Florida Journal of
Educational Research
Fall 1997, Vol. 37(1) 49 - 68

Distributed Practice: More Bang for Your Homework Buck

Marie A. Revak
United States Air Force Academy

Abstract: Homework is commonplace in math classrooms, yet little research has been conducted on the differential effectiveness of homework for students with varying aptitudes. In this study, distributed practice homework bolstered the achievement of low achieving college math students. The sample consisted of 351 US Air Force Academy cadets all in their first semester of college. An algebra/trigonometry placement exam measured prior mathematics achievement and a subset of 25 items from the Math Anxiety Rating Scale measured math anxiety (Alexander & Martray, 1989). Data were analyzed using hierarchical multiple regression. Treatment group students outscored control group students on 4 of the 6 achievement measures without regard for prior math achievement or math anxiety ($\alpha = .05$).

Homework is commonplace in college mathematics courses, yet, with the exception of the inconclusive research investigating Saxon's incremental continuous review method (Abrams, 1989; Denson, 1989; Gianniotes, 1989; Johnson & Smith, 1987; Klingele & Reed, 1984; Parker, 1990; Reed, 1983; Rentschler, 1995; Roberts, 1994; Saxon, 1982), little research has been conducted on the content or quality of mathematics homework or on homework's effect on achievement. Other than a small study conducted by Hirsch, Kapoor, and Laing (1982, 1983; $N = 52$ first semester college calculus students), there is a lack of research investigating the differential effectiveness of homework for students with varying aptitudes (Austin, 1979; Featherstone, 1985; Hirsch et al., 1982, 1983; Kohler & Grouws, 1992; Peterson, 1971; Suydam, 1985). College students placed into precalculus and algebra courses have not yet mastered the fundamentals of algebra required to succeed in calculus. Many of these students have learned algebra as a set of rules for attacking specific types of problems. Homework problems in algebra courses usually consist of a set of problems related to the most recent problem type, that is, massed practice. With massed practice, students do not practice learning to differentiate between problem types. Yet, success in calculus requires students to determine when and where to use a variety of algebraic techniques.

By assigning homework problems related only to the most current course topics, mathematics educators have ignored the findings of cognitive psychology research recommending spaced over massed practice (Dempster, 1988, 1989; Reynolds & Glaser, 1964). Distributed practice is based on the aspect of information processing learning theory known as the *spacing effect.*

The spacing effect is the phenomenon in which "for a given amount of study time, spaced presentations yield substantially better learning than do massed presentations" (Dempster, 1988,

p. 627). The spacing effect has a long history in cognitive psychology and education research and is also referred to as distributed practice, continuous review, and spaced review (Cuddy & Jacoby, 1982; Dempster, 1988; Krug, Davis, & Glover, 1990; Reynolds & Glaser, 1964; Toppino & Gracen, 1985; Underwood, 1961). According to Dempster (1988), although distributed practice is "one of the most remarkable phenomena to emerge from laboratory research" (p. 627), there is little evidence that its potential has been realized in applied settings.

Research on distributed practice is situated in information processing theory (Ausubel, 1966). For over 25 years, cognitive psychology research has documented the benefit of spaced practice (Cuddy & Jacoby, 1982; Krug et al., 1990; Melton, 1970; Modigliani, 1976; Rea & Modigliani, 1985; Toppino & Gracen, 1985; Thorndike, 1971; Underwood, 1961). The most typical finding of this research was that as spacing increased, retention also increased. However, most research pertaining to the spacing effect has investigated the learning of simple word or number lists with time lags measured in seconds. Although the spacing effect is "one of the most robust phenomena discovered in memory research" (Rea & Modigliani, 1985, p. 11), results from cognitive psychology experiments do not necessarily transfer to complex learning tasks with longer spacings between reviews (Reynolds & Glaser, 1964). According to Dempster (1988), studies conducted from a basic research perspective and those conducted from an applied perspective frame two distinct research strands.

According to Cronbach and Snow, "an interaction is said to be present when a situation has one effect on one kind of person and a different effect on another" (1977, p. 3). Salomon (1972) described aptitude-treatment interaction (ATI) research as accomplishing two functions: improving instruction and advancing instructional theory. Salomon's compensatory ATI model proposed that ATI treatments could be developed to interact with aptitudes by circumventing their debilitating effects without trying to improve them. Snow (1977) advocated the use of some measure of general ability in all instructional research. Whenever affective traits are considered, researchers should expect that the regression of the trait will vary with ability. Cronbach and Snow (1977) assert that the anxiety experienced by an individual depends on the difficulty he or she has with a task. Task difficulty depends on an individual's ability and the characteristics of the task. Therefore, a complex task is more likely to create anxiety in persons of low ability than in more able persons (Cronbach & Snow).

From an ATI standpoint, Tobias (1976, 1989) hypothesized that students with lower prior achievement require more instructional support, and conversely, that as the level of prior achievement increases, less instructional support may be required. In their review of ATI research in science education, Koran and Koran (1984) referred to task organization as a manipulation likely to have an obvious effect on learning and a clear implication for ATI research. That is, material that is well organized should result in better achievement for high anxiety students (Koran & Koran, 1984). Similarly, Tobias (1989) and Bessant (1995) recommended clearly structured instruction as beneficial to highly anxious students. According to Sieber, O'Neill, and Tobias (1977), students high in anxiety may also benefit from opportunities for repetition of selected parts of the content.

In this study, the spacing principle was applied to Precalculus homework assignments (Hirsch et al., 1982, 1983; Peterson, 1971). The purpose of the study was to explore distributed practice homework assignments as one way to provide the instructional support and task organization necessary to increase the mathematics achievement of students with low prior mathematics achievement, high levels of mathematics anxiety, or both.

