DOCUMENT RESUME

ED 046 249 TM 000 673

AUTHOR Berliner, David C.
TITLE Aptitude-Treatment Interactions in Two Studies of Learning from Lecture Instruction.
INSTITUTION Far West Lab. for Educational Research and Development, Berkeley, Calif.
PUB DATE Feb 71
NOTE 40p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, N.Y., February 1971)

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Academic Aptitude, Attention, Interaction, *Learning Activities, *Lecture, Memory, Student Behavior, Testing
IDENTIFIERS Johnson Neyman Statistical Technique

ABSTRACT

A study and replication tested whether memory aptitudes interacted with note-taking, paying attention, or responding to test-like events during lecture instruction. Regression analyses were accomplished by means of the single predictor-single criterion case of the modified Johnson-Neyman statistical technique. A disordinal interaction between test-like events and note-taking occurred in both studies. When a subject was low in memory ability, test-like events were recommended over note-taking. It also appeared that at high levels of memory ability note-taking was markedly superior to paying attention but at low levels of memory ability that difference least, and in some analyses paying attention was superior to note-taking. The study and replication demonstrated the robustness of aptitude-treatment interactions, the utility of regression analysis of this type, and some conditions under which test-like events, note-taking, and paying attention are effective student behaviors in lecture instruction. (Author/MF)

ERIC Full Text Provided by ERIC

ED0 46249

APTITUDE-TREATMENT INTERACTIONS IN TWO STUDIES OF LEARNING FROM LECTURE INSTRUCTION

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

DAVID C. BERLINER

FAR WEST LABORATORY FOR EDUCATIONAL RESEARCH AND DEVELOPMENT
1 GARDEN CIRCLE
HOTEL CLAREMONT
BERKELEY, CALIFORNIA 94705

A PAPER PRESENTED AT THE MEETING
OF THE
AMERICAN EDUCATIONAL RESEARCH ASSOCIATION,
NEW YORK CITY, FEBRUARY 1971.

EM 008 673

APTITUDE-TREATMENT INTERACTIONS IN TWO STUDIES OF LEARNING FROM LECTURE INSTRUCTION

David C. Berliner

Far West Laboratory for Educational Research and Development*

Whether "live", on film, or on television, the presentation of information by lecture is usually associated with only two acceptable classes of behavior by students: (1) note-taking, and (2) the catagory of behavior called "paying attention". Other kinds of student behavior are usually considered inappropriate while a lecture is in progress.

A critical analysis of both note-taking and paying attention has been provided by Berliner (1968, 1969).

While note-taking and paying attention may be considered two different, quite prevalent, and not necessarily optimum treatments for subjects attempting to learn from lecture instruction, still a third behavior can be suggested for subjects in a lecture situation. This behavior involves answering questions during a lecture presentation. Rothkopf (1965, 1966), Rothkopf and Bisbicos (1967) and Frase (1967), have consistently shown the facilitative effect of questions inserted into prose material. The questions seem to serve as motivational stimuli, with both arousal and associative outcomes (Frase, 1970). The covert activities

---

* Study 1 was made possible by the help of students and staff at San Jose State College. In addition, Deborah Kearney and Lois Berliner served as assistants on that study. The replication of this study was aided by members of the University of Massachusetts staff. Assistants on that study were Angelica Huber Robertson and Jack Lockhead. Mr. Fred J. Dowaliby provided invaluable assistance in the statistical analyses of these studies.

which the learner is believed to engage in when test-like events occur throughout prose material has been called mathemagenic behavior (behavior which gives birth to learning). Facilitative mathemagenic behaviors may occur in oral instruction, as with prose instruction, when test-like events are used. An extended rational for that point of view, accompanied by empirical data may be found in Berliner (1968), while a less detailed description may be found in Berliner (1969). Some preliminary supporting evidence is also offered by the work of Sanders (1970). Thus three treatments: note-taking (NT), paying attention (PA), and answering questions, i.e., test-like events (TLEs), each with its own advantages and disadvantages, may be thought applicable to promote learning from lecture instruction.

When two or more treatments can be hypothesized to be disparate enough to affect different types of people in different ways, the situation is particularly amenable to analysis of aptitude-treatment interactions (ATIs). Cronbach and Snow (1969), in the most comprehensive review and programmatic study of ATI to date, have stated the problem as follows:

> Assume that a certain set of outcomes from an educational program is desired. Consider any particular instructional treatment. <u>In what manner do the characteristics of learners affect the extent to which they attain the outcomes from each of the treatments that might be considered? Or, considering a particular learner, which treatment is best for him?</u> (1969, p.6).

Aptitude in this approach is defined as any individual difference characteristic that increases or impairs a learner's probability of success in a given instructional treatment. The researcher's task is to discover aptitudes that interact with treatments.

Snow and his colleagues (1965) successfully demonstrated ATI where the personality variables of ascendency and responsibility interacted with live or filmed treatments. Snow and Salomon (1968) pointed out the importance of the ATI approach in media research. That work along with Walberg (1969), Koran (1968), Bunderson (1969), Coffing (1970) and others cited by Cronbach and Snow (1969) and Bracht (1969) present evidence demonstrating that the search for ATIs is difficult, but neither fruitless nor pedestrian. When interactions occur, a decision maker may assign Ss to treatments so as to optimize the learning environment for different groups of students.

However, the ethereal nature of ATIs must be acknowledged. Ripple (1968) failed in numerous replications to find a hint of an ATI. Bracht's (1969) study, containing an excellent overview of the field, likewise failed to find ATIs. Neither Hamilton (1968) nor Alvord (1968), in well designed, well executed, and well analyzed studies, uncovered the ATI effects they sought. Clearly much conceptual work needs to be done after a full range of aptitudes have been explored within different learning tasks, taught in different ways.

This paper describes two studies in learning from lecture instruction wherein the NT, PA, and TLE's treatments described above were hypothesized to interact with certain memory abilities.

Study 1

Method

Subjects. The Ss were predominantly female (70%) and predominantly freshmen (84%) at San Jose State College in California. Most Ss volunteered to participate, though some Ss were required to participate in experiments for grades. After Ss chose particular days and hours, treatments and facilities were randomly assigned to time blocks. Approximately two hours were required of the 211 subjects during the immediate testing. One week later, an additional 30 minutes were required for delayed testing (N=163). Because of the time consuming and rather uninteresting nature of the first experimental session, the 19% mortality rate was not unexpected.

Materials and Procedure. Aptitude information was obtained before instruction began. All Ss took a Short Term Memory Test, a smaller version of the Auditory Letter-span Test available in the kit developed by French, Ekstrom and Price (1963). In addition a test of Memory for Ideas, developed by Seibert, Reid and Snow (1967) was used. In this test Ss were asked to listen to a short prose passage taken from a magazine and then record the story. Two stories, one technical and highly factual, the other about the arts, made up the two parts of the test. Central ideas (key words) were scored one point each. It is a test of more than memory; to some extent the ability to "integrate" a story would seem to be involved. These tests, the Short Term Memory Test of letter-span and the Memory for Ideas Test were the two aptitude variables used in this study.

The lecture material for this study was a 45 minute video-taped overview of Chinese history, comprised of 18 segments, each about 2.5 Minutes in length, and each concerned with a different dynasty or clearly delineated period of Chinese history. The activity associated with the video-taped presentation of the lecture was the independent variable. The three experimental groups had different schedules of questions to be answered in writing during the lecture. Group 1 received a short-answer recall question every 2.5 minutes, after each lecture segment. Group 2 received questions after five minutes of the lecture, one question on each of the preceding two segments. Group 3 received questions after 15 minutes of the lecture, one question on each of the preceding six lecture segments. The S's response booklet provided for knowledge of the correct response for each of the eighteen questions which they received. Two comparison groups were used to examine the effects of the various questioning techniques. Group 4 took notes and Group 5 was instructed to pay attention during the lecture. After lecture instruction had ended for them the comparison groups attempted the same questions as the experimental groups.

Two forms of a criterion test were developed (A and B). All Ss received an immediate criterion test, CT-AI or CT-BI, and one week later took the identical test as a delayed criterion test, CT-AD or CT-BD. The CTs contained 26 short-answer items, varying in point allotment of from one to eight points. Before attempting any of the CTs, Ss were instructed to review the lecture material either mentally (Group 5), through notes (Group 4) or by looking over the questions which occurred during instruction (Groups 1-3). Performance on the CTs was the dependent variable.

