DOCUMENT RESUME

ED 072 654 EM 010 770

AUTHOR Kromhout, Ora M.
TITLE Effect of Computer Tutorial Review Lessons on Exam Performance In Introductory College Physics. Tech Memo Number 64.
INSTITUTION Florida State Univ., Tallahassee. Computer-Assisted Instruction Center.
SPONS AGENCY Office of Naval Research, Washington, D.C. Personnel and Training Research Programs Office.
PUB DATE 1 Oct 72
NOTE 33p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Achievement; *College Science; Comparative Analysis; *Computer Assisted Instruction; *Intermode Differences; *Physics Instruction; *Review (Reexamination); Student Developed Materials; Student Volunteers; Tutorial Programs

ABSTRACT

Computer tutorial review lessons for a one-quarter lecture course in introductory college physics have been developed and used by approximately 500 students, on a voluntary basis, over a period of several years. The investigation reported here shows that the exam grades of students who used the computer materials are significantly higher (.01 level) than the class as a whole, if results from sections are pooled. Evidence is presented to show that the results are not primarily due to a factor of selection because of the voluntary basis for participation. (Author/JK)

CAI CENTER

TECH MEMO

EFFECT OF COMPUTER TUTORIAL REVIEW LESSONS ON EXAM PERFORMANCE IN INTRODUCTORY COLLEGE PHYSICS

Ora M. Kromhout

Tech Memo No. 64
October 1, 1972
Tallahassee, Florida

Project NR 154-280
Sponsored by
Personnel & Training Research Programs
Psychological Sciences Division
Office of Naval Research
Arlington, Virginia
Contract No. N00014-68-A-0494

FLORIDA STATE UNIVERSITY

NPO
EM

Tech Memo Series

The FSU-CAI Center Tech Memo Series is intended to provide communication to other colleagues and interested professionals who are actively utilizing computers in their research. The rationale for the Tech Memo Series is threefold. First, pilot studies that show great promise and will eventuate in research reports can be given a quick distribution. Secondly, speeches given at professional meetings can be distributed for broad review and reaction. Third, the Tech Memo Series provides for distribution of pre-publication copies of research and implementation studies that after proper technical review will ultimately be found in professional journals.

In terms of substance, these reports will be concise, descriptive, and exploratory in nature. While cast within a CAI research model, a number of the reports will deal with technical implementation topics related to computers and their language or operating systems. Thus, we here at FSU trust this Tech Memo Series will serve a useful service and communication for other workers in the area of computers and education. Any comments to the authors can be forwarded via the Florida State University CAI Center.

Duncan N. Hansen
Director
CAI Center

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Florida State University Computer-Assisted Instruction Tallahassee, Florida 32306	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

EFFECT OF COMPUTER TUTORIAL REVIEW LESSONS ON EXAM PERFORMANCE IN INTRODUCTORY COLLEGE PHYSICS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Tech Memo No. 64, October 1, 1972

5. AUTHOR(S) (First name, middle initial, last name)
Ora M. Kromhout

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1, 1972	27	11

8a. CONTRACT OR GRANT NO. N00014-68-A-0494	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. NR 154-280	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
c.	
d.	

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Personnel & Training Research Programs Office of Naval Research Arlington, Virginia

13. ABSTRACT

Computer tutorial review lessons for a one-quarter lecture course in introductory college physics have been developed and used by approximately 500 students, on a voluntary basis, over a period of several years. The investigation shows that the exam grades of students who used the computer materials is significantly higher (.01 level) than the class as a whole, if results from sections are pooled. Evidence is presented to show that this is not primarily due to a factor of selection because of the voluntary basis for participation.

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Security Classification

A-31409

EFFECT OF COMPUTER TUTORIAL REVIEW LESSONS ON EXAM PERFORMANCE IN INTRODUCTORY COLLEGE PHYSICS

Ora M. Kromhout

Tech Memo No. 64
October 1, 1972
Tallahassee, Florida

Project NR 154-280
Sponsored by
Personnel & Training Research Programs
Psychological Sciences Division
Office of Naval Research
Arlington, Virginia
Contract No. N00014-68-A-0494

EFFECT OF COMPUTER TUTORIAL REVIEW LESSONS ON EXAM PERFORMANCE IN INTRODUCTORY COLLEGE PHYSICS

Ora M. Kromhout
Florida State University

ABSTRACT

Computer tutorial review lessons for a one-quarter lecture course in introductory college physics have been developed and used by approximately 500 students, on a voluntary basis, over a period of several years. The investigation shows that the exam grades of students who used the computer materials is significantly higher (.01 level) than the class as a whole, if results from sections are pooled. Evidence is presented to show that this is not primarily due to a factor of selection because of the voluntary basis for participation.

EFFECT OF COMPUTER TUTORIAL REVIEW LESSONS ON EXAM PERFORMANCE IN INTRODUCTORY COLLEGE PHYSICS

Ora M. Kromhout
Florida State University

Exploration of the use of computers in instruction has been one aspect of the current interest in extending the technology of education. Perhaps because of the glamour of the computer, it has been a favorite object of enthusiastic support and equally fervent opposition (Hammond, 1972). The initial discussions of the sixties were hampered by a shortage of experience, with the result that the computer tended to be judged wholly "good" or "bad" for instruction depending on the individual's perception of it. Hopefully, as the amount of experience grows and is reported, the computer may be viewed objectively as one more medium for instruction, which may be a good choice for some uses, a poor one for others, depending on the situation. Such a judgment should be based not solely on speculation, but on measured results with a significant number of actual subjects. It is the purpose of this paper to report and evaluate the results of one such investigation.