Three research questions were established:

(1) Will distributed practice homework assignments have a positive effect on Precalculus achievement?

(2) Will distributed practice homework assignments have a greater positive effect on Precalculus achievement than traditional homework assignments for students with low prior mathematics achievement?

(3) Will distributed practice homework assignments have a greater positive effect on Precalculus achievement than traditional homework assignments for students with high mathematics anxiety?

Method

Participants

The sample for the study consisted of all 375 United States Air Force Academy (USAFA) cadets enrolled in Precalculus during the 1995 fall semester. Enrollment in Precalculus was based on placement exam scores. Students scoring less than 50% on the Algebra/Trigonometry placement exam were placed into Precalculus. The sample represented about 28% of the first year students. Of the remaining first year students, 519 (about 39%) were placed into Calculus I, 344 (about 26%) were placed into Calculus II, and 103 (about 8%) were placed into Calculus III. All USAFA students are required to complete a sequence of core courses which includes at least two semesters of Calculus.

Natural attrition of students resulted in a changing sample size during the semester. At the time of the first exam, 351 of the original 375 cadets enrolled in Precalculus remained. Enrollment was 341 at the time of the second exam, 338 at the time of the third exam, and 333 at the end of the semester.

The USAFA has high admission standards. To qualify for admission, students must have good grades and athletic and leadership experience (Air Force Academy Admissions Office, 1995). In addition, students must be unmarried, without dependents, and between the ages of 17 and 21 (Air Force Academy Admissions Office). The mean Scholastic Achievement Test (SAT) math achievement score for incoming Air Force Academy students was 660 (recomputed to reflect the 1995 recentering of the SAT) and the mean for the math portion of the American College Test (ACT) for incoming students was 29.3 (B. A. Branum, personal communication, September 6, 1995). The average high school grade-point average for incoming cadets was 3.85 (B. A. Branum) and 89% of entering cadets ranked in the top fifth of their high school class (Air Force Academy Admissions Office).

The USAFA class of 1999 consisted of 1367 students, 1353 from the United States and 14 from 13 foreign countries (Lockhart, 1995). Included were 238 minority members (17%) and 219 women

(16%). Of the United States students, 1086 (82%) were White, 56 (4%) were Black, 85 (6%) were Hispanic, 72 (6%) were Asian American, and 19 (1%) were Native American (B. A. Branum, personal communication, September 6, 1995).

Instruments

Prior Mathematics Achievement. The percentage correct on an Algebra/Trigonometry placement exam was used as the measure of prior mathematics achievement. The placement exam contained 35 multiple choice items (25 algebra items and 10 trigonometry items) and was machine scored. The test was validated for content in 1995 by faculty members of the USAFA math placement team. The tests were found to have high predictive validity for placing students into Precalculus as their first mathematics course, with 87% of students successfully completing Precalculus with a grade of B+ or less (A's and A-'s were considered erroneously placed; W. A. Kiele, personal communication, April 5, 1995). Many of the placement test items are *anchored*, that is, used again from year to year. The use of anchored items improves test stability and reliability.

The placement exams were administered under standardized conditions a few days after the students arrived at the Air Force Academy. Students took the exam in large lecture halls proctored by instructors. Standardized directions were printed on the first page of the exam and read aloud by the proctors. All students had identical time limits. The use of calculators was not permitted.

Mathematics Anxiety. Mathematics anxiety was measured by a subset of items from the Math Anxiety Rating Scale (MARS), college and adult version (Suinn, 1972). The MARS is a 98-item self-rating scale set in a five point Likert format designed as a diagnostic or screening tool for measuring mathematics anxiety. Scores on each MARS item represent the level of anxiety reported for a specific situation. Selections range from 1 representing *not at all anxious* to 5 representing *very much anxious*. An overall mathematics anxiety score is achieved by summing the individual item scores.

Since its publication in 1972, the MARS has been the prevailing instrument for measuring mathematics anxiety (Alexander & Martray, 1989). Alexander and Martray (1989) used a two-staged factor analysis to develop an abbreviated version of the MARS. Their first factor analysis reduced the 98-item MARS to 69 items by selecting the items most highly correlated to each of five identified factors. The 69-item MARS was again abbreviated by application of factor analysis. Items that correlated highly with each of three identified factors were selected for Alexander and Martray's 25-item abbreviated MARS. The 25-item MARS was shown to have high internal consistency within each of the three factors (Cronbach alpha of .96, .86, and .84, respectively). In addition, correlation between the 25-item and 69-item versions of the MARS was found to be high (r = .93) and test-retest reliability after two weeks was also high (r = .86). Alexander and Martray (1989) declared that the 25-item MARS was a "psychometrically equivalent alternative" to the 98-item MARS, while being more efficient, less costly, and easier to implement (p. 149).

The abbreviated MARS was administered to the control and treatment groups during the fifth week of class. A standardized set of instructions was read aloud by the instructors. Students were assured that their instructors would not have access to the individual MARS scores. The surveys were machine scored.

Precalculus Achievement. Six variables were used to measure student achievement in Precalculus. Included were four hourly exams, a final exam, and the final course percentage grade. The second, third, and fourth hourly exams included mostly new material with a few (20%) items testing material covered on earlier exams. The final exam was comprehensive. All exam items were written by members of the USAFA Department of Mathematical Sciences and the same exam was administered to all sections. Parallel make-up exams were administered to the few students who missed an exam. All exams were composed of multiple choice and open-ended items. The exams were reviewed by several mathematics instructors for content validity. Split-half reliability coefficients for all exams were calculated using the Spearman-Brown prophecy formula (Fraenkel & Wallen, 1993) and were found to be acceptable (coefficients ranged from .69 to .83).