Results and Discussion. The development of the A and B forms of the CT, the reliability of the forms and scorers, and the results of analysis of variance for these data has been presented elsewhere (Berliner, 1968; 1969). A summary of the important data and analyses is provided in Table 1. However, this re-examination of the data was undertaken to investigate the possibilities of aptitude-treatment interactions, for which analysis of variance is not the most appropriate statistical procedure. Following the recommendations of Cronbach and Snow (1969), regression analyses were used to analyze the interactions between aptitudes and treatments. The logic of the regression analyses described below is as follows: First a test of homogeneity of variance among the treatment groups was applied, followed by a test of common slope, or parallelism of regression. The hypothesis of common slope was rejected if F exceeded the .05 level of significance. Following a significant F test, the Johnson-Neyman (1936) technique was applied. In these analyses only the single predictor-single criterion case of the Johnson-Neyman technique was used. Exploratory analysis using multiple regression demonstrated no great increases in variance accounted for when variables were added to the multiple regression equations. Thus the slight increase in precision to be gained seemed not to justify the more complex analysis requiring the interpretation of ellipses or parabolas.

Computing formulas for the Johnson-Neyman technique were obtained from

Walker and Lev (1953), with modifications proposed by Potthoff (1964) to determine a simultaneous region of significance. The unmodified Johnson-Neyman technique is quite liberal in defining a region of significance, while the Potthoff modification is more conservative. In this work, in one analysis, the region described by the Johnson-Neyman technique at α=.15 was equivalent to the region described after Potthoff's modification of α=.05. This coincides with a comparison of the two techniques made by Cahen and Linn (1970). Further, the more conservative technique seemed called for because it allows inferences to be made about all points within a region, rather than for single points, and because multiple comparisons (TLEs vs NT, TLEs vs PA, NT vs PA) were used, thus slightly decreasing the level of confidence involved in each comparison. The kind of decision making which occurs in this analysis is different from the traditional inferential model. Thus Potthoff (1964) and others recommend using an 80% or 90% level of confidence. All analyses reported below were made at the 90% level of confidence. The computer program (Dowaliby and Berliner, 1970) provided tests of all assumptions as well as regression equations and plotting points for any significant interactions between aptitudes and treatments. All regression equations computed for Study 1 are reported in Table 2. Figures are provided for significant interactions.

Insert Tables 1 and 2 about here

Figure 1 displays the regression of CT-8i on the Short Term Memory Test for the NT, PA, and TLEs groups. For the NT and TLEs regression lines the assumption of homogenity of variance was met, and the hypothesis of common slope among the two treatment groups was rejected; F=7.09, df=1/36, p=.009. Application of the modified Johnson-Neyman technique (α=.10) to the disordinal interaction occurring between these two regression lines described a region of non-significance between 6.86 and 20.11 on the X axis, around a point

TABLE 1

Descriptive Statistics and Summary Analysis of the CTs Used in Study 1

		Treatments					Results of an F test from a one-way analysis of variance	Results of a Newman-Keuls Post-Hoc Comparison
		Group 1 TLEs	Group 2 TLEs	Group 3 TLEs	Group 4 NT	Group 5 PA		
CT-AI and CT-BI Combined	M	23.11	18.86	21.96	18.29	13.87	5.44**	1>5* 2>5* 3>5* 4>5*
	SD	8.54	9.63	11.87	9.69	8.03		
	N	47	43	49	34	38		
CT-AD and CT-BD Combined	M	19.98	15.91	17.82	18.14	13.96	3.05*	1>5*
	SD	7.21	8.07	9.56	10.34	7.43		
	N	42	33	33	28	27		
CT-AI	M	26.33	21.95	25.21	22.24	14.26	4.49**	1>5* 2>5* 3>5* 4>5*
	SD	8.67	10.65	11.46	9.89	9.36		
	N	24	21	24	17	19		
CT-AD	M	22.67	18.67	20.67	23.29	14.65	2.62*	1>5* 4>5*
	SD	6.67	7.53	10.64	10.39	8.37		
	N	21	15	18	14	17		
CT-BI	M	19.74	15.91	18.84	14.35	13.47	2.15	N/A
	SD	7.13	7.67	11.63	7.94	6.69		
	N	23	22	25	17	19		
CT-BD	M	17.29	13.61	14.40	13.00	12.80	1.20	N/A
	SD	6.85	7.98	6.98	6.84	5.71		
	N	21	18	15	14	10		

*P<.05
**P<.01
N/A = Not Applicable

TABLE 2

Regression Equations of the Form Y=a+bx
For Each Treatment in Study 1

Regressed Variables	Regression Equations NT a bx	TLE a bx	PA a bx
CT-AI on Short Term Memory	16.15+ .81	13.53+1.76	6.38+1.21
CT-AD on Short Term Memory	13.53+1.24	14.94+1.05	7.01+1.24
CT-AI on Memory for Ideas, Part I	-1.42+1.18	12.70+ .61	3.19+ .53
CT-AD on Memory for Ideas, Part I	1.81+1.11	12.45+ .46	-8.16+1.17
CT-AI on Memory for Ideas, Part 2	6.61+1.03	12.73+ .88	26.59- .83
CT-AD on Memory for Ideas, Part 2	6.73+1.09	14.56+ .53	27.45- .89
CT-AI on Memory for Ideas, Full Scale	-6.45+ .82	8.09+ .49	5.93+ .23
CT-AD on Memory for Ideas, Full Scale	-3.31+ .77	10.22+ .33	-12.07+ .79
Ct-BI on Short Term Memory	3.61+1.60	23.65- .60	11.61+ .26
CT-BD on Short Term Memory	3.98+1.30	19.58- .35	7.12+ .79
CT-BI on Memory for Ideas, Part 1	-8.44+1.12	20.49- .04	6.98+ .32
CT-BD on Memory for Ideas, Part 1	-6.15+ .92	22.84- .26	19.69- .34
CT-BI on Memory for Ideas, Part 2	5.11+ .62	13.13+ .39	8.36+ .33
CT-BD on Memory for Ideas, Part 2	4.47+ .57	11.66+ .32	6.15+ .40
CT-BI on Memory for Ideas, Full Scale	-11.99+ .75	16.32+ .09	6.03+ .21
CT-BD on Memory for Ideas, Full Scale	-9.31+ .63	18.66- .03	14.91- .06
Combined A and B CT-I on Short Term Memory	7.47+1.52	20.90+ .32	10.24+ .53
Combined A and B CT-D on Short Term Memory	6.26+1.59	18.40+ .23	8.72+ .80
Combined A and B CT-I on Memory for Ideas, Part 1	-4.74+1.14	16.84+ .29	4.69+ .44
Combined A and B CT-D on Memory for Ideas, Part 1	-.22+ .91	17.75+ .10	4.85+ .46
Combined A and B CT-I on Memory for Ideas, Part 2	5.24+ .87	15.92+ .44	13.20+ .04
Combined A and B CT-D on Memory for Ideas, Part 2	4.67+ .88	15.89+ .25	17.47- .23
Combined A and B CT-I on Memory for Ideas, Full	-9.46+ .79	13.88+ .24	6.12+ .22
C[illegible] A and B CT-D on Memory for Ideas, Full	-5.29+ .66	15.55+ .11	6.62+ .21

ERIC Full Text Provided by ERIC

of non-significance (the cross-over point) occurring at 9.10. The region of non-significance contained 42.5% of the Ss while 23 Ss or 57.5% of the Ss fell below the region of non-significance. With 90% confidence it can be stated that those Ss below 6.86 on the Short Term Memory Test should be assigned to the TLEs treatment. With Ss of relatively poor short term memory the use of TLEs seems to act as a memory aid. NT appears not to be an effective treatment at these levels of measured memory ability. However, from 50% confidence at 9.10 to some higher confidence level (though below 90%) it would appear that NT is a better treatment than TLEs at the higher levels of memory ability. Though "common sense" might recommend taking notes if one's memory is limited, a psychological interpretation based on these data could contradict this, suggesting that proficiency in note-taking requires continuously holding new information in memory while writing notes on material previously communicated. Note-taking is apparently quite useful for Ss skilled enough to handle the memory requirements. It appears to by dysfunctional for Ss low in short term memory ability. For these Ss, TLEs may provide the aid to memory that they require. Though not significant, ($F=2.56$, $df=1/32$, $p=.12$) the hint of a disordinal interaction in Figure 1 between PA and NT indicates that even PA might be a better treatment than NT at the very lowest levels of the short term memory continuum used in this study. Furthermore, if the range of scores in the PA group were greater, PA might ultimately be a better technique than TLE for high ability Ss. These particular relationships become clearer later in this discussion.