Because of its reliance on computation, physics was an area of particular appeal for early investigation of the instructional use of computers. Also, physics faculties had relatively easy access to computers, and were accustomed to using them in research. By 1968, a large number of projects existed throughout the nation, using the computer in varied and imaginative ways to teach physics. Computers were used in conversational as well as computational programs, as simulators, and as generators of motion pictures (Schwarz, Kromhout, & Edwards, 1969).

Despite this extensive exploration, relatively little evaluation has been reported, in the sense of measured effects of introducing the computer into instruction in undergraduate or graduate physics. Published reports have for the most part emphasized <u>how</u> it has been used, and the instructor's rationale for using the computer to improve the quality of instruction. As for evaluation, reports are generally descriptive of good results as seen by the instructor; sometimes student comments are included. However, the reports rarely include numbers reflecting a change in a variable which purports to objectively measure learning. There are several possible explanations for this. The authors are usually physical scientists whose primary interest and approach is not in educational research. In many cases the number of subjects is small, partly due to the fact that much of the material is designed for higher level courses than the large, introductory courses. In addition, most faculty members would be reluctant to tell some students that they may not use the computer materials, and others that they must use them, as is usually done in experimental designs for educational research. Lastly, and perhaps most important, the instructor of a college science course may view the skills being learned as so complex that it would be difficult to define a valid measuring tool for the dependent variable in the study.

An opportunity to observe the results of computer-assisted instruction over a period of years, with a substantial number of students, arose as an outgrowth of a project at Florida State University's Computer-Assisted Instruction (CAI) Center. The original project was to develop and implement a complete course in college physics, presented by the computer and other media, but without a human instructor. The course selected was Physics 107, a one-quarter, large enrollment introductory course for nonscience majors, mostly freshmen. The CAI multimedia version of the course was offered twice in the academic year 1967-68. The project was successfully completed, with the "CAI" students performing as well as, and often better than, the "conventional" students on the same exams (Hansen, Dick, & Lippert, 1968a; Kromhout, Hansen, & Schwarz, 1970).

As part of this project, CAI review lessons to supplement the regular physics course were begun during 1966-67, and have been offered fairly regularly since that time. For approximately one week before each course exam, students have been given the opportunity to come to the CAI Center on a voluntary basis to go through some of the 55 topics coded into the computer. Each topic consists of a series of multiple choice questions, typically six to eight, in a format similar to that used in the course exams. Response choices are included which pinpoint common misunderstandings. A tutorial approach is used, with response contingent feedback to most responses, both correct and incorrect. This approach has been shown to result in more effective learning of science concepts than simply identifying the response as correct or wrong (Gilman, 1969). The student must answer each question correctly before he can move on to the next question. The review materials are described in the appendix.

The first version of the physics review lessons was offered in late 1966 on the IBM 1440, with good response from the Physics 107 students. Within a few months, use of the system was close to saturation during the assigned hours. A 1967 study of the grades of students actively participating in the CAI review showed that they averaged about 10% higher than those students who did not (Hansen, Dick and Lippert, 1968a). However, it was recognized that an unknown factor of selectivity might be present; students who made the voluntary effort to come to the CAI Center for the review might be more motivated on the average than those who did not. On the other hand, the volunteers might be those who are most in need of help.

Since that time, the review has been made available fairly regularly before exams. It has been coded into several languages, including Coursewriter I, Coursewriter II, and FOCAL, and has been implemented on the IBM 1500, the PDP-8, and most recently on the CDC 6500 Kronos time-sharing system.

The present study sought to take advantage of the longevity and regular use of this program to investigate the effect of the review on student exam performance.

Literature Review

Although numerous articles have been written about computer-assisted instruction in college physics, it is difficult to find actual measurements of changes in student performance associated with the instruction. However, several recent articles describe such experimentation in college chemistry. The work of Castleberry, Culp, Lagowski and Rodewald at the University of Texas is well documented.

Lagowski, Castleberry, and Culp (1972) concluded that the use of supplemental computer-based instructional techniques in a general chemistry course exerted a positive influence upon student performance. They compared a computer-supplemented course with two traditional courses, all three sections initially having similar class distributions on the Standard Achievement Math Test and Standard Achievement Verbal Test. Tests for significance of differences were not reported. The superior performance of the students in the CAI-supplemented course was attributed in part to the modularization of the course itself, and development of behavioral objectives, as the result of the computer-based techniques.

Rodewald, Culp and Lagowski (1970) used control and experimental groups of 40 randomly assigned students each to study the effect of programmed modules as a supplement to a conventional organic chemistry course. All students took the same quizzes and exams which contained some questions which were related to the CAI modules, and some unrelated questions. Scores on the latter served as an internal reference for both groups. The experimental group performed significantly better ($\alpha = .05$) on the CAI-related questions, but showed no significant difference on the CAI-unrelated questions.

Positive results were reported also in a 1968-69 study at the University of Texas in which computer modules were made available to general chemistry students on a voluntary basis (Castleberry & Lagowski, 1970). About 100 students in each of the two semesters in which the study was run used the modules. Using regression models, the investigators found that in both semesters, students who took advantage of the computer modules scored significantly higher than the control group on that part of the final examination covered by the modules. Covariables were: SAT math scores, SAT verbal scores, control or experimental group membership, sex, and chemistry placement exam scores. The basis for classification as control or experimental was different in the two semesters. In the first semester, those students who logged at least 90 minutes of computer time were designated as the experimental group and the remainder as the control group. In the second semester, 149 students from one course section volunteered to use the CAI modules. Of these, 100 were randomly selected for the experimental group, about 74 of these participated. The control group consisted of those who volunteered, but who did not participate. The investigators felt that the second semester method of selection reduced the likelihood that the higher achievement of the experimental group was due to the higher motivation of volunteers.