As standard procedure at the Air Force Academy, exams were administered to the entire course population during the same period of time. Students were assigned to lecture halls and classrooms. Standardized directions were printed on the first page of the exams and read aloud by the instructors administering the exam. All students had identical time limits.

The four hourly exams were given from 7:00 to 7:50 a.m., before the start of classes. Students in both the treatment and control groups were permitted to use calculators on all four hourly exams.

The final exam was given seven days after the last class and was administered in two parts. Students were given 1 hour to complete Part I of the exam and 2 hours and 50 minutes to complete Part II. With the exception of five items, Part I was identical to the Algebra/Trigonometry Placement Exam. Part II was a cumulative exam containing mostly anchored items. Students were not permitted to use calculators on Part I of the final exam. The use of calculators was permitted on Part II.

Multiple choice exam items for all exams were machine scored. Standardized rubrics were used to score open-ended items. In most cases, one instructor was assigned to score one item on all exam papers. For exam items that were scored by more than one instructor, a sample of 30 exams (15 from the treatment group and 15 from the control group) was selected for duplicate scoring. Inter-scorer reliability was calculated and found to be high (correlation coefficients ranged from .87 to .99). All exam scores were converted to percentages.

The final course percentage grade was based on the following sub-scores: (a) four hourly exams, 45%; (b) final exam, 30%; (c) three written exercises, 5%; (d) course project, 5%; (e) three group problem solving exercises, 5%; and (f) quiz, homework, and participation points awarded by the individual instructors, 10%.

Procedures

The experiment employed the ATI compensatory instructional model. The distributed practice treatment was designed to interact with the low prior achievement and high mathematics anxiety student aptitudes by circumventing or neutralizing their debilitating effects (Salomon, 1972). As recommended in previous ATI and homework research, the duration of the treatment was one semester, the entire duration of the Precalculus course (Austin, 1979; Becker, 1970; Becker & Young, 1978; Cronbach & Snow, 1977; Holtan, 1982; Koran & Koran, 1984; Snow, 1977).

Although assignment to Precalculus sections was not purely random, student course schedules at the USAFA are computer generated and students (especially first year students) have very few choices in their schedules. The treatment group consisted of approximately 46% of the Precalculus students (161 students divided into eight sections). The control group consisted of the remaining students enrolled in Precalculus (190 students divided into nine sections).

To minimize instructor workload, each instructor was assigned either all treatment sections or all control sections. The Precalculus sections were taught by eight different instructors; three instructors taught treatment group sections and five instructors taught control group sections.

All instructors were active duty members of the United States Air Force. Degree levels for instructors ranged from bachelor to doctoral with most instructors holding a master of science degree. Instructor experience level varied from first year instructors to a seasoned instructor with over 20 years teaching experience. Although most of the instructors had some prior teaching experience, few had prior experience teaching Precalculus. Both experienced and inexperienced instructors were assigned to each group in an attempt to equalize instructor experience across groups. When weighted by the number of sections, the mean instructor experience level for each group was 2.6 years. The median experience level was 2 years.

The course topics, textbook, handouts, reading assignments, and graded assignments (with the exception of quiz, homework, and participation points) were identical for the treatment and control groups. The listing of homework assignments in the syllabus differed between groups. The control group was assigned daily homework related to the topic(s) presented that day in class. Peterson (1971) calls this the vertical model for assigning mathematics homework. The treatment group was assigned homework in accordance with a distributed organizational pattern that combines practice on current topics and reinforcement of previously covered topics. Under the distributed model, approximately 40% of the problems on a given topic were assigned the day the topic was first introduced, with an additional 20% assigned on the next lesson and the remaining 40% of problems on the topic assigned on subsequent lessons (Hirsch et al., 1983). In Hirsch's research and in this study, after the initial homework assignment, problem(s) representing a given topic resurfaced on the 2nd, 4th, 7th, 12th, and 21st lesson. Consequently, treatment group homework for lesson one consisted of only one topic; homework for lessons two and three consisted of two topics; and homework for lesson four through six consisted of three topics. This pattern continued as new topics were added and was applied to all non-exam, non-laboratory lessons.

As shown by Tables 1 and 2, the same homework problems were assigned to both groups with only the pattern of assignment differing. Because of the nature of the distributed practice model, homework for the treatment group contained fewer problems (relative to the control group) early in the semester with the number of problems increasing as the semester progressed. Later in the semester, homework for the treatment group contained more problems (relative to the control group). As shown in Tables 1 and 2, by the end of the semester, both groups had been assigned precisely the same homework problems.

Table 1
Homework Problems Assigned to the Control Group

Lesson Number											Number of Problems
1	A1	A2	A3	A4	A5	A6	A7	A8			8
2	B1	B2	B3	B4	B5	B6	B7	B8	B9		9
3	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	10
4	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	10
5	E1	E2	E3	E4	E5	E6	E7	E8			8
6	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	10
7	G1	G2	G3	G4	G5	G6	G7	G8			8
8	H1	H2	H3	H4	H5	H6	H7	H8			8
9	I1	I2	I3	I4	I5	I6	I7	I8	I9		9
10	J1	J2	J3	J4	J5	J6	J7	J8	J9	J10	10
11	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	10
12	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	10
13	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	10
14	N1	N2	N3	N4	N5	N6	N7	N8			8
15	O1	O2	O3	O4	O5	O6	O7	O8	O9		9
16	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	10
17	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	10
18	R1	R2	R3	R4	R5	R6	R7	R8	R9		9
19	S1	S2	S3	S4	S5	S6	S7	S8	S9		9
20	T1	T2	T3	T4	T5	T6					6
21	U1	U2	U3	U4	U5	U6	U7	U8			8
22	V1	V2	V3	V4	V5	V6	V7	V8			8
23	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	10
24	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	10
25	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8			8
26	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9		9
27	a1[a]	a2	a3	a4	a5	a6	a7	a8			8
28	b1	b2	b3	b4	b5	b6	b7	b8	b9		9
29	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	10
30	d1	d2	d3	d4	d5	d6	d7	d8			8
Total											269

