MIXED VS. UNMIXED LISTS IN TRANSFER AND RETROACTION

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY

Phyllis G. Peters

1961

MIXED VS. UNMIXED LISTS IN TRANSFER AND RETROACTION

By

Phyllis G. Peters

A THESIS

Submitted to the College of Arts and Sciences of Michigan State University of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Department of Psychology

1961

ABSTRACT

MIXED VS. UNMIXED LISTS IN TRANSFER AND RETROACTION

by Phyllis G. Peters

The purpose of the present study was to study transfer and retroaction as a function of mixed and unmixed list design, using a paired associate verbal task, varying response similarity along the dimensions used by Osgood (1946) while holding stimuli constant. Osgood (1946), using a mixed list design, reported that varying response similarity resulted in negative transfer and retroactive interference which increased as response similarity decreased. Bugelski and Cadwallader (1956), using an unmixed list design, and using the same response similarity dimension, reported that this response variation resulted in retroaction which decreases as similarity decreases. Recently, Twedt and Underwood (1959) found no significant differences in amount of transfer between mixed and unmixed list procedures, but failed to fully check all the response similarity variations of Osgood.

Ninety-six Ss in six experimental groups learned an original list and a transfer list, and were given one recall trial on the original list. The transfer task for the mixed list groups contained all three

response variations - similar, neutral and opposed. The transfer list for each unmixed list group represented only one response variation.

The performance measures used were number of trials to criterion on the IL task, number of correct responses during the first six IL trials, and the number of words recalled during the recall trial. Statistical analyses were performed on the results for mixed and unmixed lists together, and on each separately.

No significant differences were found between the various experimental groups on number of trials to criterion during the IL task or number of words recalled on the recall trial.

Significant differences were found when comparing number of correct responses during the first six IL trials. Neutral responses gave significantly more negative transfer than similar or opposed responses which did not differ from each other. This result indicated: as response similarity decreases, the negative transfer increases, then decreases, then increases to about the same level as the similar condition.

A check on the adequacy of the scale values of Osgood's response similarity dimension with forty Michigan State University student judges found only slight agreement with Osgood's categories. It is possible that the contradictory results may be due to a failure to adequately control response similarity.

The difference in transfer results may also be attributable to the greater number of CI and IL trials required by our Ss to meet criterion compared to those of Osgood and Bugelski.

The mixed and unmixed procedures were not significantly different - confirming the results of Twedt and Underwood.

References:

1. Bugelski, B.R. and Cadwallader, T.C. A reappraisal of the transfer and retroaction surface. J. exp. Psychol., 1956, 52, 360-366.

2. Osgood, C.E. Meaningful similarity and interference in learning. J. exp. Psychol., 1946, 36, 277-301.

3. Twedt, H.M. and Underwood, B.J. Mixed vs. unmixed lists in transfer studies. J. exp. Psychol., 1959, 58, 111-116.

Approved A. M. Barch Date 7 August 1961

ACKNOWLEDGMENT

The author wishes to express her gratitude and sincere appreciation for the guidance and assistance in the planning and execution of this research and the development of this manuscript to Dr. Abram M. Sperling, chairman of her committee.

In addition, she wishes to convey profound thanks to Dr. Paul Baker and Dr. Terrence Allen for their excellent criticism and advice, during the preparation of this thesis.

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
METHOD	10
RESULTS	16
Transfer	17
Retroactive Effects	26
DISCUSSION	32
SUMMARY	38
REFERENCES	41
APPENDIX	43
Instructions	43

LIST OF TABLES

TABLE PAGE

1. Adjectives employed in the experiment in terms of the meaningful relations between original and interpolated materials. . . 11

2. Particular blocks of adjectives employed in mixed lists in terms of the meaningful relations between original and interpolated materials. 13

3. Mean number of trials to IL criterion for mixed and unmixed lists for each block of four adjectives and blocks within similarity relations combined 18

4. Mean number of words correctly anticipated in first six IL trials for mixed and unmixed lists for each block of four adjectives and blocks within similarity relations combined 20

5. Analysis of variance of trials to criterion on IL task for mixed lists . . . 23

6. Analysis of variance of number correct during first six IL trials for mixed lists 23

7. Analysis of variance of trials to criterion of IL task for unmixed lists. . . 25

8. A factorial analysis of variance of the number of words correctly anticipated in the first six IL trials for unmixed lists. 25