At the University of Illinois, CAI was used in teaching students how to solve selected multistep general chemistry problems (Grandey, 1970). Volunteers were recruited for an experimental group to use the PLATO computer instruction. Compared with a control group, the PLATO group's score on a PLATO quiz was 14.7% higher ($p<.02$) on material covered in the computer work, but did not differ significantly on material not covered. On the final exam the PLATO group performed 23.8% higher on material covered in the computer instruction, but only 1.3% higher on material not covered.

At the State University of New York at Binghamton, CAI in the form of simulated experiments was used in 1968-69 on a voluntary basis by students in general physics (Stannard, 1970). The experimental group did much better on the material covered by the computer than did the control group (65% vs. 29%), and about equally as well (63% vs. 64%) on material not directly related to the computer work.

With the exception of the Lagowski, Castleberry, and Culp (1972) study, in which the CAI course was structured differently from the traditional course, the studies have much in common. The computer materials were used as a supplement to a traditional course in introductory physics or chemistry, usually one with a large enrollment. The experimental group performed significantly better than the control group on CAI-related questions, but not on questions unrelated to the CAI materials. The performance on the CAI-unrelated questions was used as an internal reference or control to show that the improved performance was not due to selection of superior students for the experimental group. In addition, the Castleberry and Lagowski study discredited the aptitude-selection explanation of performance by using a regression analysis which included external aptitude measures such as SAT math and verbal scores, and placement exam scores.

The present study seeks to extend this work to another area of introductory college instruction in which the computer work is again supplementary to a traditional course. As in the studies cited, performance on standard exams by students who use the review is compared with that of the students who do not. However, the methodology differs in two respects. First, the comparison of treated and untreated groups is based on a larger number of students, with data collected over a period of five quarters and ten exams, rather than from a controlled experiment with a smaller number of subjects. Secondly, a different approach is used in determining whether the voluntary participation biasses the results by selecting "better" students. In the FSU program, there are no CAI-unrelated questions to serve as an internal reference or control; the computer lessons attempt to cover all topics in the course. Therefore, the performance of the participants on CAI-unrelated questions can not be used to detect a selection effect. A different internal reference is used in the analysis: the change in relative position on the class grade distribution of a sub-group which did not use the CAI review for an early exam, but did use it in a subsequent exam. While an external measure such as SAT scores might have been used, the correlation of such measures with performance on Physics 107 exams is unknown.

A feature of the attendance pattern suggested a measure more directly related to untreated exam performance. The computer review was offered for approximately one week before each of the two or three exams in the course. During any one quarter, the percent of the class which participated in the review increased from the first exam to the next, as shown in Figure 1. It was therefore possible to identify, within a class section, a group of students who did not use the review for the first exam, but did for a later exam. This group will be referred to as sub-group A. By locating the mean performance of sub-group A on the class exam distribution for the two exams (untreated and treated), and recording the mean standardized score (group mean minus class mean, divided by class standard deviation), on the two exams it should be possible to find out two things. The standardized score for the untreated exam should show whether these volunteers are coming from a particular part of the class distribution. The difference between the standardized scores on the untreated and treated exams, for the same group of subjects, should show whether use of the review makes a difference in student exam performance, relative to the class as a whole.

The research hypotheses were formulated as follows:

1. The mean exam score of the treated group is significantly higher than that of the untreated group.

2. For a "sub-group A" which was untreated on an early exam, but treated on a later one, the mean standardized score for the