Note. A1 represents the first problem in topic A, A2 represents the second problem, etc.
[a]Homework on topic "T" was not distributed due to a late syllabus change.
[b]Upper and lower case letters represent different topics.

Table 2
Homework Problems Assigned to the Treatment Group

Lesson Number																Number of Problems
1	A1	A2	A3													3
2	A4	B1	B2	B3												4
3	B4	B5	C1	C2	C3	C4										6
4	A5	C5	C6	D1	D2	D3	D4									7
5	B6	D5	D6	E1	E2	E3										6
6	C7	E4	F1	F2	F3	F4										6
7	A6	D7	F5	F6	G1	G2	G3									7
8	B7	E5	G4	H1	H2	H3										6
9	C8	F7	H4	I1	I2	I3										6
10	D8	G5	I4	I5	J1	J2	J3	J4								8
11	E6	H5	J5	J6	K1	K2	K3	K4								8
12	A7	F8	I6	K5	K6	L1	L2	L3	L4							9
13	B8	G6	J7	L5	L6	M1	M2	M3	M4	M5						10
14	C9	H6	K7	M6	M7	N1	N2	N3	N4							9
15	D9	I7	L7	L8	N5	O1	O2	O3	O3							9
16	E7	J8	M8	O4	O5	O6	P1	P2	P3	P4						10
17	F9	K8	N6	P5	P6	P7	Q1	Q2	Q3	Q4						10
18	G7	L9	O7	Q5	Q6	Q7	R1	R2	R3							9
19	H7	M9	P8	R4	R5	R6	S1	S2	S3	S4						10
20	I8	N7	Q8	S5	S6	T1[a]	T2[a]	T3[a]	T4[a]	T5[a]	T6[a]					11
21	A8	J9	O8	R7	U1	U2	U3	U4								8
22	B9	K9	P9	S7	U5	U6	V1	V2	V3	V4	V5					11
23	C10	L10	Q9	V6	W1	W2	W3	W4	W5	W6						10
24	D10	M10	R8	U7	W7	W8	X1	X2	X3	X4	X5	X6				12
25	E8	N8	S8	V7	X7	X8	Y1	Y2	Y3	Y4	Y5					11
26	F10	O9	W9	Y6	Y7	Z1	Z2	Z3	Z4	Z5						10
27	G8	P10	U8	X9	Z6	Z7	Z8	a1[b]	a2	a3	a4					11
28	H8	Q10	V8	Y8	a5	a6	a7	b1	b2	b3	b4	b5	b6			13
29	I9	R9	W10	Z9	b7	b8	b9	c1	c2	c3	c4	c5	c6	c7		14
30	J10	S9	X10	a8	c8	c9	c10	d1	d2	d3	d4	d5	d6	d7	d8	15
Total																269

Note. A1 represents the first problem in topic A, A2 represents the second problem, etc.
[a]Homework on topic "T" was not distributed due to a late syllabus change.
[b]Upper and lower case letters represent different topics.

Because homework was the key manipulated variable in this experiment, and because larger effects on achievement were sometimes found when homework was graded (Austin, 1979; Lai, 1994; Paschal, Weinstein, & Walberg, 1984), instructors were directed to collect all homework. Homework was checked and coded for correctness and completion on a three point scale (0 = less than one-third complete and correct, 1 = one-third to two-thirds complete and correct, and 2 = more than two-thirds complete and correct).

Instructors in both groups were encouraged to use class time to discuss and review the assigned homework problems. Prior to the second, third, and fourth exam, and at the end of the semester, both groups spent one lesson in review. Review lessons were planned by the individual instructors. Classroom observations and student and instructor surveys were used to ensure that the treatment was administered as planned and directed.

Results

The means and standard deviations for the entire sample and for the treatment and control groups on measures of prior achievement, mathematics anxiety, and Precalculus achievement are reported in Table 3. Hierarchical multiple regression was employed to test the hypotheses[1]. Three sets of independent variables were defined. Set A, the covariate set, contained two variables: (a) prior math achievement, and (b) mathematics anxiety. Set B contained the group membership variable (treatment group or control group). Set C, the two-way interaction set, contained two interaction variables: (a) Prior Achievement × Treatment, and (b) Anxiety × Treatment. The dependent variable in this study was Precalculus achievement. Precalculus achievement was measured as the semester progressed and produced six scores: four hourly exam scores, a final exam score, and a final course percentage grade. By analyzing each measure of achievement separately, the goal was to determine whether the length of treatment had an impact on achievement with the expectation that the distributed practice treatment would have a cumulative effect (Austin, 1979).

[1] Multiple regression was selected as the analysis tool due to its ability to handle unequal cell sizes and quantitative independent variables (Cohen & Cohen, 1983).