Insert Figure 1 about here

Figure 2 displays essentially the same relationships between short term memory and the B form of the delayed criterion test, CT-BD,

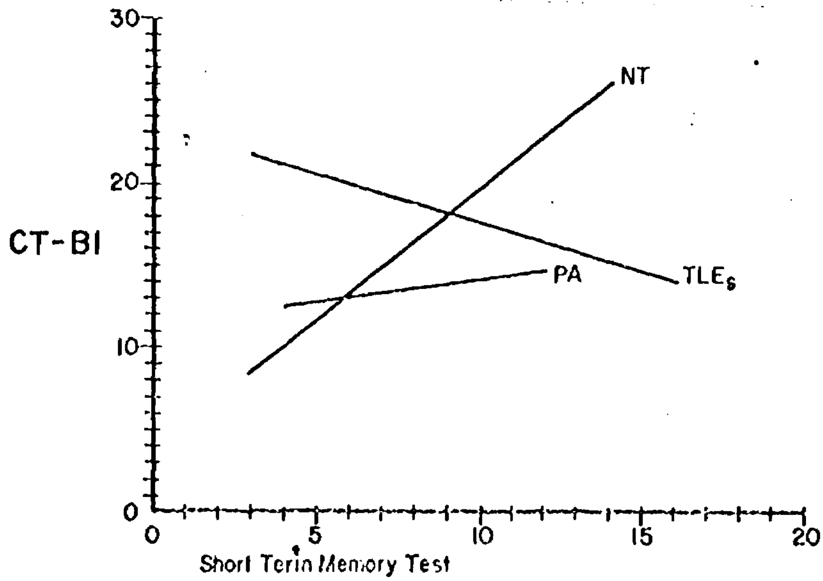

Fig. 1. Regression of CT-BI on Short Term Memory Test (Study 1).

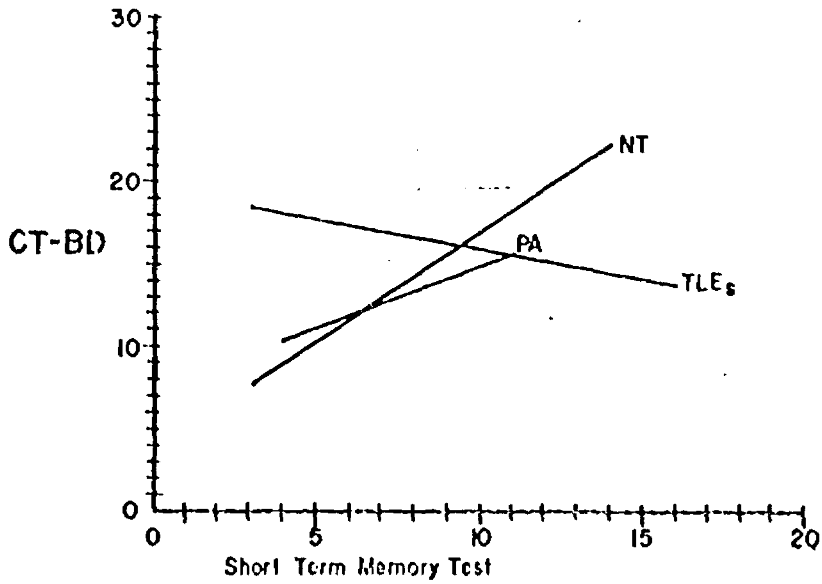

Fig. 2. Regression of CT-BD on Short Term Memory Test (Study 1).

administered 1 week after instruction. The hypothesis of common slope among the NT and TLE treatments was rejected, F = 4.84, df = 1/32, p = .03, and the assumption of homogenity of variance was met. However, application of the modified Johnson-Neyman technique (α = .10) for the disordinal interaction indicated by the regression lines of the TLEs and NT groups failed to define a region of significance. The stability of the slopes over one weeks time was reassuring, as was the continued rejection of parallelism. With confidence ranging from 50% at the point of cross-over, to something less than 90% confidence at the ends of the regression slopes it could again be noted that when short term memory ability is low, TLEs seem to act as a memory aid. When memory ability is quite high, NT seems to be a more effective treatment. PA appears to be relatively ineffective, but extension of this regression line, assuming that the slope would remain the same if Ss existed at the high and low ends of the scale, might indicate that PA was a more effective treatment than NT when memory aptitude was low and that PA was a more effective technique than TLEs when memory aptitude was quite high.

Insert Figure 2 about here

In Figure 3 are the regression lines of CT-BI on the Memory for Ideas Test, Part 1. Two disordinal interactions appear to occur among the three regression lines. The hypothesis of common slope between the NT and TLEs treatments was rejected; F = 9.09, df = 1/36, p = .005. The assumption of homogenity of variance was met. Application of the modified Johnson-Neyman procedure (α = .10) to the disordinal interaction which occurs between these two regression lines yields a point of non-significance at 24.97, with a region of non-significance between

20.87 and 39.55. Twenty-one Ss, 52.5% of the sample fell below the region of non-significance, indicating that with 90% confidence Ss below 20.87 on the X axis will benefit more from the TLEs treatment than from the NT treatment. These data, using a related but quite different aptitude measure, substantiate the data presented in Figures 1 and 2. Again the usefulness of TLEs is demonstrated when memory aptitude is low, in this case the more cognitive memory ability measured in the Memory for Ideas Test. Once again NT appears to be the more effective treatment for Ss with high memory ability, though the confidence in that statement is considerably less than 90%.

Insert Figure 3 about here

Of particular interest in Figure 3 is the apparent disordinal interaction between regression slopes for the PA and NT treatments. The hint of this relationship existed in Figures 1 and 2 when short term memory was the aptitude variable. Now, in Figure 3, this interaction is clearer when the Memory for Ideas Test, Part 1, is the aptitude variable. However, the hypothesis of parallelism was not rejected; $F = 3.20$, $df = 1/32$, $p = .08$. The modified Johnson-Neyman technique applied to this situation failed to yield a region of significance, all cases falling within a region of non-significance. This was not the case when the delayed CT was analyzed, as Figure 4 makes clear.

Figure 4 displays the regression of CT-BD on the Memory for Ideas Test, Part 1, for the NT, PA, and TLEs treatments. Once again two disordinal interactions appear to occur among the three regression lines. The hypothesis of common slope between the NT and TLEs treatment was rejected; $F = 10.09$, $df = 1/32$, $p = .003$. The assumption of homogeneity of variance was met and the modified Johnson-Neyman technique

ERIC

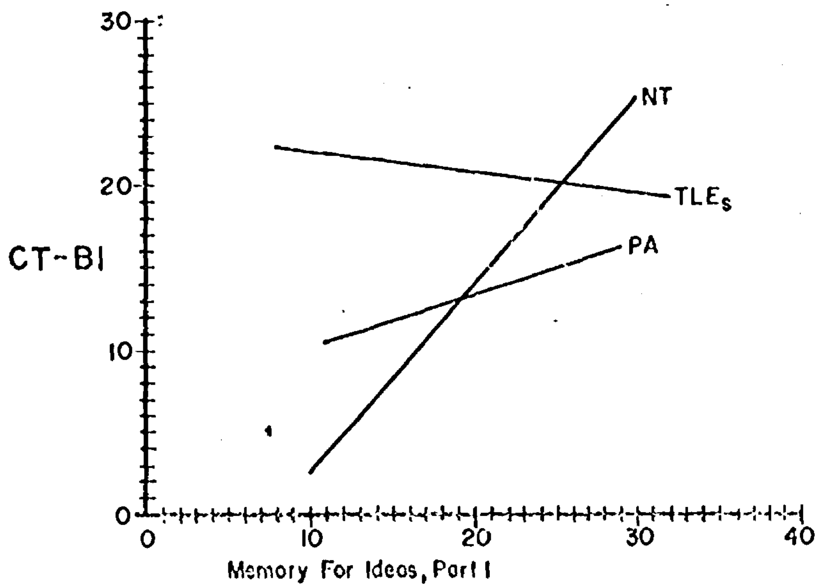

Fig. 3. Regression of CT-BI on Memory for Ideas, Part 1 (Study 1).

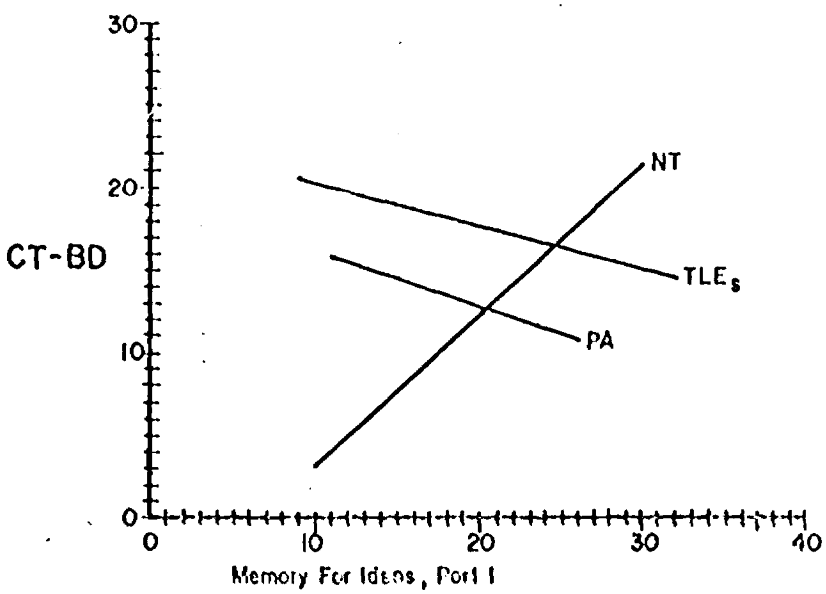

Fig. 4. Regression of CT-BD on Memory for Ideas, Part 1 (Study 1).