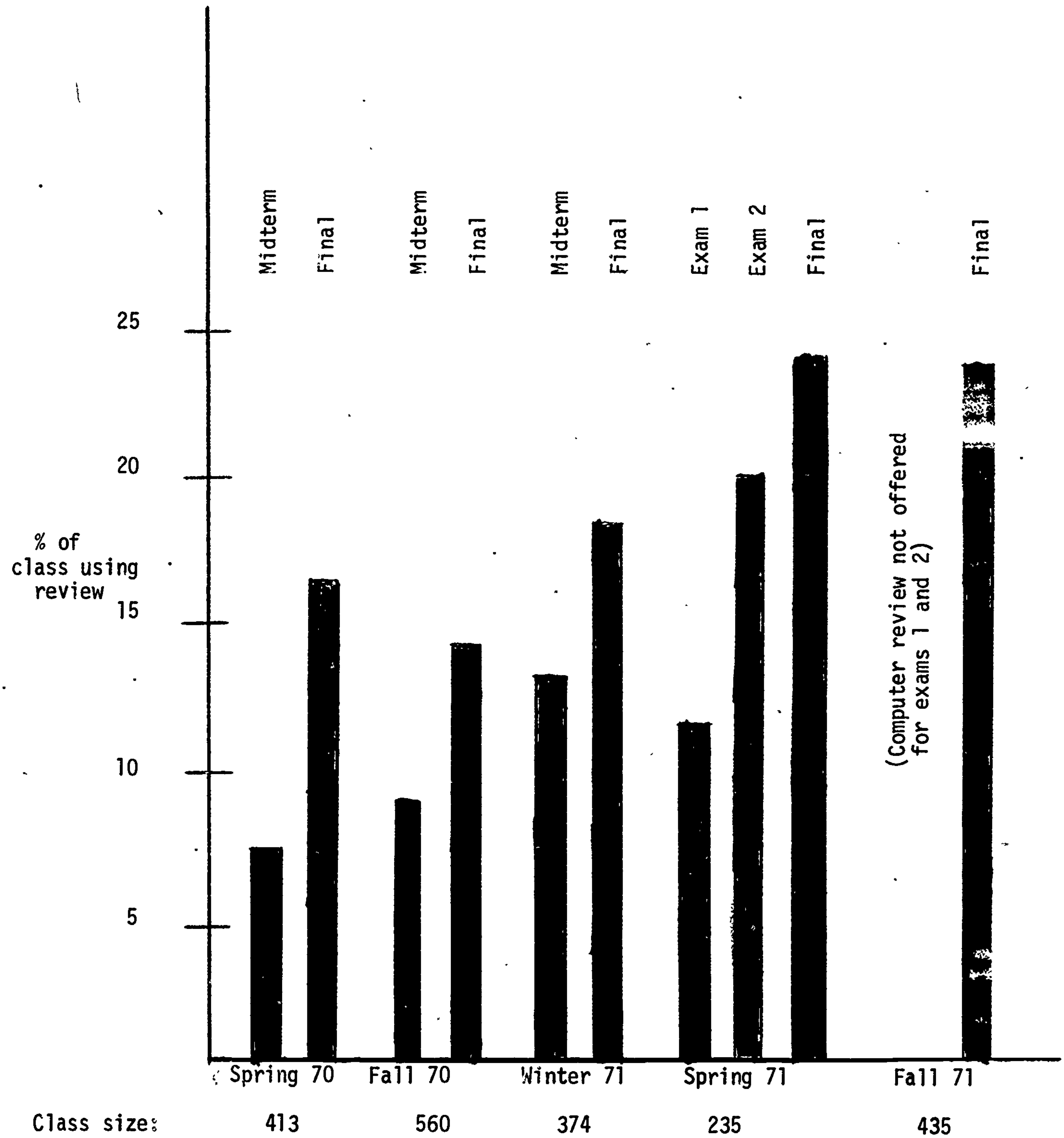

Figure 1 Usage of Review for Each Exam

sub-group on the untreated exam is not significantly different from that of the class as a whole.

3. There is a positive change in the standardized score of this sub-group from the untreated exam to the treated exam.

Method

Subjects

The subjects were volunteer participants from Physics 107 classes at Florida State University. They were not required to participate nor were they made aware that a study of results was being made. They came from all three sections of Physics 107. All sections have the same course outline and exams, but different instructors.

Materials

The materials are multiple choice questions grouped by topics, 55 topics available in all. Feedback to correct and incorrect student responses is immediate, usually including tutorial information. Students can select any topic from the list, a typical topic having six or eight questions, and taking 15 to 25 minutes to complete.

Dependent Variable Measure

The dependent variable in the study is the student's grade on the regular Physics 107 exams. These are multiple choice exams, of 20 or 40 items. The reliability range was .51 to .65 for the 20-item exam, .56 to .82 for the 40-item exams, with a mean value of .56 and .73 for the 20-and 40-item exams, respectively.

Procedure

The availability of the computer review was announced to each section of Physics 107 approximately one week before each exam. The announcement stated that participation was voluntary, no prior knowledge of computers was necessary, and that students in the past had found it helpful. The review was available for a limited period each day, usually 5 or 6 hours.

When students arrived at the CAI Center, a proctor recorded their names, showed them how to use the teletype terminal, and indicated he would be on hand to answer questions about the operation or the physics content. On the wall was a list of the topics recommended by the instructor for that exam; students could choose topics in any order, and could stay as long as they wished, unless others were waiting to use the terminal.

All students, treated and untreated, took the same multiple choice exams under the same conditions. These exams, which constituted the total basis for the student's course grade were given twice during the course prior to the Spring, 1971 quarter. Thereafter three exams were given during the course.

After each exam, individual student scores, class means, standard deviations, and, in some cases, reliability were obtained from the instructors. These data were collected for all exams in the spring and fall quarters, 1970, and winter and spring quarters, 1971. Because of other demands on the computer, the review was not offered for the first two exams of the fall 1971 quarter. Beginning with the spring 1972 quarter, the review was moved to the FSU Computing Center's CDC 6500.

Results

Comparison of Scores of Treated and Untreated Groups.

Table 1 shows the mean exam score of the CAI (treated) group and the untreated group for the exams in five quarters. The total number of student-users was 580; however, the same student may be counted more than once if he came for more than one exam. The total class population from which these students came totalled about 2000. The average level of participation was about 15%; this varied considerably from one exam to another, as shown in Figure 1, rising during a given quarter as more students heard about the service, then dropping back at the start of each new quarter. Three sections of the course were given each quarter, with three different instructors. However, the same instructors stayed with the same section, in most cases, throughout the period. All sections used the same text, course outline, and exams.

A comparison of column X, the mean score for the CAI group, with column U, the mean for the untreated group, shows that the CAI group scored higher in 27 of the 29 sections. The ratio of CAI group mean to untreated group mean is shown in another column, X/U. Although this ratio is greater than 1.00 in all but two of the sections, a t-test on the difference (X-U) shows a significant difference in only 12 of the 29 sections ($p<.05$); this could be due to the small number of CAI students (n) in any one section. The consistently higher CAI scores strongly suggest a significant difference. If the results of all sections are combined for a given exam, in order to increase the size of n and thereby the resolution needed to distinguish the differences, the results are as in Table 2. For 8 of the 10 exams for which data was collected, the

Table 1. Mean exam scores for treated and untreated groups, by section.

