective questions in this case were based wholly on the United States history text book, "The Rise of American Democracy" by Casner and Gabriel. The intelligence test was given March 9, first reading test March 10, and the final reading test April 18.

RESULTS

The reliability of the intelligence test, found by the split half method, was .923 for the Manhattan group and .952 for the Junction City group.

The reliability of the reading test found by correlating the two forms of the same tests, was .710 for the Manhattan group and .689 for the Junction City group. Calculating the reliability for the control and experimental groups of Manhattan separately, the reliability of the control group test was .770 and of the experimental group test .570. This may be due to the fact that the experimental group is undergoing changes in reading habits.

Table 1 gives the number of pupils in each of the different groups, and the mean intelligence test score made by each group. For example, in the left hand column it shows that the Manhattan control group had a total of 33 pupils and their mean score on the intelligence test was 33.78. The Junction City control group had a total of 27 pupils and their mean score on the intelligence test was 33.55. Comparable data for the semi-experimental and experimental groups are shown in the middle and right hand colums respectively.

Table 1. Data Concerning the Paired Groups.

	Control group	Semi-experimental group	Experimental group
Manhattan			
Number	33	25	29
Mean score on intelligence test	33.78	32.20	32.27
S. D. Mean	2.49	2.05	2.44
Junction City			
Number	27	60	65
Mean score on intelligence test	33.55	33.23	33.14
S. D. Mean	2.37	1.55	2.49

In both the Manhattan and Junction City schools, the control groups are slightly superior to the experimental groups in intelligence.

The weekly tests showed gains by the experimental group of from one to four times their standard errors. In five of the eight tests, the experimental group showed statistically valid gains over the control group. In every case, the performance of the semi-experimental group surpassed that of the control group, and the performance of each experimental group surpassed that of any other group. In the Manhattan final review test, the experimental group showed statistically valid gains over the semi-experimental group. The experimental group students also made higher scores on the daily work than did the semi-experimental group.

Tables 2 and 3 give the mean scores made by the pupils on the weekly tests, the difference between means and the standard error of the difference between means. In these tables, Group A is the control group; Group B the semi-experimental group; and Group C the experimental group. This table is to be read as follows: On the Manhattan weekly test for March 17, the control group, Group A, made a mean score of 35.30, the standard deviation of the distribution was 4.70 and the standard error of the mean .831; the semi-experimental group, Group B, made a mean score of 36.91, the standard deviation of the distribution was 4.56 and the standard error of the mean .930; the difference between the mean of the control group and the semi-experimental group was 1.61, the standard error of the difference between means 1.24 and the critical ratio 1.29. The corresponding data for all other groups are listed in the tables in the same manner. The semi-experimental group not only surpassed the control group in performance, but with one exception showed an increasing gain each week. The Manhattan group showed an increasing gain each week. In the Junction City group, the test scores of the experimental group crowded the ceiling so that such gains were practically impossible.