9. Mean number of words correctly recalled in the one recall trial for mixed and unmixed lists 27

10. Analysis of variance of number of words recalled for mixed lists. 30

11. A factorial analysis of variance of words recalled for unmixed lists 30

LIST OF FIGURES

FIGURE PAGE

1. Recall scores after interpolated learning 5

2. Mean number of words correctly anticipated in first six IL trials for mixed and unmixed lists. 21

3. Number of words correctly recalled for mixed and unmixed lists 28

INTRODUCTION

Müller and Pilzecker (1900) began the work on retroactive interference when they discovered that learning new material during the interval between original learning and the recall trial made the amount of material recalled less than the amount retained when simple rest filled the interval.

The topic became a popular one for study in the field of learning. One of the factors found to be particularly significant in affecting the amount of interference or facilitation is the amount of similarity between the interpolated and original learning.

Robinson (1927) was the first to clearly specify the effects of similarity in retroaction studies. He proposed the following relationship: "As similarity between interpolation and original memorization is reduced from near identity, retention falls away to a minimum and then rises again, but with decreasing similarity it never reaches the level obtained with maximum similarity. " (pp. 298-299). This relationship gives the well-known Skaggs-Robinson curve of retroactive inhibition.

McGeoch and McDonald (1931) and McGeoch and McGeoch (1937) failed to support the Skaggs-Robinson hypothesis and stated that the greater the similarity

in meaning, the greater the interference. Both the McGeoch studies used judged similarity as the definition of the meaningful relationship, whereas Robinson used identical elements. The method of identical elements defines the similarity dimension by the proportion of elements common to both initial and transfer tasks. (For example, cef and def share two elements for the first degree of similarity; cef and deh share one element for the second degree of similarity; cef and doh share no elements for the third or most remote degree of similarity.) In the method of judged similarity, the judges must rate on some scale the extent of similarity of various words to the reference word. (For example, when "free" is the reference word and direct synonymity is excluded, "open" is judged to have the first degree of similarity, "odd" to have the second degree of similarity, and "closed" to be the most remote.)

Gibson (1940) in a theoretical article on verbal learning distinguished between the similarity of the response in transfer studies. This suggested a reconsideration of Robinson's hypothesized curve in that he failed to distinguish between the similarity of the various parts of the task.

The distinction between stimulus similarity and response similarity found experimental validation in experiments by Gibson (1941) and Hamilton (1943) varying stimulus similarity. Both studied the effect of varying stimulus similarity in a paired-associate task with geometric forms as stimuli and verbal material as responses. They obtained opposite results. Gibson reported that retroactive interference increased with increasing similarity of stimuli. Hamilton reported that retroactive interference decreased with increasing similarity of stimuli. However, Gibson used different and neutral responses while Hamilton used identical responses. Thus, Gibson varied both stimuli and responses while Hamilton varied only stimuli which further stressed the necessity for the distinction between similarity of parts of the task.

Osgood (1946) varied the similarity of the response, holding the stimulus constant. He reported retroactive interference increased as response similarity decreased.

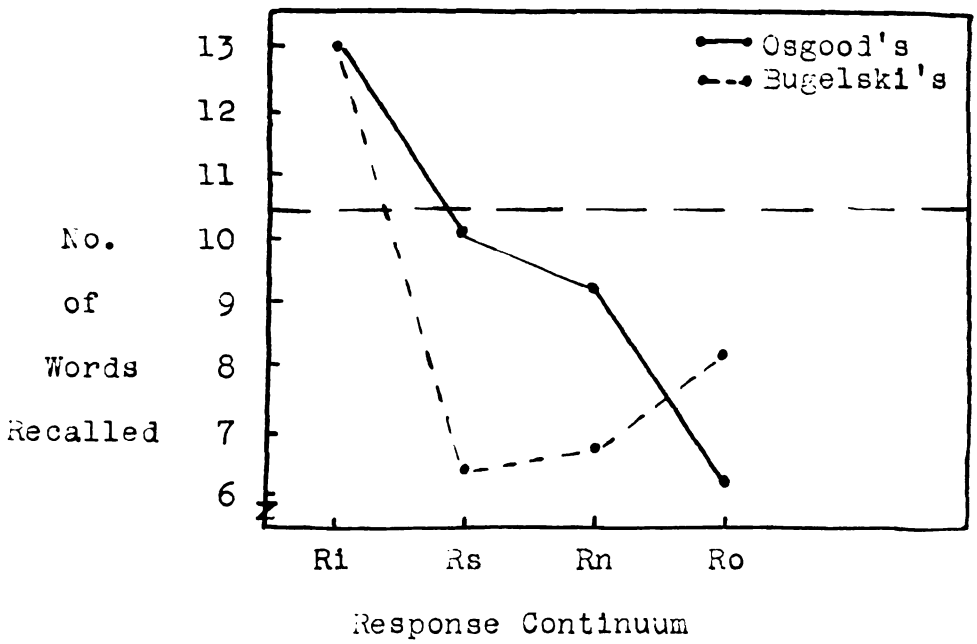
In 1949, Osgood proposed a resolution of the similarity paradox in his transfer and retroactive surface. In this surface, he relates the effect of varying stimulus similarity, response simi-