(α= .10) applied to this interaction described a region of non-significance between 20.79 and 35.30, with the point of non-significance occurring at 24.65 on the X axis. Seventeen Ss or 47.2% of the sample fell below the region of non-significance. For these Ss we can predict with 90% confidence that a treatment incorporating TLEs would be more beneficial than a treatment calling for note-taking while learning from lecture instruction. No significant differences between these two treatments is evident above 20.79, but once again the previously described trends are substantiated and with somewhat greater confidence than chance it can be predicted that NT is more effective than TLEs when memory aptitude is high.

Insert Figure 4 about here

The disordinal interaction occurring between the PA and NT treatments is even more interesting from a methodological point of view. The modified Johnson-Neyman procedure (α= .10) defines a region of non-significance between 14.04 and 27.30 around a point of non-significance at 20.48 on the X axis. Eighty-four percent of the Ss in this sample fell within this region. However, 8% of this sample fell below this region and 8% fell above this region of non-significance. Thus we may conclude with 90% confidence that for Ss with a memory ability score below 14.04 PA is the more beneficial treatment and again with 90% confidence we may conclude that for Ss with a memory score above 27.30 NT is the more beneficial treatment. The trends noted in Figures 1-3 with regard to the relationship of PA and NT are accentuated here because in this analysis the slope of the PA regression line has become negative. Within the region of non-significance, from 14.04 to

27.30, one can assign Ss to one of the two treatments as one moves in towards the point of non-significance (20.48). This is done with decreasing confidence until at that point, the two treatments yield equal predictions and only chance or 50% confidence can be put into any decision about assignment to these two treatments.

The results and discussion of Figures 1 through 4, concerned with the regression of the B form of immediate and delayed criterion tests on the Short Term Memory Test and the Memory for Ideas Test indicate that the following generalizations are appropriate: 1) When memory ability is high, NT is an effective learning strategy. For Ss with low memory scores it is not recommended; perhaps it is even dysfunctional. 2) When memory ability is low, the presence of TLEs aid learning from lecture instruction. As memory scores increase, the usefulness of this technique is vitiated, perhaps even becoming dysfunctional. 3) In comparing the two "standard" forms of student behavior, NT and PA, it appears that PA is a more effective learning strategy than NT when memory ability is low. The reverse is true if memory aptitude is high, where NT appears to be markedly superior to PA; 4) TLEs appear to be a more useful strategy than PA at almost all levels of memory ability, though at the very highest levels the difference is least. PA may even be superior to TLEs at the very highest levels of memory ability if extrapolation of the PA regression line is possible.

Using the Memory for Ideas Test, Part 2, which is less technical and provides a more general story for Ss to remember, we found no interactions with treatments when the B form of the CT was used.

The full scale Memory for Ideas Test was analyzed by regressing CT-BI and CT-BD upon it. The full scale score was the sum of parts 1 and 2 of the test. Figures 5 and 6 display these data.

A test of common slope for the regression lines of the NT and TLEs treatments presented in Figure 5 was rejected; F=5.76, df=1/36, p=.02. The assumption of homogeneity of variance was met. A region of non-significance was described extending from 35.54 to 132.24 on the X axis around a point of non-significance at 42.94. Nineteen Ss or 47.5% of the cases in this sample were below the region of non-significance. The PA and NT interaction did not reach significance; F=3.33, df=1/32, p=.07.

Insert Figure 5 about here

The analysis of the delayed CT presented in Figure 6 was similar. A test of common slope of the regression lines for the NT and TLEs treatments was rejected; F=5.91, df=1/32, p=.02. Homogeneity of variance existed. One third of this sample, 12 Ss, fell below the region of non-significance which extended from 34.86 to 104.75 around a point of non-significance at 42.36 on the X axis. The PA and NT interaction did not reach significance; F=3.90, df=1/21, p=.06.

Insert Figure 6 about here

The results and generalizations made with regard to Figures 1-4 are congruent with the data reported in Figures 5 and 6 using the regression of CT-BI and CT-BD on the full scale Memory for Ideas Test. With 90% confidence it is predicted that Ss whose memory for ideas scores fall below 35 points would benefit more from treatments incorporating TLEs than from taking notes. However, the point at which NT significantly exceeds TLEs is far beyone the range of these data. The relationship of PA to NT described above appears to hold, but in these analyses was non-significant.

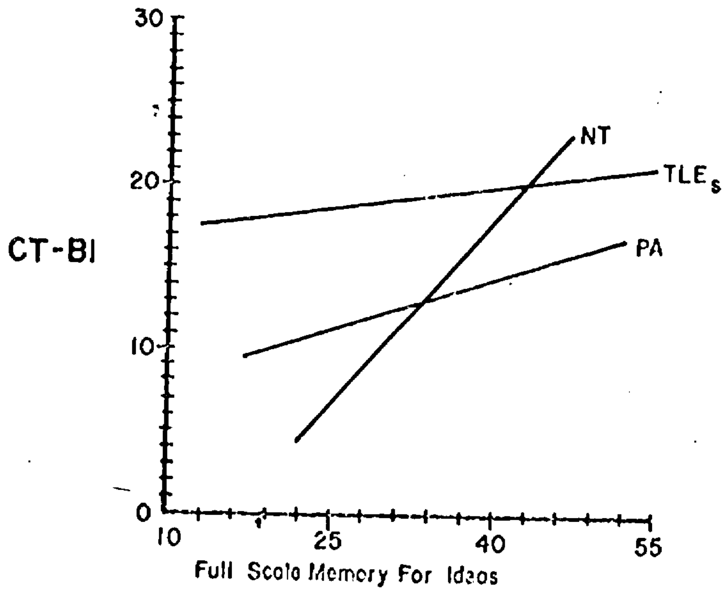

Fig. 5. Regression of CT-BI on full scale Memory for Ideas (Study 1).

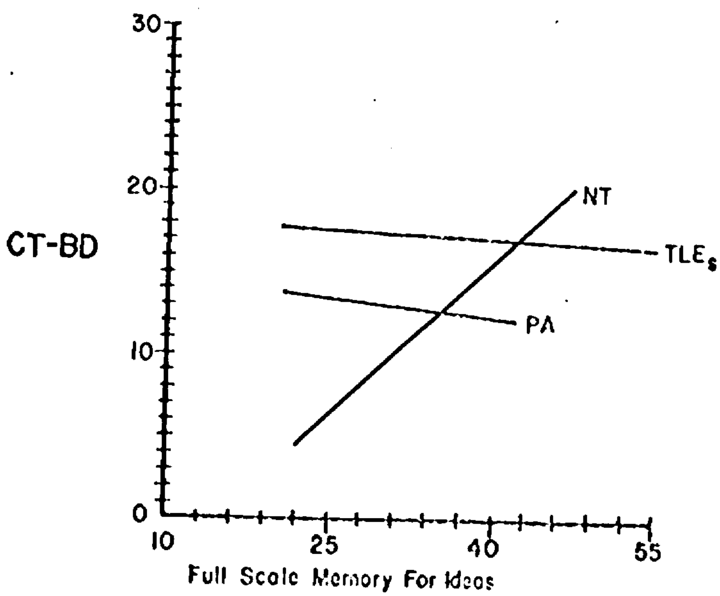

Fig. 6. Regression of CT-BD on full scale Memory for Ideas (Study 1).

All analyses of Form A of the CT provided non-significant results. The regression equations in Table 2 provide all the necessary data for graphing these findings. Essentially six of the eight analyses using Form A of the CT yielded non-significant interactions between the NT and TLEs treatments. However, the regression lines did cross and showed the TLEs group to have a higher predicted CT score at low memory ability levels and the NT group to have a higher predicted CT score at high levels of memory ability. Thus, these analyses are supportive of the analyses of the B form data. No interpretation of the Form A data with regard to interactions involving the PA group is warranted, since the regression slope of the PA group was quite variable.

As noted above the A and B forms of the CT are randomly parallel tests following modern test theory procedures outlined in Magnusson (1967). The tests do not meet the classical criteria for parallel tests. If classical notions of parallel forms may be ignored, the pooled data from these two forms is appropriate to use. These analyses were, therefore, also performed.