Figures 1 and 2, show the mean score in percentage

Table 2. Mean Scores Made by the Manhattan Pupils on the Weekly Tests.*

Group A	Group B	Group C
	March 17	
M=35.30	M=36.91	M=37.39
S.D. dis.= 4.70	S.D. dis.= 4.56	S.D. dis.= 4.90
S.D. mean= 0.831	S.D. mean= 0.930	S.D. mean= 0.926
M_B-M_A= 1.61	S.D. difference=1.24	Critical Ratio= 1.298
M_C-M_A= 2.09	S.D. difference=1.24	Critical Ratio= 1.685
M_C-M_B= 0.34	S.D. difference=1.23	Critical Ratio= 0.276
	March 24	
M=29.85	M=30.76	M=31.86
S.D. dis.= 4.516	S.D. dis.= 5.38	S.D. dis.= 4.90
S.D. mean= 0.786	S.D. mean= 1.06	S.D. mean= 0.854
M_B-M_A= 0.91	S.D. difference= 1.33	Critical Ratio= 0.684
M_C-M_A= 2.01	S.D. difference= 1.15	Critical Ratio= 1.747
M_C-M_B= 1.10	S.D. difference= 1.38	Critical Ratio= 0.719
	March 31	
M=26.36	M=28.37	M=29.96
S.D. dis.= 4.52	S.D. dis.= 4.92	S.D. dis.= 4.60
S.D. mean= 0.689	S.D. mean= 1.00	S.D. mean= 0.770
M_B-M_A= 2.01	S.D. difference= 1.30	Critical Ratio= 1.546
M_C-M_A= 3.60	S.D. difference= 1.03	Critical Ratio= 3.495
M_C-M_B= 1.59	S.D. difference= 1.33	Critical Ratio= 1.196
	April 7	
M=25.62	M=27.20	M=29.96
S.D. dis.= 4.70	S.D. dis.= 4.70	S.D. dis.= 4.36
S.D. mean= 0.831	S.D. mean= 0.931	S.D. mean= 0.824
M_B-M_A= 1.58	S.D. difference= 1.24	Critical Ratio= 1.25
M_C-M_A= 4.76	S.D. difference= 1.23	Critical Ratio= 3.87
M_C-M_B= 2.76	S.D. difference= 1.34	Critical Ratio= 2.059
	Final Review	
M=64.34	M=69.36	M=75.21
S.D. dis.= 13.8	S.D. dis.= 10.00	S.D. dis.= 9.96
S.D. mean= 2.44	S.D. mean= 2.04	S.D. mean= 1.72
M_B-M_A= 5.02	S.D. difference= 3.18	Critical Ratio= 1.578
M_C-M_A= 10.87	S.D. difference= 2.98	Critical Ratio= 3.647
M_C-M_B= 5.85	S.D. difference= 1.94	Critical Ratio= 3.010

*For explanation of terms, see page 12.

Table 3. Mean Scores Made by Junction City Pupils on the Review Tests.*

Group A	Group B	Group C
	Problem 21	
M= 30.50	M= 34.85	M= 36.15
S.D. dis.= 6.27	S.D. dis.= 8.10	S.D. dis.= 7.11
S.D. mean= 1.23	S.D. mean= 1.06	S.D. mean= 0.89
M_B-M_A= 4.33	S.D. difference= 1.67	Critical ratio= 2.59
M_C-M_A= 6.35	S.D. difference= 1.52	Critical ratio= 4.18
M -M = 1.22	S.D. difference= 1.39	Critical ratio= 0.877
	Problem 22	
M= 34.96	M= 39.55	M= 40.51
S.D. dis.= 5.88	S.D. dis.= 7.36	S.D. dis.= 6.45
S.D. mean= 1.13	S.D. mean= 0.95	S.D. mean= 0.81
M_B-M_A= 4.59	S.D. difference= 1.48	Critical ratio= 3.10
M_C-M_A= 5.55	S.D. difference= 1.39	Critical ratio= 3.99
M_C-M_B= 0.96	S.D. difference= 1.33	Critical ratio= 0.722
	Problem 23	
M= 36.34	M= 41.62	M= 41.98
S.D. dis.= 7.65	S.D. dis.= 8.34	S.D. dis.= 6.57
S.D. mean= 1.50	S.D. mean= 1.09	S.D. mean= 0.83
M_B-M_A= 5.28	S.D. difference= 1.86	Critical ratio= 2.83
M_C-M_A= 5.64	S.D. difference= 1.71	Critical ratio= 3.30
M_C-M_B= 0.36	S.D. difference= 1.38	Critical ratio= 0.268

*M = Mean score.
S.D. dis. = Standard deviation of the distribution.
S.D. mean = Standard deviation of the mean.
S.D. difference = Standard deviation of the difference between two means.
M_A = Mean score of Group A.
M_B = Mean score of Group B.
M_C = Mean score of Group C.
M_B-M_A = Mean score of Group B minus mean score of Group A.