larity and both types of similarity simultaneously to transfer and retroaction effects. He did not feel it was necessary to distinguish between the methods of identical elements and judged similarity in their effects on transfer and retroaction.

The significant features of the surface can be stated in three empirical laws, based on then available experimental evidence. Of particular interest for this study is the second law which deals with response variation: "Where stimuli are functionally identical, and responses are varied, negative transfer and retroactive interference are obtained, the magnitude of both decreasing as similarity between the responses increases. " (Osgood, 1949, p.135) (See Figure 1). His 1946 experiment was a major experimental support for this law. Osgood regards the contrary results of the McGeoch studies, previously cited, to be due to the simultaneous variation of both stimulus and response similarity.

Bugelski and Cadwallader (1956) made a comprehensive attempt to test all of Osgood's surface in a single experiment. Their results confirmed two of Osgood's empirical laws: (a) Law One, dealing with stimulus similarity, and (b) Law

Figure 1. Recall scores after interpolated learning. A comparison of Osgood's theoretical values with empirical data. The theoretical points were plotted on the basis of a control group performance for the zero effect level and a possible peak score of 13. Recall of original learning material is shown after learning interpolated material consisting of identical stimuli and identical, similar, neutral and opposed responses. (Bugelski and Cadwallader, 1956)



Three, dealing with stimulus similarity and response similarity simultaneous variation.

Contrary to the second law, Bugelski and Cadwallader found that when stimuli are identical, negative transfer is greatest with the most similar responses and least with opposed responses, (See Figure 1). Their results approximated very closely the Skaggs-Robinson hypothesis.

Several differences in procedure might be responsible for the obtained differences.
(1) Osgood used letter pairs (such as c.m.) as his stimuli; Bugelski and Cadwallader used geometric figures from Gibson (1941).
(2) Osgood used a four-second presentation rate for his stimuli; Bugelski and Cadwallader used a two-second rate.
(3) Osgood used mixed lists; Bugelski and Cadwallader used unmixed lists.

Of the three differences, the one involving mixed and unmixed lists seemed the most likely to explain the results. The difference in stimuli was not considered responsible for the difference in results because of the marked agreement between Osgood and Bugelski & Cadwallader on Osgood's first and third laws, both of which involved stimulus

similarity variation. Also, since Osgood used a weighted response measure which involved giving extra credit for correctly anticipating within the first two seconds of the stimulus presentation, the importance of the timing difference is less than would appear at first inspection.

In the unmixed list design each interpolated list represents only one similarity variation. Thus, one group of Ss has an interpolated list in which all the items are similar to the original; another group has an interpolated list in which all the items are neutral in relation to the items of the original; a third group has an interpolated list in which all the items are opposed in meaning or feeling tone. Since the same relation holds throughout the whole interpolated list, S might discern the relation, thus contributing to positive transfer and confounding the experiment. (Slamecka and Ceraso, 1960)

The mixed list procedure attempts to avoid or reduce the possibility of S developing such a special set for the transfer list by incorporating more than one similarity relation into the interpolated list.(Osgood, 1946). Thus, one-third of the items in the interpolated list are similar to the

original responses, one-third are neutral and one-third opposed.

Mixed list designs are also preferred because:

(1) Fewer subjects are needed since it is not necessary to have a separate group of subjects for each condition in the experiment.