The trends noted in Figures 1 and 2 using only the Form B data held for the regression of combined A and B form data on short term memory in both immediate and delayed tests. In this analysis, however, the results were non-significant. The regression of the combined CT-I data on Memory for Ideas, Part 1, is displayed in Figure 7, while the regression of CT-D combined data on Memory for Ideas, Part 1, is displayed in Figure 8. These two figures show similar findings. The hypothesis of common slope between the NT and TLEs regression lines in Figure 7 was rejected; $F=7.70$, $df=1/77$, $p=.007$. In Figure 8 the hypothesis of common slope for these two treatments was also rejected; $F=6.36$, $df=1/67$, $p=.01$. The assumptions of homogeneity of variance were met in both cases. In Figure 7 a region of non-significance was described by the modified Johnson-Neyman technique extending from 20.67 to 45.29 around a point of non-significance at 25.33 on the X axis. With 90% confidence it can be stated that the 43 Ss, or 53.1% of this sample who fell below 25.33 on the aptitude test would benefit more from a treatment incorporating TLEs than one requiring NT. In Figure 8 a region of non-significance was described extending from 15.02 to 38.80 around a point of non-significance at 22.32 on the X axis. In this case 13 Ss, or 18.3% of the sample fell below the region of non-significance, indicating that with 90% confidence Ss scoring below 15.02 points on this memory aptitude test would benefit more from a treatment incorporating NT than one utilizing TLEs. This relationship is stable over the one week delay between tests as may be noted in the figures. The relationship between PA and NT described in Figures 7 and 8 is quite weak, but compatible with previous discussions.

Insert Figure 7 about here

Insert Figure 8 about here

The regression equations used in analysis of the Memory for Ideas Test, Part 2, are presented in Table 2. Though all analyses were non-significant, no reversal in the trends already noted occurred when graphic displays of the data were analyzed.

The full scale Memory for Ideas Test was analyzed by regressing immediate and delayed combined CTs upon that memory test and the results are presented in Figures 9 and 10. For the immediate CT, Figure 9, two disordinal interactions which occurred among the three treatments were significant. The hypothesis of common slope was rejected between the NT and TLEs treatments; $F=6.38$, $df=1/77$, $p=.01$. The modified

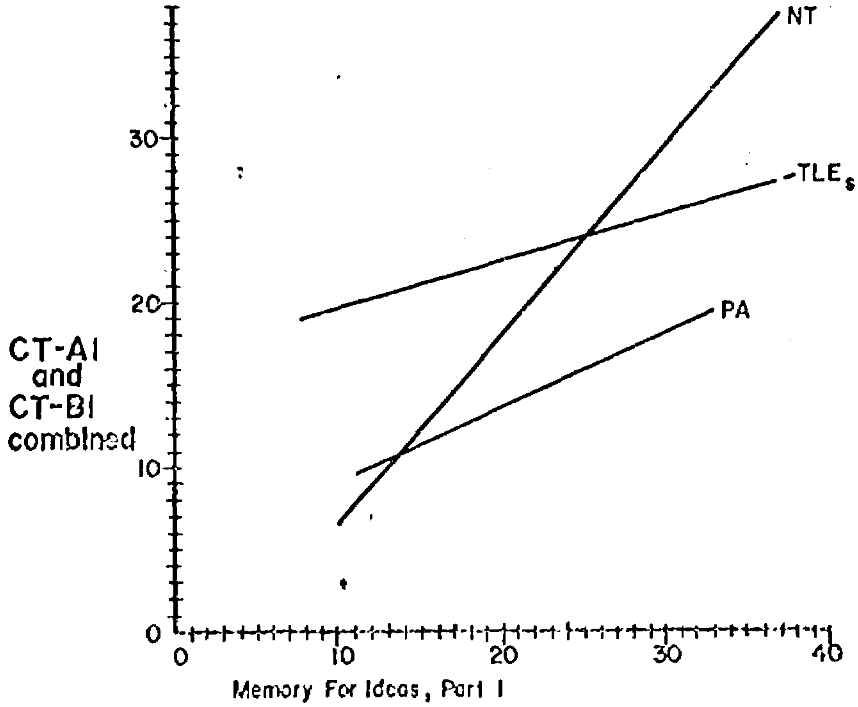

Fig. 7. Regression of CT-AI and CT-BI, combined data, on Memory for Ideas, Part 1 (Study 1).

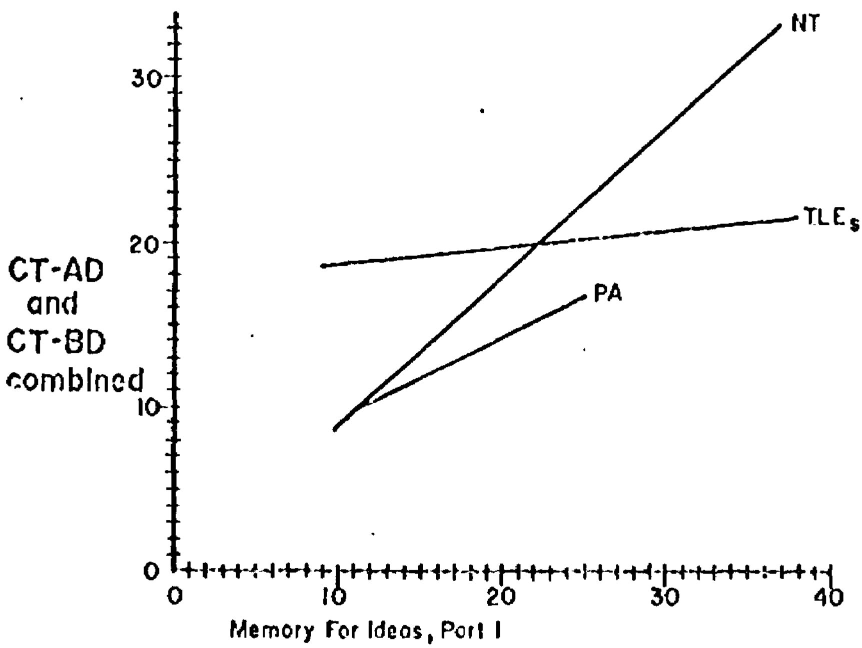

Fig. 8. Regression of CT-AD and CT-BD, combined data, on Memory for Ideas, Part 1 (Study 1).

Johnson-Neyman technique described a region of non-significance from 35.46 to 91.39 around a point of non-significance at 42.81. The assumption of homogeneity of variance was tested and accepted. For the 39 Ss, or 48.1% of the cases which fell below the region of non-significance, it can be stated with 90% confidence that they would benefit more from a treatment incorporating TLEs rather than a treatment requiring NT.

If the comparison of treatments is between NT and PA, it is hypothesized that, with 90% confidence, Ss would benefit more from NT than PA if their memory aptitude test score is above 34.36. For this comparison the hypothesis of common slope for the two regression lines was rejected; F=5.19, df=1/68, p=.02. The modified Johnson-Neyman technique described a region of non-significance between -184.84 and 34.36 around a point of non-significance at 27.21 on the X axis. Thirty-five Ss, or 48.6% of this sample were above the region of non-significance.

Insert Figure 9 about here

Once again it appears that TLEs is a viable technique for Ss low in this memory aptitude, while NT is a viable technique for Ss high in this memory aptitude. These relationships are not as strong when the delayed CT was examined, as illustrated in Figure 10. In that case, the interaction between NT and TLEs was significant, with a test of parallelism between the two regression slopes providing an F ratio of 5.94, df=1/67, p=.02. The region of non-significance was quite extensive, from 25.74 to 64.58 around a point of non-significance of 38.03 on the X axis. Nevertheless, seven Ss, 9.9% of the cases, did fall below the region of non-significance. Thus a weaker, but essentially similar relationship exists as that described for the data from the immediate test presented in Figure 9. The interaction between NT and PA was not significant, but as illustrated in Figure 10 it is compatible with the previous discussion.

Insert Figure 10 about here

The summary of Study 1 will be postponed until after the results of Study 2 are presented and discussed.

Study 2

The findings of Study 1, if replicable, would be important theoretically and could have practical implications for the design and conduct of lecture instruction. With some modifications, a replication of Study 1 was thus undertaken.

Method

Subjects. A total of 211 Ss were drawn from among those students enrolled in basic psychology at the University of Massachusetts. Mean age was 18.6 years with a range of 17-28 years. Most Ss were sophomore liberal arts majors, though all classes and majors were represented. Forty-three percent were male and 57% were female. A personal history form was completed by the Ss. Analysis of those data revealed that there were no differences among Ss in the various treatments for the variables: year in school; academic major; sex; age; and previous history of exposure to the subject matter used in the experimental lecture. The Ss signed up for two sessions which were given on the same weekday, on successive weeks during the fall of 1969. Conditions were assigned randomly to time blocks.

Materials and Procedure. In addition to the two aptitude tests described in Study 1, a third aptitude test, Memory for Sentences, was added. This aural test presented Ss with complete sentences of various length, and required verbatim recall immediately after each sentence.