Exam	Section	Mean Score: Treated X	Mean Score: Untreated U	Ratio of Means X/U	Group Size: Treated n	Group Size: Untreated N	Standard Deviation: Treated S_t	Standard Deviation: Total Class S_u	t	df (n+N-2)	$F=(S_u/S_t)^2$ df=(n+N-1, n-1)
Spring '70 Midterm	1	14.2	13.0	1.09	13	117	2.07	2.69	1.53	128	1.69
	2	14.6	13.0	1.12	10	84	1.96	3.15	1.52	92	2.58
	3	15.0	13.2	1.14	7	182	2.73	2.62	1.78	187	0.92
Spring '70 Final	1	11.7	10.7	1.09	21	111	2.22	2.57	1.65	130	1.34
	2	12.2	11.0	1.11	22	72	2.70	2.80	1.73	92	1.08
	3	10.8	10.6	1.02	24	170	2.29	2.54	0.34	192	1.23
Fall '70 Midterm	1	12.6	13.1	.96	32	159	2.55	3.21	-0.85	[illegible]	1.58
	2	13.0	12.5	1.04	9	192	2.94	3.15	0.48	[illegible]	1.15
	3	13.0	12.4	1.05	9	161	2.94	3.30	0.56	[illegible]	1.26
Fall '70 Final	1	12.3	10.7	1.16	36	152	2.38	2.67	3.35 ✓✓	186	1.26
	2,3	11.2	10.8	1.04	42	328	2.82	2.52	1.06	368	0.80
Winter '71 Midterm	1	11.0	9.96	1.10	17	99	2.87	2.85	1.39	114	0.99
	2	10.7	10.7	1.00	23	121	2.69	2.88	0.00	142	1.15
	3	13.2	12.6	1.04	10	104	2.32	2.77	0.60	112	1.43
Winter '71 Final	1	11.3	9.95	1.13	20	96	1.82	2.39	2.26 ✓	114	1.72
	2	11.4	10.5	1.08	30	114	2.87	2.70	1.48	142	0.89
	3	11.6	10.4	1.12	20	94	1.49	2.23	2.20 ✓	112	2.24 ✓
Spring '71 Exam 1	1	15.9	13.0	1.23	13	51	2.23	3.02	3.14 ✓✓	62	1.83
	2	15.1	13.7	1.10	8	98	3.30	3.10	1.24	104	0.88
	3	15.7	12.7	1.23	7	62	2.18	3.05	2.43 ✓	67	1.96
Spring '71 Exam 2	1	15.6	14.0	1.11	24	38	2.31	2.65	2.26 ✓	60	1.32
	2	15.4	12.8	1.20	11	93	2.57	3.4	2.33 ✓	102	1.75
	3	13.0	12.9	1.01	11	55	3.36	3.1	0.12	64	0.85
Spring '71 Final	1	13.8	11.8	1.17	20	41	2.36	3.17	2.30 ✓	59	1.80
	2	14.4	10.8	1.33	25	85	3.42	3.30	4.75 ✓✓✓	108	0.93
	3	14.6	11.0	1.33	12	56	2.10	3.22	3.52 ✓✓	66	2.35
Fall '71 Final	1	12.6	11.0	1.15	61	119	2.59	2.96	3.50 ✓✓	178	1.30
	2	14.4	11.9	1.21	29	92	2.40	2.78	4.22 ✓✓✓	119	1.34
	3	12.0	11.5	1.04	14	54	3.04	3.10	0.48	66	1.04
Total					580	3200					
Mean		12.78	11.69	1.093			2.536	2.849	8.45 ✓✓✓	3778	1.265 ✓

✓✓✓ $p < .001$
✓✓ $p < .01$
✓ $p < .05$

Table 2. Mean exam scores for treated and untreated groups, by exam, all three sections combined.

Quarter	Exam	Mean Score: Treated X	Mean Score: Un-Treated U	Ratio of Means X/U	Group Size: Treated n	Group Size: Un-Treated N	Standard Deviation: Treated S_t	Standard Deviation: Total Class S_u	t	df (n+N-2)	$F=(S_u/S_t)^2$ df=(n+N-1, n-1)
Spring '70	Mid-term	14.5	13.1	1.11	30	383	2.19	2.76	2.72✓✓	411	1.60
	Final	11.6	10.7	1.08	67	353	2.40	2.61	2.36✓	418	1.18
Fall '70	Mid-term	12.7	12.6	1.01	50	512	2.69	3.22	0.19	560	1.43
	Final	11.7	10.7	1.09	78	480	2.62	2.57	3.16✓✓	556	0.96
Winter '71	Mid-term	11.3	11.1	1.02	50	324	2.68	2.84	0.48	372	1.12
	Final	11.4	10.3	1.11	70	304	2.18	2.46	3.37✓✓	372	1.28
Spring '71	No.1	15.6	13.2	1.18	28	211	2.52	3.06	3.88✓✓	237	1.47
	No.2	14.9	13.1	1.14	46	186	2.62	3.11	3.55✓✓	230	1.41
	Final	14.2	11.1	1.28	57	182	2.77	3.24	6.36✓✓	237	1.37
Fall '71	Final	13.0	11.4	1.14	104	265	2.62	2.93	4.78✓✓	367	1.25

✓✓ $p < .01$

✓ $p < .05$

CAI group (combined across 3 sections) scored significantly higher ($p<.05$) than the class mean, based on a two-tailed t-test.