(2) Each subject serves as his own control, thus making for more sensitive statistical tests.

Recently, Twedt and Underwood (1959) reported no significant differences in transfer effects as a function of mixed vs. unmixed list design. They compared mixed and unmixed lists in three conditions: (a) identical stimuli - neutral responses, (b) identical stimuli - identical but re-paired responses, (c) neutral stimuli - identical responses. However, the point of greatest disagreement in the results of Osgood and Bugelski occurred in the variation with identical stimuli - similar responses. This variation was omitted in the Twedt and Underwood study. The condition of identical stimuli - opposed responses was also omitted.

Thus, it was still feasible to hypothesize that the differences in the results of Osgood's

work and the Bugelski and Cadwallader experiment might be due to this difference in experimental design.

The purpose of the present experiment is to study transfer and retroaction as a function of mixed and unmixed list design, using a paired-associate verbal task, which varies response similarity along the dimensions used by Osgood, (1946) while holding stimuli constant.

METHOD

Experimental Design. This study followed the traditional retroactive interference design (A - B, A - K, A - B) in which the experimental variable is the meaningful relation of B and K responses in paired-associate learning. Three degrees of meaningful relations were used: similar, (S), neutral, (N), and opposed (O).

All groups had the same list for original learning and for the recall trial.

For the unmixed list design, three groups had an interpolated list with only one degree of similarity - all twelve responses being either similar, neutral or opposed to the responses of the original list.

For the mixed list design, three groups of subjects learned a twelve item interpolated list with a block of four responses similar, four neutral and four opposed to the responses of the original list.

There were a total of six experimental groups.

Lists. The lists were taken from Osgood (1946) and consisted of letter pairs as stimuli, and meaningful adjectives as responses. Table 1

Table 1. Adjectives employed in the experiment in terms of the meaningful relations between original and interpolated materials.

OL	IL		
	Similar	Neutral	Opposed
f.s. - tense	hard	basic	soft
g.l. - free	open	odd	closed
h.v. - dainty	clean	curious	dirty
j.y. - robust	solid	long	flimsy
k.t. - neat	clear	numb	hazy
d.m. - skillful	quick	sour	slow
q.r. - slender	airy	daily	solid
r.h. - noisy	excited	equal	calm
t.g. - drowsy	dead	necessary	alive
w.p. - boorish	rough	near	smooth
x.n. - rounded	graceful	lucky	clumsy
y.b. - pale	sickly	similar	healthy

presents the adjectives used in the experiment in terms of their meaningful relations. Three of Osgood's stimulus-response pairs were removed from each list to reduce the time required to complete the experiment. Table 2 presents the particular combination of adjectives used in the three mixed lists.

Three mixed lists were used so that each response adjective used in the three unmixed lists would appear in the same meaningful relation in a mixed list. Thus, the three sets of four similar adjectives in the unmixed similar list were used each once, the first set in the first mixed list, the second set in the second, etc. This was also true for the neutral and opposed conditions. The order of presentation of the S-R pairs of the mixed lists during interpolated learning was arranged to minimize runs of responses of the same degree of similarity.

Apparatus. For each list learned, three different orders of the S-R pairs were shown on a Lafayette memory drum, with a two-second presentation rate, and a two-second inter-item time. Inter-trial time was four seconds. The original list was shown in window one (left side when facing

Table 2. Particular blocks of adjectives employed in mixed lists in terms of the meaningful relations between original and interpolated materials.

OL	IL Mixed 1	2	3
	Similar	Neutral	Opposed
f.s. - tense	hard	basic	soft
g.l. - free	open	odd	closed
j.y. - robust	solid	long	flimsy
w.p. - boorish	rough	near	smooth
	Neutral	Opposed	Similar
q.r. - slender	daily	solid	airy
t.g. - drowsy	necessary	alive	dead
x.n. - rounded	lucky	clumsy	graceful
y.b. - pale	similar	healthy	sickly
	Opposed	Similar	Neutral
h.v. - dainty	dirty	solid	curious
k.t. - neat	hazy	clear	numb
d.m. - skillful	slow	quick	sour
r.h. - noisy	calm	excited	equal

machine) for all groups. The unmixed lists - similar, neutral and opposed - were in windows two, three and four respectively. The three mixed lists were shown in windows two, three and four respectively. Two memory drums were used; lists were interchanged daily to control for variation due to machines.