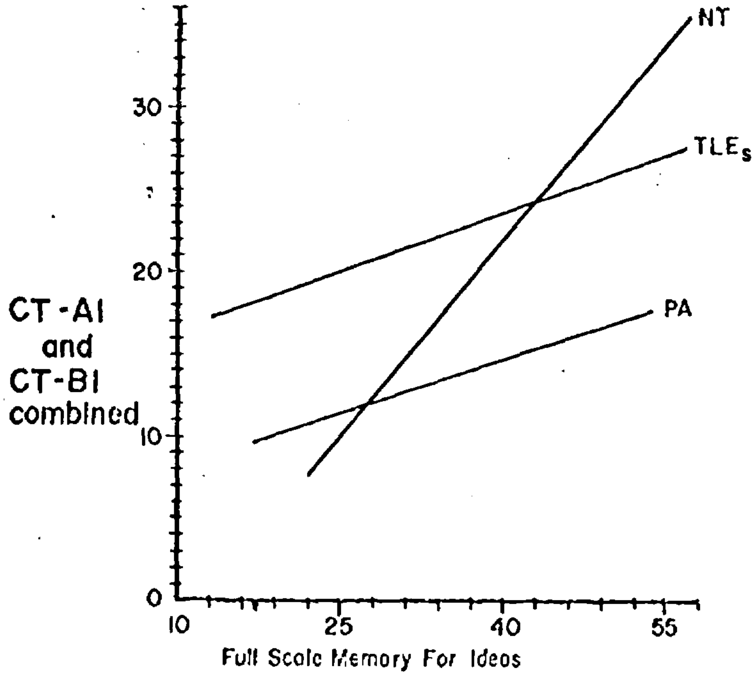

Fig. 9. Regression of CT-AI and CT-BI, Combined data, on full scale Memory for Ideas (Study 1).

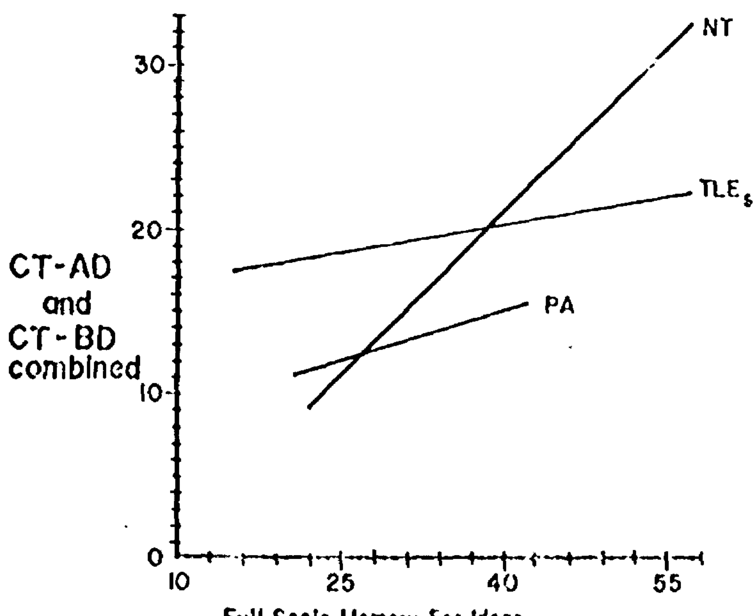

Fig.10. Regression of CT-AD and CT-BD, Combined data, on full scale Memory for Ideas (Study 1).

This test, like the Memory for Ideas Test, was developed by the Purdue research group (Siebert, Reid and Snow, 1967). This test was added to represent short term memory ability for meaningfully connected verbal material, less complex than that represented by the Memory for Ideas Test. Unfortunately the battery of aptitude tests had to be given after instruction because of time limitations on the use of the Ss. All tests and instructions were administered by audio tape.

The lecture material was a 30-minute overview of Chinese history, in which the 18 segments used in Study 1 were reduced to 12 two-and-one-half minute segments. The lecture was presented via videotape through large television monitors in the experimental room. Only three experimental groups were used: TLEs, NT and PA. The B form of the test described in Study 1 was used as a CT; it was administered as CT-I (immediate) or CT-D (delayed). Procedures for processing the three groups were essentially the same as in Study 1.

Results and Discussion. Pearson correlations between scorers of the CTs and aptitude tests average about .90 in separate checks of reliability. Test-retest reliability was also high. For the three treatments the correlations between immediate and delayed tests were .90, .92 and .84. Table 3 presents means, standard deviations and the numbers of Ss per treatment. The results of an F test from an analysis of variance and associated significance levels are also presented for these data. Post hoc analyses using the Newman-Keuls procedure are presented (column N-K).

Insert Table 3 about here

In analysis of both CT-I and CT-D a significant between treatment difference was found. Overall the PA group, in both immediate and delayed testing, did poorest again. The TLEs group which is

TABLE 3

Descriptive Statistics and Summary Analysis of the CTs Used in Study 2

Study 2

		Treatment				
		TLEs	NT	PA	F test	N-K
	M	24.60	24.53	19.16	7.79**	TLEs>PA*
CT-I	SD	9.59	8.12	10.87		NT>PA*
	N	67	58	86		
	M	21.87	24.53	17.48	11.95**	TLEs>PA*
CT-D	SD	8.34	8.15	9.44		NT>PA*
	N	67	58	86		

*P .05
**P .01

ERIC Full Text Provided by ERIC

as powerful as the NT group in immediate testing appeared less effective on the delayed test. As indicated in column N-K of Table 3, in both immediate and delayed testing the TLEs and NT groups did not differ statistically among themselves but were significantly different from the PA treatment. These findings are quite simiiar to the findings reported in Study 1, The NT group's remarkable retention on the CT after one week's time deserves special mention. An earlier investigation (Berliner, 1968) reported a reminiscence effect with virtually identical material, and once again this material by treatment interaction seems to yield a reminiscence effect. Replication of conditions producing reminiscence are themselves worth further investigation.

Since aptitude tests were administered one week after instruction and immediate testing, aptitude scores were examined for possible treatment effects. No significant differences between treatments were found using simple one way analyses of variance. As noted in Table 4, means and standard deviations for the tests in each treatment are quite homogenous.

Insert Table 4 about here

Regression analyses were performed using the procedures noted above. All regression equations for all analyses of Study 2 are presented in Table 5. Figures 11 and 12 illustrate the important results. Essentially the same relationships as those in Figures 1 and 2 are displayed. The stability of the ATIs described in Study 1 is quite distinct. For the regression of CT-1 on the Short Term Memory Test (Figure 11) the hypothesis of common slope is rejected with regard to the NT and TLEs interaction ($F=5.17$, $df=1/121$, $p=.02$). The assumption of homogeneity of variance was met and application of the modified

TABLE 4

Descriptive statistics for the Aptitude Tests Used in Study 2

		Treatments in Study 2		
		TLEs	NT	PA
Short Term Memory Test	M	6.81	6.83	7.29
	SD	2.74	2.23	2.73
	Range	3-15	2-13	2-18
Memory for Ideas, Part 1	M	19.52	19.66	20.42
	SD	6.76	6.61	7.17
	Range	6-35	6-38	0-38
Memory for Ideas, Part 2	M	13.67	12.27	13.76
	SD	4.18	4.55	4.85
	Range	4-22	0-22	3-27
Memory for Ideas, Full Scale	M	33.19	31.93	34.17
	SD	9.22	9.43	9.83
	Range	17-54	13-56	9-59
Memory for Sentences	M	2.04	1.84	1.88
	SD	1.34	1.25	1.34
	Range	0-6	0-5	0-6

Johnson-Neyman technique (α=.10) described a region of non-significance between -.38 and 14.85 around a point of non-significance of 6.86. Only one case, .8% of the sample fell above the region of non-significance. Nevertheless, the similarity between the regression analyses for Study 1 and Study 2 is clear. With confidence varying between 50% and 90% it may be repeated, with increased assurance, that when short term memory ability is low, treatments incorporating TLEs are recommended. Conversely, when this memory ability is high, NT is recommended.

Insert Table 5 about here

Insert Figure 11 about here

Insert Figure 12 about here

Regression data for the delayed CT on the Short Term Memory Test (Figure 12) once again supports previous findings. In this case the hypothesis of common slope between the NT and TLEs regression lines was (marginally) rejected; F=3.75, df=1/121, p=.052. The conservatism of the modified Johnson-Neyman technique at α=.10 worked against defining any regions of significance. Figures 11 and 12 show that the PA and TLEs relationships noted earlier were replicated, though again were weak. The PA and NT interactions reported earlier were not as distinct in this study.