Finally, if the results of all sections and exams are combined, the mean CAI score was 12.78, or 9.3% higher than the untreated group mean of 11.76. A t-test on the means indicates that the treated group's score was significantly higher than that of the untreated group ($t=8.45$, $df=3778$, $p<.001$).

<u>Change in Mean Standardized Scores for Volunteer Subjects from Untreated Exam to Treated Exam.</u>

While the data in the previous section show higher performance by the CAI group, the question remains, is this due to selectivity, since participation is on a voluntary basis? Would these volunteers have performed better even without the CAI review? To answer this question, the attendance list was checked to look for a sub-group of the participants, in any section, who had not come to the review for an early exam, but did come for a later exam. The mean standardized score of this sub-group on the untreated exam should give an indication of whether the volunteers tend to come from the higher half of the distribution. The mean standardized score of the same subjects on the later exam, with treatment, can then be compared to the untreated standardized score. Such a sub-group was identified in data from several of the sections, and will be referred to hereafter as sub-group A.

Looking at column 2 in Table 3, which shows the "untreated" standardized scores for a sub-group A in several sections, it can be seen that the scores are not always positive, nor always negative. They range from -.4 standard deviation below class mean, to +.6 standard deviation above it. The mean standardized score for these subgroups is +.12 standard deviation ($t=1.97$, $df=2000$, $p=.05$).

Table 3. Standardized scores* of volunteers on untreated and treated exams.

		Sub-group A (see text)					Sub-group B (see text)			
Quarter	Section	Class Size N	Early exam, Untreated	Later exam, Treated	Change	n	Early exam, Treated	Later exam, Untreated	Change	n
	Column:	1	2	3	4	5	6	7	8	9
Spring '70	1	130	+.23	+.27	+.04	11	+.15	.00	-.15	6
Fall '70	1	191	+.19	+.83	+.64	14	-.50	-.24	+.26	9
Fall '70	2	371	-.17	+.18	+.35	13	-.08	+.26	+.34	9
Winter '71	1	116	-.46	+.19	+.65	9	+.25	+.70	+.45	6
	2	144	+.17	+.38	+.21	14	+.28	+.33	+.05	6
Spring '71	1a**	64	-.31	+.24	+.55	11				
	1b	64	-.38	+.28	+.66	12				
	2b	110	+.58	+.70	+.12	18				
	2c	110	+.39	+.64	+.25	14				
Fall '71	1b	182	+.16	+.36	+.20	60				
	1c	182	+.15	+.36	+.20	60				
	2b	121	+.18	+.67	+.49	29				
	2c	121	+.48	+.67	+.19	29				
	3b	68	-.41	+.11	+.52	14				
	3c	68	-.13	+.11	+.24	14				
Total		2042				322				36
Means			+.12✓	+.42//	+.30//		-.03	+.18	+.21	
t, df			2.00,2041	7.00,2041	5.4,321		-.18,2041	1.07,2041	1.3,35	

*Standardized score=(sub-group mean - class mean)/class standard deviation

**For those quarters in which 3 exams were given, the letters a,b, and c indicate which pair of exams is compared: <u>a</u> compares exam 1 and exam 2, <u>b</u> compares exam 1 and the final, <u>c</u> compares exam 2 and the final. Where no letters are given, the comparison is between the midterm and the final.

// $p < .01$; ✓ $p = .05$

The mean standardized scores of these same subjects on the later exam, when they did use the computer review, is given in Column 3. The change in position of the sub-group on the class distribution is listed in Column 4. In each case, the change is positive, indicating that the relative position of the sub-group on the class distribution has improved on the later exam, when they used the review. The mean standardized score for these students on the treated exam is +.42, an average change of about 0.3 standard deviation from the untreated exam. The mean standardized score for the sub-group on the treated exam, $\pm$.42 standard deviation, is significantly different from the total class mean ($p<.001$).

An apparent effect which was surprising was noticed with respect to those participants who came to the computer review for the early exam, and did not come for the later exam. These will be called sub-group B; their data are also shown in Table 3, columns 6 through 9. It appeared that they also improved their exam grade, relative to the class as a whole, on the second exam; the mean change was +.2 standard deviation (from -.03 to +.18). However, the number of subjects in this sub-group was small, and the change is not significant at the .01 or .05 level, nor is the untreated standardized score significantly different from the class as a whole.

Discussion

The results of this study on an introductory general physics course support previous findings on introductory chemistry courses; students who use computer modules as a voluntary supplement to a conventional course score higher on course exams ($p<.001$). The size of the effect appears to be about 9%. The hypothesis that the "untreated" exam performance is not significantly different from that of the class as a whole was rejected, but marginally ($p=.05$). The third hypothesis, that there is a positive change in the standardized score of the volunteer participants, from the untreated exam to the treated exam, was upheld ($p<.001$). The size of this improvement was approximately four-tenths of a standard deviation.