Procedure. The general procedure was as follows:

(1) The S was given instructions (presented in Appendix A). The anticipation method was used for all lists.

(2) Ss learned the original list to a criterion of two successive correct repetitions.

(3) A two-minute rest was given in which the S left the machine, and was engaged in conversation by E.

(4) The interpolated list was learned to a criterion of two successive correct repetitions.

(5) A two-minute rest was given as in (3) above.

(6) The S was given one recall trial on the original list. The recall trial was not begun until the S understood that he should anticipate the first time through the list.

Three female experimenters ran all Ss. The

author ran one-half of the subjects kept in the analysis.

Subjects. Subjects were volunteers from the introductory psychology course, who were given research credit for participating in the experiment. A total of 124 subjects were run. Twenty-eight were not used in the analysis, leaving a total of 96 Ss.

Ss were eliminated for the following reasons: one S for not following instructions, fifteen Ss for not reaching criterion on the original list in 50 trials, twelve Ss for not reaching criterion on the interpolated learning at the end of their experimental session.

There were eight females and eight males in each of the six groups. An order of conditions was prepared before the experiment began with each of the six conditions represented once in each block of six Ss, and the order counter-balanced from block to block. Ss were assigned in order of their appearance at the laboratory.

All Ss were naive in verbal learning experiments.

RESULTS

This section is divided into two parts: the first part presents the analyses of the transfer task, the second presents the analyses of the retroactive effects.

Bugelski and Cadwallader (1956) reported a high negative correlation (rho = -.85) between the trials taken to learn the interpolated list, and the number of items of the original list recalled in the one recall trial. Therefore, he presented only the results of the recall trial in his analyses. Osgood (1946) reported the results of both his transfer and retroaction measures.

In the present study, $r = -.62$ for the three unmixed lists, -.48 for the three mixed lists, and -.54 for all lists combined for trials to criterion and items recalled. All correlations were significantly different from zero, ($p < .05$), but somewhat less than expected. Therefore, analyses are presented below for both interpolated learning and the recall trial.

Original learning was fairly uniform and an analysis of variance revealed no significant differences between groups. ($F = 1.05$, $p > .05$).

Transfer

Three measures of interpolated learning were used to ascertain the amount and kind of transfer. (1) Number of trials to the criterion of two successive correct repetitions of the twelve items. (2) Number of trials to the criterion of two successive correct repetitions of blocks of four adjectives common to the mixed and unmixed lists. (3) Number of correct responses in the first six trials of the interpolated list for blocks of four adjectives. (All Ss had at least six trials of interpolated learning.)

a. Mixed vs. Unmixed Lists

Table 3 gives the mean number of trials to criterion for each block of four adjectives for each of the similarity conditions in the mixed and unmixed lists. Each block of adjectives contains the four adjectives that are found in the mixed lists and that are in the corresponding unmixed list.

The difference between the mixed and unmixed list design was tested for each block of four adjectives by means of t tests. Only one of the nine values of t reached the .05 level of signif-

Table 3. Mean number of trials to IL criterion for mixed and unmixed lists for each block of four adjectives and blocks within similarity relations combined.

Similar Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	13.1	15.7	14.7	14.5
Unmixed	13.2	16.1	12.5	14.0

Neutral Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	12.1	13.6	17.4	14.4
Unmixed	17.4	13.9	13.7	15.0

Opposed Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	12.2	15.6	17.6	15.1
Unmixed	12.2	15.4	12.6	13.4

icance: Block 1 of the neutral condition, ($t=2.21$, $p < .05$).

The difference between mixed and unmixed lists was also tested for each condition with blocks combined. The three t tests were not significant.

Table 4 gives the mean number of words correctly anticipated in the first six trials of interpolated learning for mixed and unmixed lists for each block of shared adjectives. Figure 2 presents the condition means for mixed and unmixed lists graphically.

Nine t tests were calculated on the nine blocks testing for differences between mixed and unmixed design with this measure. Only one value of t approached the .05 level of significance: Block 1 of the neutral condition had a t of 3.71, ($p < .01$).

Three t tests were performed on the means of the conditions with blocks combined. None were significant.

b. Mixed Lists

A simple analysis of variance was performed comparing trials to criterion on the transfer task for the three mixed lists. Lists were not significant, ($F = 1.08$, $p > .05$).