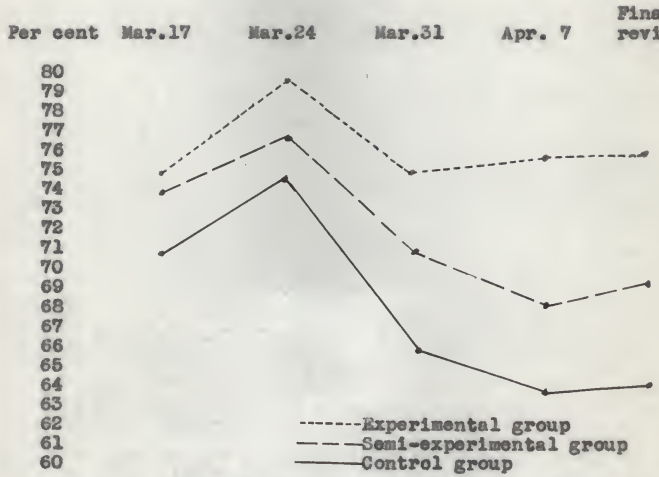

Fig. 1. Graph showing mean percentage grades made by the Manhattan groups on the weekly tests.

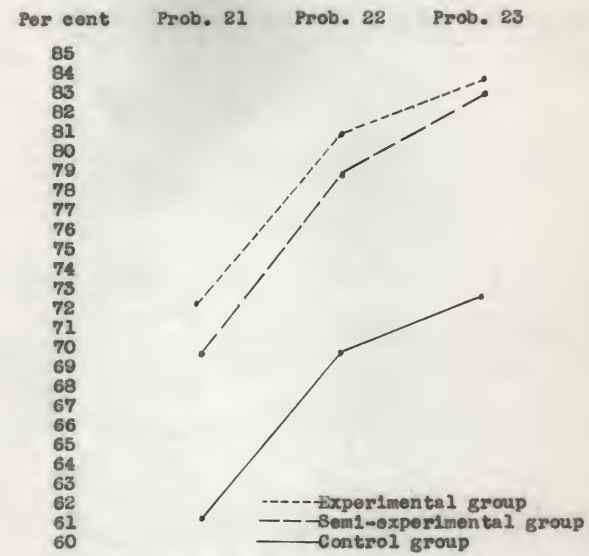

Fig. 2. Graph showing mean percentage grades made by the Junction City groups on the review tests.

made by each group on the weekly tests. The scores were changed to percentages to make the graph, because the tests did not all consist of the same number of questions. The tests were not of equal difficulty, therefore the fluctuation of the lines of the graph have no significance, except to show the relative standing of the control, semi-experimental, and experimental groups. The comparative gains of the Junction City experimental groups are affected by the closeness with which the scores crowd the upper limit of the range.

Table 4 gives the mean scores made by the semi-experimental and experimental groups on the daily objective questions, and the difference between means. Group B is the semi-experimental group, and Group C the experimental group. For the second week work of the Manhattan group, the S. D. of the difference between means is 2.85, and the critical ratio 2.33. Many of the daily papers of the Manhattan group were not labeled in the classroom, so that the grades could not be given in weekly groups, and therefore are not entered on the table.

For problem 23 of the Junction City group, the S. D. of the difference between means is 2.30, and the critical ratio 1.97.

Table 4 shows that in every case, students using the

Table 4. Mean Scores Made by the Students on the Daily Objective Questions.

	Group B	Mean C-Mean B	Group C	Possible score
Manhattan				
First week			44.76	63
Second week	61.28		68.07	96
		6.69		
Third week			71.76	100
Fourth week				
Junction City				
Problem 21	71.63		77.29	100
		5.56		
Problem 22	84.88		87.32	100
		2.44		
Problem 23	74.77		79.29	90
		4.52		

chemo-score answer sheets did slightly better work than did those using the perfo-score answer sheets on the daily work, indicating that there is probably some motivating influence in the self-checking answer sheets. Although the differences are not statistically significant according to usual standards, they are all fairly large and consistent in showing the better performance of the experimental groups.

The Manhattan semi-experimental and experimental groups did not show improvement in reading ability as compared with the control group. The Junction City semi-

experimental and experimental groups did gain slightly over the control group. Tables 5 and 6 give the mean scores made by the different groups on the first and final reading tests.

Table 5. Mean Scores Made by the Manhattan Pupils on Reading Tests.