The analysis of the Memory for Ideas Test, by parts, and also using the full scale score, uncovered no interactions with treatments. The regression equations for these analyses are found in Table 5. Graphing these regression lines reveals neither support for nor marked contradictio[n]

TABLE 5

Regression Equations of the Form Y=a+bx
For each Treatment in Study 2

Study 2

Regressed Variables	Regression Equations		
	NT	TLE	PA
	a bx	a bx	a bx
CT-I on Short Term Memory	19.91+ .68	30.14- .81	19.85- .09
CT-D on Short Term Memory	21.45+ .45	26.81- .73	16.39+ .15
CT-I on Memory for Ideas, Part 1	13.60+ .56	9.31+ .78	4.93+ .70
CT-D on Memory for Ideas, Part 1	13.73+ .55	10.63+ .58	4.21+ .65
CT-I on Memory for Ideas, Part 2	17.31+ .59	16.37+ .60	14.13+ .37
CT-D on Memory for Ideas, Part 2	16.98+ .62	14.49+ .54	13.18+ .31
CT-I on Memory for Ideas, Full Scale	11.46+ .41	6.53+ .54	3.46+ .46
CT-D on Memory for Ideas, Full Scale	11.37+ .41	7.92+ .42	3.08+ .42
CT-I on Memory for Sentences	20.30+2.30	22.31+1.12	17.44+ .91
CT-D on Memory for Sentences	21.20+1.81	20.64+ .60	15.26+1.18

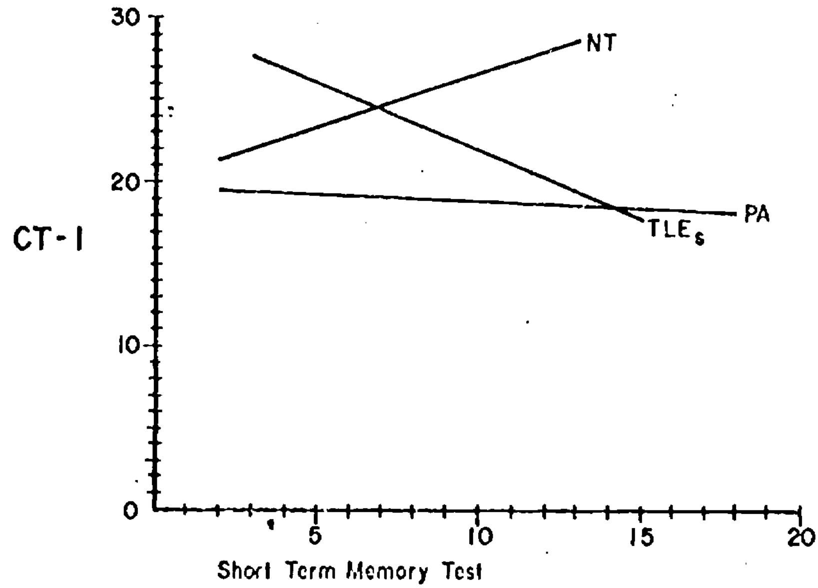

Fig.11. Regression of CT-I on Short Term Memory Test (Study 2).

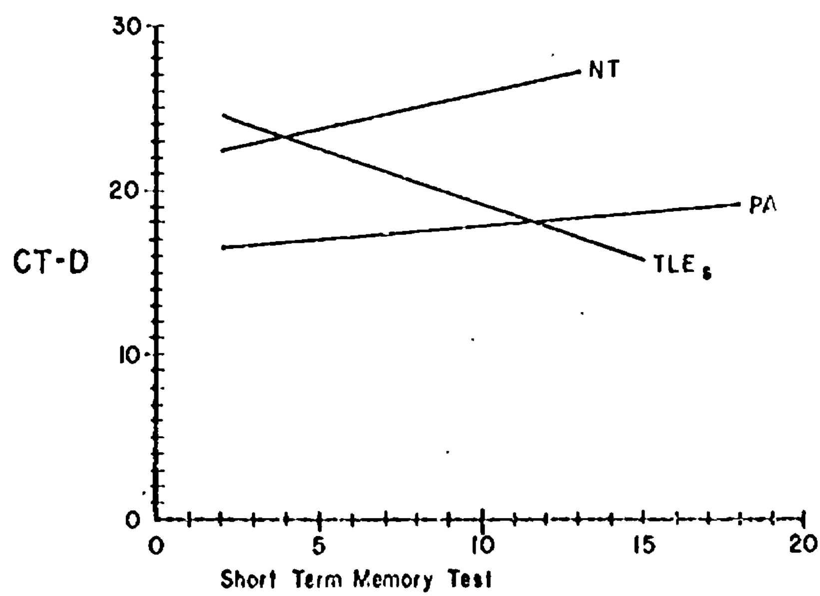

Fig.12. Regression of CT-D on Short Term Memory Test (Study 2)

of the findings in Study 1.

The additional aptitude test used in this study, Memory for Sentences, was also analyzed by procedures previously described. The hypothesis of parallelism between treatment regression lines was not rejected, thus no regions of significance were found when the Johnson-Neyman technique was used. The data, while not confirming previous memory aptitude-treatment interactions is nevertheless supportive. Figures 13 and 14 display the obtained relationships. In the case of the regression of CT-I on Memory for Sentences a non-significant disordinal interaction between TLEs and NT treatments occurs. This suggests the same kind of decision rules developed in Study 1 about when to recommend TLEs and NT treatments for known scores on memory aptitude tests. In the case of the regression of CT-D on Memory for Sentences a non-significant ordinal interaction occurs, suggesting that at higher scores on memory aptitude tests NT is a superior treatment to TLEs, a finding which is also in line with the results of Study 1.

Insert Figure 13 about here

Insert Figure 14 about here

In this study the PA treatment was so weak that virtually any possibility of disordinal interactions occurring with other treatments was remote. However, as noted in Figures 11 and 13, at the lower levels of memory abilility the difference in predicted CT score for the NT and PA treatments is considerably less than when memory ability is high. This trend was noted in Study 1. Though this relationship was not obvious in the delayed testing (Figures 12 and 14), the results do not contradict interpretations which were offered in Study 1.

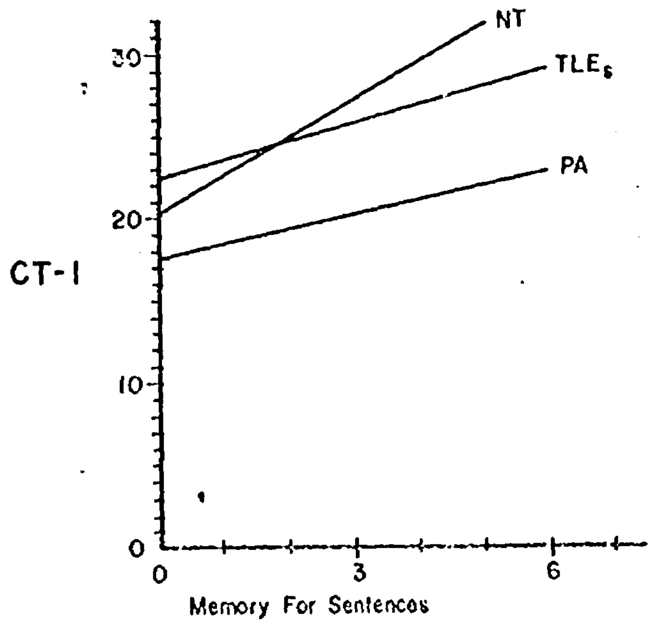

Fig. 13. Regression of CT-I on Memory for Sentences (Study 2).

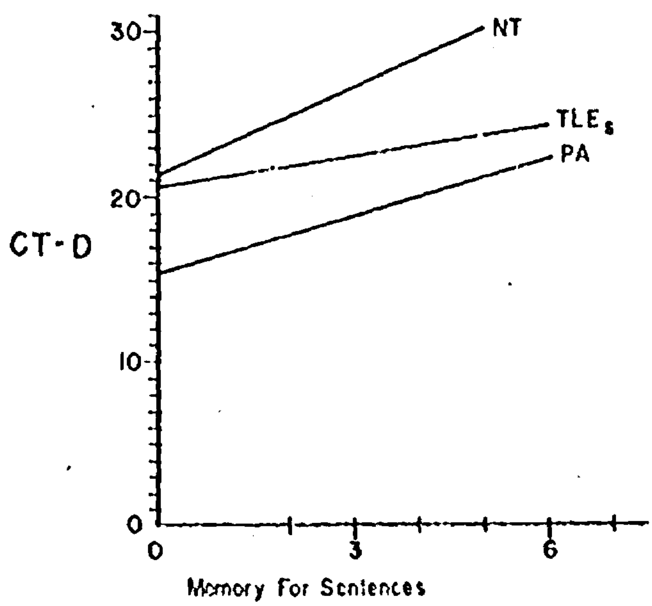

Fig. 14. Regression of CT-D on Memory for Sentences (Study 2).

Summary and Conclusion

Work has already been done on a third study, using related learning materials. Studies 1 and 2 have demonstrated ATIs that are stable over time (immediate and delayed) and over samples (San Jose State and University of Massachusetts). The third study of this series will investigate if the ATIs generalize across subject matter areas. At this time we may state that previous discussions of the main effects (Berliner, 1968, 1969) must be modified in light of the ATIs found in these studies. Snow's (1970) classification of ATIs as being either of the preferential or compensatory models, leads to the conclusion that the TLEs - NT interaction seems to fit the compensatory model. TLEs may function as a memory aid to Ss low in memory ability (e.g. figures 1-4, 11 and 12). The typically high correlation between memory scores and CT scores found in the NT treatment (e.g. r=.72 between Memory for Ideas, Part 1, and CT-BI) indicates a strong relationship between the memory test and CT performance after taking notes. But that relationship is vitiated when an S uses TLEs (e.g. r= -.03 between Memory for Ideas, Part 1, and CT-BI). Study 1 therefore, supports the belief that the TLEs treatment compensates for the lower memory ability of certain Ss, and provides a mechanism for them to do well on the CT. This interpretation was also supported in Study 2, but not as strongly.