Table 4. Mean number of words correctly anticipated in first six IL trials for mixed and unmixed lists for each block of four adjectives and blocks within similarity relations combined.

Similar Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	8.9	10.8	9.5	9.7
Unmixed	10.6	9.7	11.8	10.7

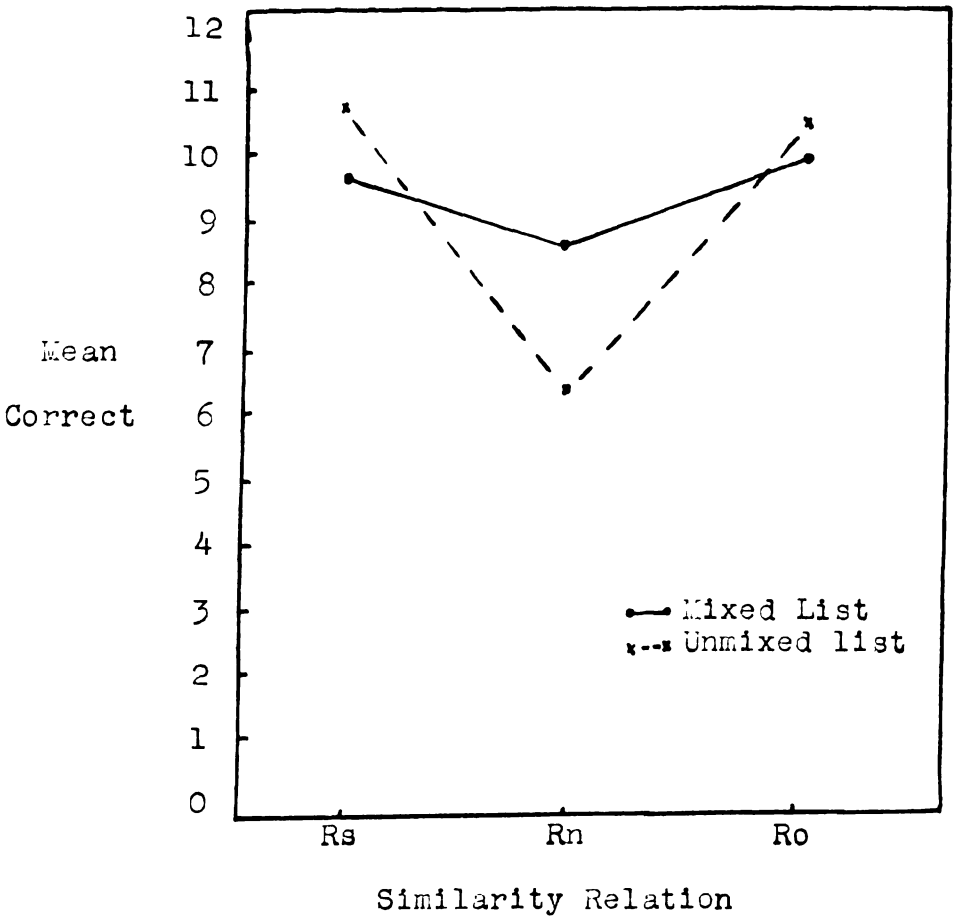
Neutral Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	10.8	7.7	7.7	8.7
Unmixed	5.7	6.1	7.8	6.5

Opposed Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	11.4	10.0	8.0	9.8
Unmixed	11.8	9.6	10.1	10.5

Figure 2. Mean number of words correctly anticipated in first six trials for mixed and unmixed lists.



This was an over-all check of the equivalence of the three mixed lists. No difference was expected as the only difference in the construction of the three lists were the particular words used for the similarity condition.

The similar condition was compared with the neutral (S/N) and with the opposed (S/O), and the neutral condition with the opposed (N/O). Three matched *t* tests were completed on the number of trials to criterion for the three blocks of four adjectives with the combined totals for a condition from all three lists. No value of *t* approached the .05 level. Osgood also performed these tests in his analysis, and found the same results for the S/N and N/O comparisons. However, he found the opposed condition took significantly more trials to reach criterion than the similar condition.

Since the number of men and women were equal in number in each group, a factorial analysis of variance was used to increase the precision of the statistical tests and to compare the results of men and women.