Table 3
Descriptive Statistics for Measures of Prior Achievement, Anxiety, and Precalculus Achievement

	Prior achievement	Math anxiety	1st Exam	2nd Exam	3rd Exam	4th Exam	Final Exam	Course grade
				All Students				
N	351	351	351	341	338	333	317	333
M	35.88	51.51	80.43	70.67	70.48	65.21	70.43	74.83
SD	8.74	14.44	13.25	13.67	13.10	13.55	11.13	8.55
min	5.00	28.00	14.81	21.48	29.63	23.70	20.33	35.00
max	50.00	99.00	99.26	96.30	100.00	100.00	94.67	96.76
				Treatment Group				
n	161	161	161	160	157	155	144	155
M	36.51	49.48	82.69	73.58	70.71	68.28	71.70	76.96
SD	8.09	12.96	11.89	12.79	12.99	12.73	10.60	7.84
min	5.00	28.00	28.99	37.78	29.63	23.70	28.61	46.43
max	50.00	93.00	99.26	95.56	98.52	100.00	93.56	94.83
				Control Group				
n	190	190	190	181	181	178	173	178
M	35.36	53.23	78.51	68.10	70.27	62.54	69.41	72.97
SD	9.24	15.42	14.05	13.93	13.23	13.71	11.48	8.72
min	5.00	28.00	14.81	21.48	30.37	28.15	20.33	35.00
max	47.50	99.00	99.26	96.30	100.00	99.26	94.67	96.76

Note. All prior achievement and achievement scores are measured in percent.

Hypothesis Testing

Table 4 shows the results of the step-by-step hierarchical regressions as the three sets of independent variables were added.

Effect of the Covariates

Step one of the hierarchical multiple regression analyses tested the effect of the covariates (Set A, prior mathematics achievement and mathematics anxiety) on Precalculus achievement. Set A was regressed on each of the six measures of Precalculus achievement. A significant proportion of variance in all six measures of Precalculus achievement was explained by prior mathematics achievement and mathematics anxiety (see Table 4).

Table 4
Hierarchical Multiple Regression Analysis - Main Effect and Interaction Effect

Independent variable sets	Cumulative R^2	*df*	*F*	Variable sets added	Increment to R^2	*df*	*F* of the increment
				First Exam			
A	.239	2, 348	54.66***	A			
A, B	.249			B	.010	1, 347	4.73*
A, B, C	.251			C	.001	2, 345	0.30
				Second Exam			
A	.169	2, 338	34.42***	A			
A, B	.193			B	.023	1, 337	9.78**
A, B, C	.198			C	.005	2, 335	1.07
				Third Exam			
A	.069	2, 335	12.46***	A			
A, B	.070			B	.000	1, 334	0.13
A, B, C	.073			C	.003	2, 332	0.56
				Fourth Exam			
A	.093	2, 330	16.97***	A			
A, B	.124			B	.031	1, 329	11.57**
A, B, C	.128			C	.004	2, 327	0.71
				Final Exam			
A	.121	2, 314	21.61***	A			
A, B	.125			B	.004	1,313	1.27
A, B, C	.126			C	.001	2,311	0.24
				Final Course Grade			
A	.203	2, 330	41.90***	A			
A, B	.234			B	.031	1, 329	13.48**
A, B, C	.236			C	.002	2, 327	0.36

Note. Set A = placement test score and math anxiety score.
Set B = group membership.
Set C = two-way interactions.
**p* < .05 ***p* < .01 ****p* < .001

Main Treatment Effect

Step two of the hierarchical analyses tested for a main effect due to the distributed practice treatment. The covariates (Set A) and the group membership variable (Set B) were regressed on each of the six measures of Precalculus achievement. Tests of the semi-partial correlation coefficients revealed that, when the covariates were controlled for, the distributed practice treatment accounted for a statistically significant proportion of the variance in Precalculus achievement in all but the third exam and final exam (see Table 4).

Two-Way ATI Effects

Step three of the hierarchical regression analysis added the two aptitude-treatment interaction variables (Set C). The semi-partial correlation coefficients were tested to determine whether the interactions accounted for any variance in Precalculus achievement above what had already been accounted for by prior achievement, anxiety, and the distributed practice treatment. The effect of the two-way ATIs was not statistically significant for any of the six measures of Precalculus achievement (see Table 4).

Instructor Effects

Regression analysis was also used to determine whether there was a significant effect due to instructor after prior achievement, anxiety, and the treatment were controlled for. A two-step hierarchical regression was employed with the covariate and group membership variables (Set A') entered in the first step and the dummy-coded instructor variable set (Set B') added in the second step. Semi-partial correlation coefficients were calculated and F-tests were conducted. This analysis revealed that the instructors did not contribute to the variance in Precalculus achievement beyond what was already accounted for by prior achievement, anxiety, and the distributed practice treatment.

Other Analyses

Study Time

The USAFA routinely collects study time data. After each exam, a large sample of cadets (at least 60% of the course population) anonymously reported the amount of time (in minutes) spent studying for the exam. Time spent studying was approximately equal for both groups (see Table 5). Descriptive data revels that, for both the treatment and control group, study time for the third exam was at least 16% greater than study time for any other exam. Study time for the final exam was at least 68% greater than study time for any of the hourly exams (see Table 5).[2]

[2] Since the group of students sampled for study time for one exam was not necessarily independent of the group of students sampled for study time for other exams, inferential statistical tests of study times between exams are not appropriate.

Table 5
Analysis of Study Times for Exams

Exam	Treatment mean (in minutes)	Control mean (in minutes)	*df*	*t*
1	88.4	84.5	333	0.59
2	95.4	97.4	296	0.23
3	117.6	116.9	305	0.08
4	100.8	93.2	274	0.77
Final	198.1	235.9	128	1.30

All p values > .20.