The data from these studies are supportive of the notion that TLEs have a place in learning from lecture instruction, provided that a Ss memory abilities are known. To illustrate this an expectancy table, following Cahen's (1969) recommendations, has been developed. Table 6 gives the predicted CT score for the NT and TLEs treatments in Study 1, taking into account the Short Term Memory aptitude (Figure 1). This relationship was the one replicated in Study 2 (Figure 11). Predicted CT scores for the two treatments differ by approximately 1.5 to 2 standard deviations at the extremes. This quite clearly illustrates the importance and usefulness of incorporating the ATI model into instructional research.

Insert Table 6 about here

With regard to note-taking behavior these studies may help explain why the literature generally points out the ineffectiveness of note-taking. Only at very high levels of memory ability is note-taking an effective learning strategy. In most instances the NT condition has been characterized by a steeply rising regression line. However, the PA treatment has been generally characterized by a flatter slope, thus giving rise to situations in which, when memory ability is low, the advantage of NT over PA is minimum. At these levels of memory ability, in some instances, the PA treatment is even superior to the NT treatment (e.g. Figures 1-6). A psychological analysis of note-taking leads to the hypothesis that only when memory aptitude is high does one possess the ability to accurately store orally transmitted information for the time necessary to transcribe it accurately. Likewise, only at such aptitude levels can one store and attend to the new information that is being transmitted while transcribing any previously obtained information. When memory aptitude is low, these abilities are likely not to be present and thus it is as efficient or more efficient for the learner to pay attention to the lecture.

Aptitude treatment interactions are not often found in the literature. The Johnson-Neyman (1936) technique has rarely been used for analysis, though now that some computer programs exist (Carroll and Wilson, 1969; Dowaliby and Berliner, 1970) this technique might become more common. In fact, there is a trend toward analysis of data through the ATI approach

TABLE 6

Predicted Score on CT-BI for Certain Aptitude Scores in the NT and TLEs Treatments of Study 1

Short Term Memory Aptitude Score	Predicted CT Score in NT Treatment	Predicted CT Score in TLEs Treatment	Difference in Predicted Score
3	8.41	21.85	13.44
5	11.61	20.65	9.04
7	14.81	19.45	4.64
9	18.01	18.25	.24
11	21.21	17.05	4.16
14	26.01	15.25	10.76

and methodology. This paper demonstrates a number of interesting ATIs in Study 1, and a replication of some of these in Study 2.

The conditions under which certain ATIs hold or do not hold must be investigated further, and many more aptitudes must be examined for interactive effects. However, even at this preliminary stage it is noted that the ATI approach to instructional research has great potential in educational psychology. Clearly, the treatment main effects noted in the analyses of variance reported above were inadequate to represent the key trends in these data. Further, using a simple effects analysis, after finding an interaction with an analysis of variance, would not have produced the precision in locating the points at which confident decisions could be made. Using the aptitudes an S brings to an instructional situation, and incorporating these into a regression approach to analysis, provides information to a decision maker such that Ss may be assigned to different treatments, each with the greatest potential for individuals with similar aptitude scores.

REFERENCES

Alvord, R. W. Learning and transfer in a concept-attainment task: A Study in individual differences. Unpublished doctoral dissertation, Stanford University, 1967.

Berliner, D. C. The effects of test-like events and note-taking on learning from lecture instruction. Unpublished doctoral dissertation, Stanford University, 1968.

Berliner, D. C. Effects of test-like events and note-taking on learning from lecture instruction. Paper presented at the meeting of the American Psychological Association, Washington, D. C. September, 1969. (Issued also as Technical Report Number 2, Center for Educational Research, School of Education, University of Massachusetts, Amherst, Massachusetts. March, 1970).

Bracht, G. H. The relationship of treatment tasks, personalogical variables and dependent variables to aptitude-treatment interactions. Unpublished doctoral dissertation, University of Colorado, 1969.

Bunderson, C. V. Aptitude x Treatment interactions: Of what use to the ins instructional designer? Paper presented at the meeting of the American Psychological Association, Washington, D. C., September, 1969.

Cahen, L. S. Aptitude treatment interactions: some conceptual and methodological considerations. Research Bulletin 69-96. Princeton, New Jersey: Educational Testing Service, 1969.

Cahen, L. S., & Linn, R. L. Regions of significant criterion differences in aptitude-treatment interaction research. Research Bulletin 70-46. Princeton, New Jersey: Educational Testing Service, 1969.

Carroll, J. B., & Wilson, G. G. An interactive-computer program for the Johnson-Neyman technique in the case of two groups, two predictor variables and one criterion variable. Research Bulletin 69-68. Princeton, New Jersey: Educational Testing Service, 1969.

Coffing, D. G. Eye movement preferences as individual differences in learning. Unpublished doctoral dissertation, Stanford University, 1970.

Cronbach, L. J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cronbach, L. J., & Snow, R. E. Individual differences in learning ability as a function of instructional variables. Final Report, U.S. Office of Education, Contract No. OEC-4-6-061269-1217. Stanford, California: Stanford University, School of Education, March, 1969.

Cronbach, L. J., & Gleser, G. C. Psychological tests and personnel decisions. (2nd ed.) Urban, Illinois: University of Illinois Press, 1965.

Dowaliby, F. J., & Berliner, D. C. ANALATI: A computer program for the application of the Johnson-Neyman technique in the case of 1 predictor and 1 criterion variable. Technical Report No. 8, Center for Educational Research, School of Education, University of Massachusetts, Amherst, Massachusetts, May, 1970.

Frase, L. T. Learning from prose materials: Length of passage, knowledge of results, and position of questions. Journal of Educational Psychology, 1967, 58, 266-272.

Frase, L. T. Boundary conditions for mathemagenic behaviors. Review of Educational Research, 1970, 40, 337-347.

French, J., Ekstrom, Ruth B., & Price, L. A. Manual for kit of reference tests for cognitive factors. Princeton, New Jersey: Educational Testing Service, 1963.

Hamilton, Nancy R. Differentiated response to instruction designed to call for spatial and verbal aptitudes. Unpublished doctoral dissertation, Stanford University, 1968.

Johnson, P. O., & Neyman, J. Tests of certain linear hypotheses and their application to some educational problems. In J. Neyman and E. S. Pearson (Eds.), Statistical Research Memoirs, 1936, 1, 57-93.

Koran, M. L. The effect of individual differences in observational learning in the acquisition of a teaching skill. Unpublished doctoral dissertation, Stanford University, 1969.

Magnusson, D. Test theory. Reading, Massachusetts: Addison-Wesley, 1967.

McDonald, F. J. A model of mathemagenic behaviors as intervening variables in classroom communication. Stanford, California: Stanford Center for Research and Development in Teaching, Research Memorandum No. 30, April, 1968.

Potthoff, R. F. On the Johnson-Neyman technique and some extensions thereof. Psychometrika, 1964, 29, 241-256.

Ripple, R. E., Millman, J., & Glock, M. D. Learner characteristics and instructional mode: A search for disordinal interactions. Journal of Educational Psychology, 1969, 60, 113-120.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965. Pp. 193-221.

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

Rothkopf, E. Z., & Bisbicos, E. E. Selective facilitative effects of interspersed questions on learning from written materials, Journal of Educational Psychology, 1967, 58, 56-61.

Sanders, J. R. Short-term and long-term retention effects of adjunct questions in aural discourse: An extension of research on mathemagenic behavior. Boulder, Colorado: Laboratory of Educational Research, University of Colorado, Research Paper No. 43, July, 1970.

Seibert, W. F., Reid, J. C., & Snow, R. E. Studies in cine-psychometry II: continued factoring on audio and visual cognition and memory. Layfayette, Indiana: Audio Visual Center, Purdue University, Final Report, December, 1967.

Snow, R. E. Toward a model of teacher-learner interaction. Stanford, California: Stanford Center for Research and Development in Teaching, Research Memorandum No. 18, January, 1968.

Snow, R. E. Research on Media and Aptitudes, Viewpoints, (Bulletin of the School of Education, Indiana University), 1970, 46, 63-89.

Snow, R. E., Tiffin, J., & Seibert, W. F. Individual differences and instructional film effects. Journal of Educational Psychology, 1965, 56, 315-326.

Snow, R. E., & Saloman, G. Aptitudes and instructional media. Audio-Visual Communications Review, 1968, 16, 341-357.

Walberg, H. J. Basic educational research in a curriculum project. Paper persented at the meeting of the American Educational Research Association, Los Angeles, California, February, 1969.

Walker, H. J., & Lev, J. Statistical Inference. New York: Holt, Rinehart, and Winston, 1953.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.