A "Type 1" analysis after Lindquist (1953) was used. Analysis of variance of trials to criterion is shown in Table 5. While the three conditions of similarity were not significantly different from

Table 5. Analysis of variance of trials to criterion on Il task for mixed lists.

Source	df	M.S.	F
Between-Subjects	47		
Sex	1	529.00	4.73 *
error (b)	46	111.83	
Within-Subjects	96		
Similarity (Sim)	2	8.50	.47
Sex x Sim	2	27.91	1.53
error (w)	92	18.17	
Total	143		

Table 6. Analysis of variance of number correct during first six IL trials for mixed lists.

Source	df	M.S.	F
Between-Subjects	47		
Sex	1	100.55	2.96
error (b)	46	33.97	
Within-Subjects	96		
Similarity (Sim)	2	17.06	1.23
Sex x Sim	2	76.29	5.71*
error (w)	90	13.36	
Total	143		

* .05 level of significance

** .01 level of significance

each other, females took significantly fewer trials to criterion then males.

Table 6 gives the analysis comparing the number of words correctly anticipated in the first six IL trials. Neither sex differences nor similarity conditions were significant. However, their interaction was found significant. Individual comparisons revealed that men and women were significantly different only with the condition of "opposed" similarity.

c. Unmixed Lists

The mean number of IL trials to criterion for the similar, neutral and opposed lists are 18.2, 19.8 and 18.3 respectively.

Analysis of variance of number of trials to criterion for men and women for each of the similarity conditions yielded nonsignificant results for the similarity variation. The women took significantly fewer trials than the men. The interaction was not significant. (See Table 7)

However, no significant difference in trials to criterion was found. The substantial correlation ($r=.67$) between original learning and IL suggested the use of analysis of covariance.

Analysis of variance of the number of words correctly anticipated in the first six IL trials for men and women for each of the similarity conditions,

Table 7. Analysis of variance of trials to criterion of IL task for unmixed lists.

Source	df	M.S.	F
Between	5		
Similarity (Sim)	2	11.52	.21
Sex	1	744.18	14.72**
Sim x Sex	2	64.18	1.27
Within	42	50.54	
Total	47		

Table 8. A factorial analysis of variance of the number of words correctly anticipated in the first six IL trials for unmixed lists.

Source	df	M.S.	F
Between	5		
Similarity (Sim)	2	793.75	5.46*
Sex	1	1564.08	10.76**
Sim x Sex	2	66.06	.46
Within	42	145.32	
Total	47		

* .05 level of significance

** .01 level of significance

shown in Table 6, revealed both similarity and sex differences to be significant. The interaction was not significant. Individual comparisons showed the neutral lists required significantly more trials to reach criterion than the similar or opposed lists. The similar and opposed lists did not differ significantly. Women learned the lists faster than the men.

Retroactive Effects

The retroactive measure used is the number of words correctly recalled in the one recall trial. The reader will remember that all groups have the same recall trial, but will be labelled with the relation name they had during interpolated learning.

a. Mixed vs. Unmixed List Design

Table 9 gives the mean number of words correctly recalled for mixed and unmixed lists for each similarity condition for each block of four adjectives. Figure 3 presents the mean of the totals for each similarity condition graphically.

The difference between mixed and unmixed list design was tested for each block of four adjectives by means of *t* tests. Block 1 of the neutral condition again was the only comparison to reach the .05 level. Again, no significant differences were found between

Table 9. Mean number of words correctly recalled in the one recall trial for mixed and unmixed lists for each block of four adjectives and blocks within similarity relations combined.

Similar Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	1.8	1.6	1.6	1.7
Unmixed	1.7	1.8	1.7	1.7