Group A	Group B	Group C
	First reading test	
$M = 18.85$	$M = 17.11$	$M = 19.00$
$M_B - M_A = -1.74$		
$M_C - M_A = 0.15$		
	Final reading test	
$M = 20.71$	$M = 18.96$	$M = 20.65$
$M_B - M_A = -1.75$		
$M_C - M_A = -0.06$		

Table 6. Mean Scores Made by the Junction City Pupils on Reading Tests.

Group A	Group B	Group C
	First reading test	
$M = 20.33$	$M = 20.70$	$M = 19.40$
$M_B - M_A = 0.37$		
$M_C - M_A = -0.93$		
	Final reading test	
$M = 19.64$	$M = 20.58$	$M = 19.50$
$M_B - M_A = 0.94$		
$M_C - M_A = -0.14$		

Table 6 shows that the semi-experimental gained .37 of a point, and the experimental .79 of a point as compared with the control. The correlation between the first and

final reading test for the control group was .75. The correlation between the first and final reading test of the experimental group was .804 and the standard error of the difference in gains .81, which gives a critical ratio of (.79 ÷ .81) = .97.

The lower scores on the final reading test are due to the fact that the second form of the test given was more difficult than the first. The experimental group is the only one to make a gain in mean score from the first reading to the final reading test in the Junction City group.

CONCLUSION

Throughout the experiment, the experimental groups have excelled all the other groups in performance. The chemo-score answer sheets evidently provide some motivation for the pupil and also give him a means of knowing the right answer. The differences in gain in favor of the experimental groups were not as great as those found in previous experiments, but this may be due to the fact that the questions were made up by teachers who had very little experience in making up objective questions, and little time to spend on the construction of the questions, and also to the crowding of the test ceiling by the experimental groups. In the Manhattan group where the ceiling was not

crowded, the semi-experimental and experimental group performance did show a regular gain.

The Manhattan experimental group must be undergoing some change of reading habits, because the variability in the final reading test is much greater for this group than the variability of their first reading test scores. Such a change was shown in the Junction City experimental group by a slight gain in the second test.

ACKNOWLEDGMENTS

The writer wishes to express his appreciation to Doctor J. C. Peterson for help and advice, and for the use of his tests and answer sheets in the experiment; to Superintendent W. A. Wood of Junction City and Superintendent W. E. Sheffer of Manhattan for their interest and cooperation in starting the experiment; and to Miss Gertrude Pulcher and Mr. Victor W. Clough from Manhattan and Junction City respectively, for assistance in securing data from their classes.

LITERATURE CITED

(1) Book, W. F. and Norvelle, L.
An experimental study of learning incentives. Ped. Sem. 29:305-362. 1922.

(2) Fleenor, B. H.
The self-instructor method in correspondence courses. Kans. Acad. Sci., Trans. 36:166-171. 1933.

(3) Jenkins, John G.
Psychology in business and industry. New York. John Wiley and Sons, 235 p. 1935.

(4) Kellogg, W. N. and Payne, Bryan.
The true-false question as an aid in studying. J. Educ. Psychol. 29:581-589. 1938.

(5) Keys, N.
The influence on learning and retention of weekly as opposed to monthly tests. J. Educ. Psychol. 25:426-436. 1934.

(6) Marx, Lawrence M.
Immediate check-up devices and their value in learning. Unpublished thesis. Kansas State College of Agr. and Appl. Science. p. 1-31. 1933.

(7) Noll, Victor H.
The effect of written tests upon achievement in college classes. J. Educ. Res. 32:345-358. 1939.

(8) Parr, F. W.
Teaching ways of study. J. Higher Ed. 5:377-381. 1934.

(9) Peterson, J. C.
The value of guidance in reading for information. Kans. Acad. Sci., Trans. 34:291-296. 1931.

(10) Ross, C. C.
An experiment in motivation. J. Educ. Res. 18:337-346. 1929.

(11) Salisbury, R.
Some effects of training in outlining. Engl. J. 24:111-116. 1935.

(12) Symonds, Percival M. and Chase, Dorris Harter.
Practice vs. motivation. J. Educ. Psychol. 20:351. 1929.

(13) Washburne, John N.
The use of questions in social science material. J. Educ. Psychol. 20:321. 1929.