These results are in agreement with those of previous studies in chemistry, in showing superior performance on exams by those students who participate in the computer review. The question of whether the students who voluntarily supplement their studies with computer work tend to be the initially superior students is less clear. This study suggests that they are, on the average, slightly higher-performance students, by approximately one-tenth of a standard deviation. Previous studies have concluded that they are not different from the class as a whole. The apparent disagreement may be due to a difference in power to detect such a small difference, with the present study having a larger number of subjects and consequently a greater power of resolution. On the other hand, there may actually be a difference between the volunteers in those studies and in this one. Even in the present study, the size of the selection effect is several times smaller than the difference between the CAI group and the class as a whole (one-tenth as opposed to four-tenths of a standard deviation). Students who use the review improve their position relative to the class distribution of exam scores.

The categorization of "treated" and "untreated" subjects is an important decision in studies of volunteer subjects, in which the length of time spent with the computer materials is left up to the student. In some previous studies the division was based on the number of hours used. In the present study, any student who came to the Center to use the review at all was considered treated, all others untreated. The first assumption is probably reasonable, since few students stayed less than one hour or more than three, the average being about two hours. The second assumption, that students who did not come to the Center are "untreated", is known to be fallacious to some extent. Participants take their computer typeouts with them when they leave the Center, to study at a later time, and often share them with other "untreated" students. The extent of this effect is not known, but would presumably operate in the direction of making the apparent difference between treated and untreated groups smaller than it actually is.

The method used in this study to detect small differences depends on combining results across three sections (instructors) and ten different exams, with five different groups of students. This procedure is thought to be justified in view of the stable course content and exam content in all sections. Also, the hypothesis regarding improvement of performance relative to the class was tested on related samples, with the same subjects being compared with and without treatment. The decision not to reject that hypothesis is therefore not jeopardized by the process of combining results or by the small selection effect due to volunteer participation.

Conclusions and Implications

The finding that the use of the computer review improves performance on examinations indicates that the availability of this service is useful to students in a large introductory physics course for nonscience majors. Although students have been using the service enthusiastically prior to these findings, instructors and administrators may be interested in documentation of its effectiveness. Comments from students who use the service indicate that hope for an improved exam grade, and curiosity about the computer, are the two most common reasons for voluntarily spending this time with the computer materials.

In studies of this type, measuring the effect of a voluntary-participation treatment, it is difficult to know just how to separate out the selection effect, if there is one. This study has attempted one approach -- that of measuring a "gap" or change in performance of the same subjects on two different exams, in terms of the individual's position relative to the class. This has been proposed as a more valid measure of untreated performance in such a study than external aptitude measures such as SAT. The importance of large samples, such as those obtained here by observation over a number of quarters, is shown in the increasing power to detect differences as more sections are combined.

The findings do not, in themselves, explain why the review improves grades. It might be interpreted as a form of drill-and-practice, with the effect due simply to the increased time of study. If so, it would be interesting to see if a similar amount of time spent with other materials (such as the textbook, programmed texts, computer typeouts) is as effective. At the computer terminal, the student is faced with making a series of decisions, on multiple choice questions, and committing himself to an answer. The level of concentrated attention on the part of students at the terminals is impressive, as they make decisions and receive feedback on them. For these reasons, the computer might be expected to be a more efficient form of study than more conventional media. Most interesting, however, is the fact that students do voluntarily use the computer review, quarter after quarter, asking for it if for any reason it is unavailable or not yet announced.

The study indicated that the average participant improved his exam performance by using the computer review. Further research should investigate the relationship between the effect of the computer review and the individual's initial or untreated performance.

REFERENCES

Castleberry, S., & Lagowski, J.J. Individualized instruction using computer techniques. Journal of Chemical Education, 1970, 47, 91-96.

Gilman, D.A. The effect of feedback on learner's certainty of response and attitude toward instruction in a computer-assisted instruction program for teaching science concepts. Journal of Research in Science Teaching, 1969, 6, 171.

Grandey, R.C. An investigation of the use of computer-aided instruction in teaching students how to solve selected multistep general chemistry problems. Thesis, University of Illinois School of Education, Computer-Based Education Laboratory, November 1970.

Hammond, A.L. Computer-assisted instruction: many efforts, mixed results. Science, 1972, 176, 1005-1006.

Hansen, D.N., Dick, W., & Lippert, H.T. Research and implementation of collegiate instruction of physics via computer-assisted instruction. (Technical Report No. 3, 3 vols.) Tallahassee, Fla: Florida State University, 1968. (a)

Hansen, D.N., Dick, W., & Lippert, H.T. Semiannual progress report, July 1, 1967 through December 31, 1967. (CAI Center Report No. 6) Tallahassee, Fla.: Florida State University, 1968. (b)

Kromhout, O.M., Hansen, D.N., & Schwarz, G. A computer-assisted and -managed course in physical sciences. Proceedings of the Conference on Computers in Undergraduate Science Education. Illinois Institute of Technology, Chicago, August 1970.

Lagowski, J.J., Castleberry, S.J., & Culp, G.H. The impact of computer-based instructional methods in general chemistry. Proceedings of the 1972 Conference on Computers in Undergraduate Curricula. Georgia Institute of Technology, Atlanta, 1972, 77-88.

Rodewald, L.B., Culp, G.H., & Lagowski, J.J. The use of computers in organic chemistry instruction. Journal of Chemical Education, 1970, 47, 134-136.

Schwarz, G., Kromhout, O.M., & Edwards, S. Computers in physics instruction. Physics Today, 1969, 22, 41-49.

Stannard, C. Computer-assisted instruction in General Physics. Proceedings of the 1970 Conference on Computers in the Undergraduate Curricula. University of Iowa, Iowa City, 1970, 10.21-10.25.