Effect of Homework on Exam Scores

Five separate regressions were performed to determine whether homework scores could predict a significant proportion of variance in exam scores. Block homework scores explained a statistically significant proportion of variance in all hourly exam scores. Similarly, the total homework score explained a statistically significant proportion of variance in the final exam score (see Table 6).

Table 6
Effect of Homework on Exam Scores

Exam	*r*	R^2	df	F
1	.39	.151	1, 349	62.07***
2	.33	.109	1, 339	41.54***
3	.33	.109	1, 336	41.22***
4	.30	.090	1, 331	32.67***
Final	.39	.153	1, 315	56.96***

***$p < .001$

Discussion and Conclusions

Distributed Practice Effect

The distributed practice treatment produced a statistically significant main effect on four out of six measures of Precalculus achievement (three hourly exams and the final course percentage grade). These findings are in agreement with results reported by Friesen (1975), Parker (1990), Peterson (1970), Reed (1983; Klingele & Reed, 1984), and Saxon (1982). The treatment did not produce a statistically significant main effect on the third exam or final exam.

Effect sizes were calculated to better interpret the practical significance of the distributed practice treatment. The treatment produced an effect size (f^2) of 0.013 on the first exam, 0.029 on the second exam, 0.035 on the fourth exam, and 0.040 on the final course percentage grade. Although the effect sizes appear to be small, the treatment group outscored the control group in every case. A mean difference of 5.13 percentage points on the first, second, and fourth exam translates to an advantage of about a third of a letter grade for students in the treatment group. In addition, higher minimum scores earned by the treatment group may indicate that the distributed practice treatment served to eliminate the extremely low scores (refer to Table 3). As postulated by Austin (1979), the distributive practice treatment appeared to have a cumulative effect.

Because the distributed practice treatment produced a significant main effect on all but one of the hourly exams, a plausible explanation for this aberration was sought. The treatment and control groups achieved nearly equal scores on the third exam (treatment mean = 70.71 and control mean = 70.27). Although the two groups spent nearly equal time studying for the exam (treatment mean = 117.6 minutes and control mean = 116.9 minutes), both groups reported spending much more time studying for the third exam than they spent studying for any of the other three hourly exams. The third exam occurred after mid-semester progress reports which may have motivated students to devote more time to studying. It is possible that the additional study time imitated the distributed practice treatment by allowing for more repetitions of problem types.

Oddly, the distributed practice treatment did not produce a significant effect on final exam scores. One possible cause for the disparity was the USAFA policy exempting the top performers from the final exam. Of the 16 exempted students, 11 were from the treatment group with only 5 from the control group. It is likely that the treatment group would have outscored the control group on the final exam if these top performers had taken the exam. In addition, increased study time for the final exam may have influenced the results. Because the final exam was scheduled late during final exam week, study time for the exam was not only longer, but more widely distributed. The benefits of the longer and more dispersed study time may have been similar to the benefits created by the distributed practice treatment.

Aptitude-Treatment Interaction Effects

Two significant two-way interactions were expected: (a) Prior Mathematics Achievement × Treatment, and (b) Mathematics Anxiety × Treatment. Neither of these interactions was found to explain a significant proportion of variance in Precalculus achievement beyond what had already been explained by the covariates and the distributed practice treatment.

The sample in this study, first year students on the low mathematics ability track at the Air Force Academy, may provide some explanation for the lack of significant interaction effects. Students on the average track are typically enrolled in Calculus I during the Fall semester and Calculus II during the Spring semester. Similarly, those with high math ability are usually enrolled in Calculus II or Calculus III during the Fall semester. Because mathematics achievement has been found to correlate negatively with mathematics anxiety (Berenson, Carter, & Norwood, 1992; Clute, 1984; Coleman, 1991; Cooper & Robinson, 1989; Covington & Omelich, 1987; Frary & Ling, 1983; Gliner, 1987; Hembree, 1990; Lawson, 1993; McCoy, 1992; Richardson & Suinn, 1972), the students placed into Precalculus were probably relatively high in mathematics anxiety. Aptitude-treatment interactions are not expected to be as strong when students have comparable aptitudes. The homogeneity of this group may have nullified the expected two-way interaction effects.

The results of this study challenge the results reported by Hirsch and his colleagues (1982, 1983). Hirsch et al. found significant Prior Achievement × Treatment ATIs on three out of five measures of Calculus I achievement. In all three cases, the distributed practice treatment was beneficial to students scoring at or below the mean on an algebra and analytic geometry pre-test. It is not known whether the students in Hirsch's study were grouped homogeneously.

Limitations

This study was limited by the length of the semester and the number of homework assignments. By following the homework pattern advocated by Hirsch et al. (1982, 1983), homework for topics introduced after the tenth lesson could not be fully distributed. Homework for each topic was assigned in the order listed in the textbook, in which the easier problems preceded the more difficult ones. For the treatment group, this meant that the easiest problems were assigned early in the distribution pattern with the hardest problems assigned in the later stages of the distribution. The treatment may have been more effective if the difficulty level of problems within each assignment was mixed. Similarly, the distributed practice treatment may be more effective when applied to courses of longer duration.

Several factors may limit the generalizability of this study. Although the sample was large, the subjects, being military academy cadets, may not be representative of typical high school or college students. Overall, students attending the USAFA are a fairly homogeneous group with similar academic and career goals. The limited external validity due to the controlled atmosphere at the Air Force Academy serves to strengthen the internal validity of the study. Threats due to subject characteristics, mortality, location, history, and subject attitude have been minimized due to the controlled environment at the USAFA (Fraenkel & Wallen, 1993).