APPENDIX

APPENDIX

LIST OF TOPICS

1 Algebra review
2 Algebraic manipulation of units of measure
3 Scientific notation
4 Order of magnitude
5 Problems to practice scientific notation
6 Mathematical functions
7 Scaling
8 Film, "Measuring small distances"
9 Film, "Change of scale"
10 Film, "Crystals"
11 Vectors
12 Refraction of light
13 Reflection of light
14 General properties of waves
15 Wave model of light
16 Particle model of light
17 Film, "Speed of light"
18 Film, "Simple waves"
19 Film, "Introduction to optics"
20 Film, "Frames of reference"
21 Film, "Conservation of energy"
22 Film, "Deflecting forces"
23 Film, "Inertia"
24 Film, "Inertial mass"
25 Film, "Forces"
26 Force
27 Energy
28 Mass and weight
29 Acceleration
30 Force=mass x acceleration
31 Momentum
32 26 review questions on course as a whole
33 Sample exam on first fourth of course
34 Sample exam on light
35 Film, "Energy and work"
36 Film, "Coulomb's Experiment"
37 Film, "Millikan's Experiment"
38 Film, "Charges and Kinetic Energy Transfer"
39 Film, "Mass of the electron"
40 Film, "Interference of photons"
41 Film, "Photons"
42 Film, "The Franck-Hertz experiment"
43 Work
44 Atom models - Thomson, Rutherford, Bohr
45 Induction
46 Modern physics, Photons, particles and waves
47 Magnetism
43 Light (particle and wave models)
49 Properties of gases
50 Electric currents, power and energy
51 More practice, algebraic manipulation of units

SAMPLE QUESTIONS

```
/GET,T16/UN=LIBRARY
/T16
 THE NEXT THREE QUESTIONS WILL HELP YOU REVIEW THE
 PARTICLE MODEL OF LIGHT.   1.  THE PARTICLE MODEL OF
 LIGHT ASSUMES THAT LIGHT PARTICLES            1.  ARE TOO
 LARGE TO BE DEFLECTED BY AIR.  2.  TRAVEL SO FAST WE
 DON'T NOTICE GRAVITATIONAL BENDING OF THEIR
 PATHS.          3.  ARE REFLECTED FROM BOTH HARD AND SOFT SURFACES.
?1
 NO, JUST THE OPPOSITE IS TRUE. THE PARTICLE MODEL
 OF LIGHT ASSUMES THAT LIGHT PARTICLES ARE EXTREMELY SMALL.
 KNOWING THIS, ANSWER THE QUESTION CORRECTLY.
?3
 NO, THE PARTICLE MODEL OF LIGHT ASSUMES THAT LIGHT PARTI-
 CLES ARE    REFLECTED ONLY FROM HARD SURFACES.
 TRY AGAIN.
?2
 CORRECT
 2.  DIFFICULTIES WERE ENCOUNTERED WITH THE PARTICLE MODEL
 OF LIGHT WHEN IT WAS DISCOVERED THAT            1.  LIGHT
 EXERTS A PRESSURE,            2.  LIGHT TRAVELS FASTER IN A
 VACUUM THAN ANY OTHER MEDIUM.  3.  LIGHT CAN BE BENT
 AT RIGHT ANGLES.
?1
 WRONG.  THE FACT THAT LIGHT EXERTS A PRESSURE IS
 COMPATIBLE WITH THE IDEA THAT LIGHT IS COMPOSED OF SMALL
 PARTICLES TRAVELING AT HIGH SPEED.  TRY AGAIN.
?3
 WRONG.  PARTICLES CAN BE BOUNCED OFF OF SMOOTH SURFACES
 MUCH AS A TENNIS BALL WOULD BOUNCE. NOW ANSWER CORRECTLY.
?4
 REMEMBER THAT IN REFRACTION, LIGHT BENDS TOWARD THE NORMAL
 WHEN ENTERING WATER FROM AIR. IF LIGHT WERE COMPOSED OF
 PARTICLES, THIS WOULD MEAN THAT THEY WOULD BE GOING
 FASTER IN THE AIR.
?2
 CORRECT
 3.  THE PARTICLE MODEL OF LIGHT CANNOT EXPLAIN
 1.  REFLECTION   2.  INTENSITY OF SOURCES OF LIGHT
    3.  DIFFRACTION          4.  STRAIGHT LINE LIGHT PATHS
    5.  HINT
?1
 WRONG.  RELFECTION CAN BE EXPLAINED IN THE PARTICLE MODEL
 OF LIGHT.  THE ANGLE OF INCIDENCE IS EQUAL TO THE ANGLE
 OF REFLECTION FOR A STEEL BALL HITTING A STEEL PLATE AT
 HIGH SPEED.
?2
 WRONG.  SOURCE STRENGTH CAN BE EXPLAINED BY THE PARTICLE
 MODEL IN TERMS OF HOW MANY PARTICLES ARE EMITTED PER
 SECOND.  NOW ANSWER CORRECTLY.
?4
 WRONG.  THE PARTICLE MODEL OF LIGHT ASSUMES THE PARTICLES
 TO BE TRAVELING SO FAST THAT ANY BENDING DUE TO GRAVITY
 COULD NOT BE SEEN TO ANY PRACTICAL EXTENT.  NOW ANSWER
 CORRECTLY.
?5
```

```
WOULD A STREAM OF PARTICLES BEND GOING THROUGH A SLIT?
?3
 CORRECT
 END OF THIS TOPIC
T16.
/JI+\ ]++
```