Certain threats to internal validity remain. Although it cannot be assumed that instructors with similar experience levels are equally effective, this study and a previous study conducted at the USAFA found that instructor experience was not a significant contributor to achievement variance (Thompson, Mitchell, Coffin, & Hassett, 1979). It is possible that one or more instructors were biased, either for or against the distributed practice treatment. A Hawthorne effect may have resulted if the students in the treatment group recognized that they were receiving special treatment in the way of distributed practice homework assignments (Fraenkel & Wallen, 1993). Conversely, students assigned to the control group may have suffered a demoralization effect (Fraenkel & Wallen). In addition, the treatment may have had a negative impact on the achievement of the treatment group if exam items were related to homework problems not yet assigned due to the distributed practice syllabus. Finally, it is possible that the treatment was not fully confined to the treatment group. Although survey responses indicated that students rarely studied with students who used a different syllabus, it is possible that cadets discussed homework problems with students from other sections.

Recommendations for Future Research

Distributed practice homework has been shown to be beneficial to students on the low mathematics track at the USAFA. Testing of the distributed practice treatment on medium and high ability students is recommended. In addition, different variations of spaced review should be investigated across a wide variety of students, institutions, and mathematics courses. Because the collection and grading of homework may have caused a higher than average homework completion rate, this study should be replicated in an environment where homework is not collected.

Future studies of this kind should include the study time variable. The study time data in this experiment indicate that the distributed practice treatment had the greatest impact when less time was devoted to studying for an exam. This finding appears to support the theory that distributed practice assignments receive more attention than massed assignments. An analysis of how students use their study time could help shed light on why and how this phenomenon occurs.

According to Holtan (1982), the value of the distributed practice treatment may well be in the delayed retention of the skills and concepts practiced. Follow-up retention tests are recommended for the students taking part in this study.

The multiple correlations revealed in this study accounted for less than 26% of the variance in all measures of achievement. This suggests that the contribution of other variables such as motivation, attitude, and study habits should be examined. Systematic research in this area should help identify the students who will benefit most from distributed practice assignments and contribute to the theoretical structure of ATI.

Summary

This study has documented a significant positive correlation between homework scores and exam scores. Homework scores were found to account for between 10% and 15% of the variability in exam scores. Meaningful homework may be viewed as an important component in mastering mathematics course material.

Enrollments in remedial mathematics college courses are on the rise (Berenson et al., 1992) and 90% of college mathematics enrollments are in elementary calculus, elementary statistics, and courses prerequisite to them (National Research Council, 1989). There is great potential for application of the distributed practice model. Mathematics achievement is still the principal gateway for students preparing to enter technical and scientific careers, and distributed practice may help foster success in these pivotal math courses.

References

Abrams, B. J. (1989). A comparison study of the Saxon algebra text. *Dissertation Abstracts International, 51,* 2551-A.

Air Force Academy Admissions Office (1995). *United States Air Force Academy Catalog, 1995-1996.* United States Air Force Academy, CO.

Alexander, L., & Martray, C. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development, 22,* 143-150.

Austin, J. D. (1979). Homework research in mathematics. *School Science and Mathematics, 79,* 115-121.

Ausubel, D. P. (1966). Early versus delayed review in meaningful learning. *Psychology in the schools, 3,* 195-198.

Becker, J. P. (1970). Research in mathematics education: The role of theory and of aptitude-treatment interaction. *Journal for Research in Mathematics Education, 1,* 22.

Becker, J. P., & Young, C. D., Jr. (1978). Designing instructional methods in mathematics to accommodate different patterns of aptitude. *Journal for Research in Mathematics Education, 9,* 4-19.

Berenson, S. B., Carter, G., & Norwood, K. S. (1992). The at-risk student in college developmental algebra. *School Science and Mathematics, 92,* 55-58.

Bessant, K. C. (1995). Factors associated with types of mathematics anxiety in college students. *Journal for Research in Mathematics Education, 26,* 327-345.

Clute, P. S. (1984). Mathematics anxiety, instructional method, and achievement in a survey course in college mathematics, *Journal for Research in Mathematics Education, 15,* 50-58.

Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coleman, B. L. (1991). A study of the prevalence and intensity of mathematics anxiety in college students and preservice teachers at a large southern university. *Dissertation Abstracts International, 52,* 4253-A.

Cooper, S. E., & Robinson, D. A. G. (1989). The influence of gender and anxiety on mathematics performance. *Journal of College Student Development, 30,* 459-461.

Covington, M. V., & Omelich, C. L. (1987). "I knew it cold before the exam": A test of the anxiety blockage hypothesis. *Journal of Educational psychology, 79,* 393-400.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and Instructional Methods: A Handbook for Research on Interactions.* New York: Irvington Publishers, Inc.

Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Learning and Verbal Behavior, 21,* 451-467.

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results. *American Psychologist, 43,* 627-634.

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*(4), 309-330.

Denson, P. S. (1989). A comparison of the effectiveness of the Saxon and Dolciani texts and theories about teaching of high school algebra. *Dissertation Abstracts international, 50,* 3173-A.

Featherstone, H. (1985). What does homework accomplish? *Principal, 65(2),* 6-7.

Fraenkel, J. R., & Wallen, N. E. (1993). *How to design and evaluate research in education.* New York: McGraw-Hill.

Frary, R., & Ling, J. (1983). A factor analytic study of mathematics anxiety. *Educational and Psychological Measurement, 43,* 985-993.

Friesen, C. D. (1975). The effect of exploratory and review homework exercises upon achievement, retention, and attitude in a first-year algebra course. *Dissertation Abstracts International, 36,* 6527-A.

Gianniotes, K. K. (1989). Cumulative versus massed practice in Algebra I. *Dissertation Abstracts International, 50,* 1587-A.

Gliner, G. S. (1987). The relationship between mathematics anxiety and achievement variables. *School Science and mathematics, 87,* 81-87.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21,* 33-46.

Hirsch, C. R., Kapoor, S. F., & Laing, R. A. (1982). Alternative models for mathematics assignments. *International Journal of Mathematical Education in Science and Technology, 13,* 243-252.

Hirsch, C. R., Kapoor, S. F., & Laing, R. A. (1983). Homework assignments, mathematical ability, and achievement in calculus. *Mathematics and Computer Education, 17*(1), 51-57.

Holtan, B. (1982). Attribute-Treatment interaction research in mathematics education. *School Science and Mathematics, 82,* 593-602.

Johnson, D. M., & Smith, B. (1987). An evaluation of Saxon's algebra text. *Journal of Educational Research, 81*(2), 97-102.

Klingele, W. E., & Reed, B. W. (1984). An examination of an incremental approach to mathematics. *Phi Delta Kappan, 65*(10), 712-713.

Kohler, M. S., & Grouws, D. A. (1992). Mathematics teaching practices and their effects. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (115-126). New York: MacMillan.

Koran, M. L., & Koran, J. J. (1984). Aptitude-treatment interaction research in science education. *Journal of Research in Science Teaching, 21,* 793-808.

Krug, D., Davis, T. B., & Glover, J. A. (1990). Massed versus distributed repeated reading: A case of forgetting helping recall? *Journal of Educational Psychology, 82,* 366-371.

Lai, W. Y-K. (1994). The influence of written teacher comments and differing amounts of homework upon student achievement in basic mathematics. *Dissertation Abstracts international, 54,* 4021-A.

Lawson, V. J. (1993). Mathematics anxiety, test anxiety, instructional methods, and achievement in a developmental mathematics class. *Dissertation Abstracts International, 53,* 3479-A.

Lockhart, M. (1995, July 6). Class of '99 enters Academy. *The Falcon Flyer,* 8.

McCoy, L. P. (1992). Correlates of mathematics anxiety. *Focus on Learning Problems in Mathematics, 14,* 51-57.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal learning and Verbal Behavior, 9,* 596-606.

Modigliani, V. (1976). Effects on later recall by delaying initial recall. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 609-622.

National Research Council (1989). *Everybody Counts.* Washington, DC: National Academy Press.

Parker, J. K. (1990). Effects of incremental continuous review homework format on seventh-grade mathematics achievement. *Dissertation Abstracts International, 52,* 834-A.

Paschal, R. A., Weinstein, T., & Walberg, H. J. (1984). The effects of homework on learning: A quantitative synthesis. *The Journal of Educational Research, 78,* 97-104.

Peterson, J. C. (1970). Effect of exploratory homework exercises upon achievement in eighth grade mathematics. *Dissertation Abstracts international, 30,* 4339-A.

Peterson, J. C. (1971). Four organizational patterns for assigning mathematics homework. *School Science and Mathematics, 71,* 592-596.

Rea, C. P., & Modigliani, V. (1985). The effect of expanded Vs. massed practice on the retention of multiplication facts and spelling lists. *Human Learning, 4,* 11-18.

Reed, B. W. (1983). Incremental, continuous-review versus conventional teaching of algebra. *Dissertation Abstracts International, 44,* 1716-A.

Rentschler, R. V., Jr. (1995). The effects of Saxon's incremental review on computational skills and problem-solving achievement of sixth-grade students. *Dissertation Abstracts International, 56,* 484-A.

Reynolds, J. H., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology, 55,* 297-308.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19,* 551-554.

Roberts, F. H. (1994). The impact of the Saxon mathematics program on group achievement test scores. *Dissertation Abstracts International, 55,* 1498-A.

Salomon, G. (1972). Heuristic models for the generation of aptitude-treatment interaction hypotheses. *Review of Educational Research, 42,* 327-343.

Saxon, J. (1982). Incremental development: A breakthrough in mathematics. *Phi Delta Kappan, 63,* 482-484.

Sieber, J. E., O'Neill, H. F., & Tobias, Sigmund. (1977). *Anxiety, Learning, and Instruction.* Hillsdale, NJ: Erlbaum.

Snow, R. E. (1977). Individual differences and instructional theory. *Educational Researcher, 6*(10), 11-15.

Suinn, R. M. (1972). *Mathematics Anxiety Rating Scale Informational Brief.* [Brochure]. Rocky Mountain Behavioral Sciences Institute, Fort Collins, CO.

Suydam, M. N. (1985). Research report: Homework Yes or No. *Arithmetic Teacher, 32*(5), 56.

Thompson, S. B., Mitchell, C. R., Coffin, R. C., & Hassett, M. J. (1979). The H2 experiment: A comparison of homogeneous and heterogeneous aptitude sectioning in core mathematics. (Tech. Rep. No. 79-9). United States Air Force Academy, CO, Dean of the Faculty.

Thorndike, E. L. (1971). *The Fundamentals of Learning.* New York: AMS Press. (Original work published 1932).

Tobias, Sigmund. (1976). Achievement treatment interactions. *Review of Educational Research, 46,* 61-74.

Tobias, Sigmund. (1989). Another look at research on the adaptation of instruction to student characteristics. *Educational Psychologist, 24,* 213-227.

Toppino, T. C., & Gracen, T. F. (1985). The lag effect and differential organization theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 185-191.

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychological Review, 68,* 229-247.