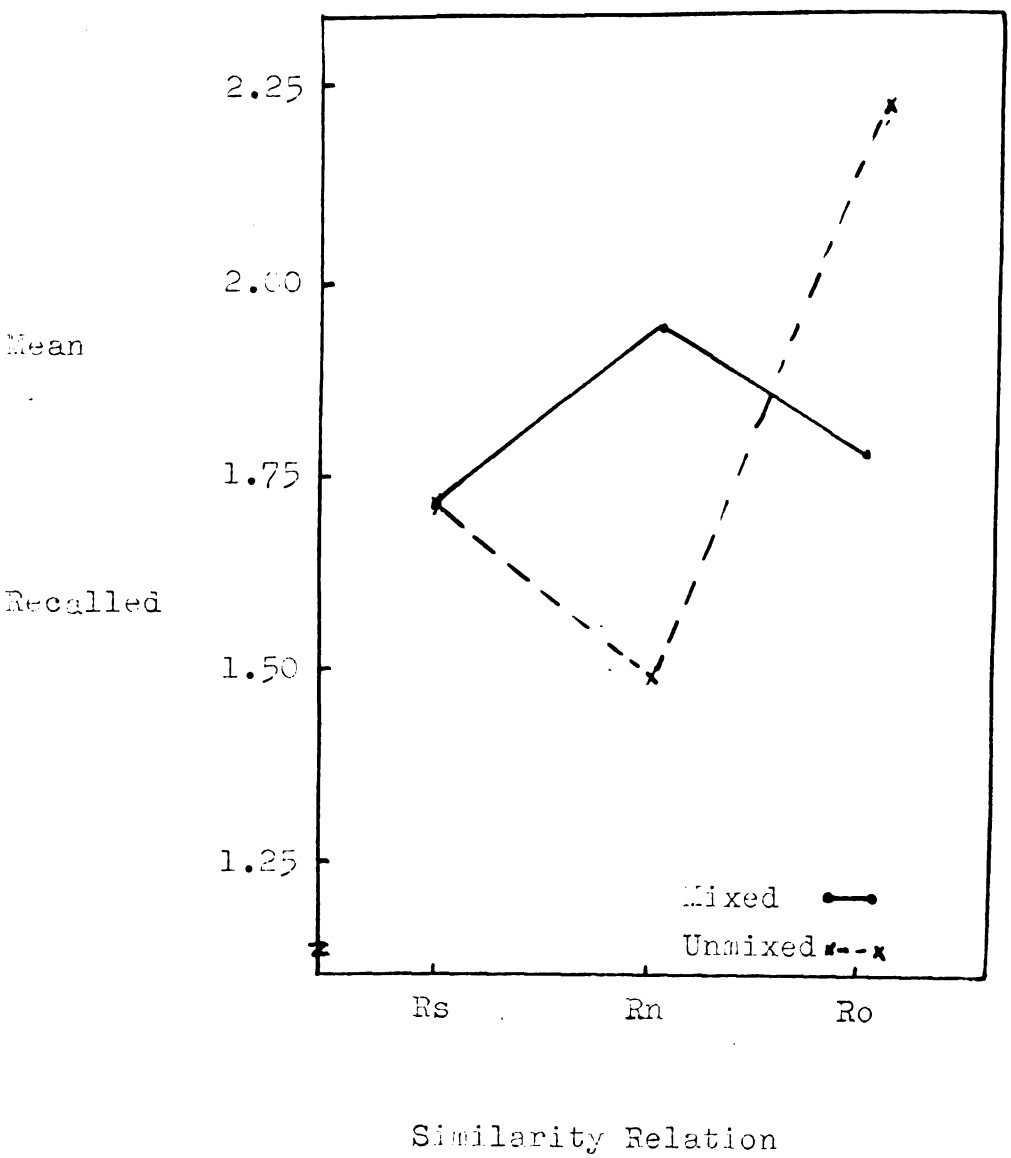
Neutral Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	2.3	1.5	1.8	1.9
Unmixed	1.2	1.7	1.6	1.5

Opposed Relation

	Block 1	Block 2	Block 3	Blocks Combined
Mixed	1.8	1.9	1.6	1.8
Unmixed	1.9	2.3	2.3	2.2

Figure 3. Number of words correctly recalled for mixed and unmixed lists.



mixed and unmixed lists when tested for each condition with blocks combined.

b. Mixed Lists

An analysis of variance of the number of items recalled in the recall list revealed no significant differences between lists. This test was a rough check of the equivalence of the mixed lists.

An analysis of variance using the number of words recalled tested the effects of the interpolated similarity conditions and found no significant differences. (See Table 10)

To check the possibility that the type of analysis used by Osgood might give different results, matched t tests were applied, comparing each similarity condition with the other two. No value of t approached the .05 level of significance. This confirmed Osgood's failure to find significant differences on the first recall trial.

c. Unmixed Lists

The mean number of words correctly recalled of the similar, neutral and opposed groups are 5.0, 4.7 and 6.6 respectively.

An analysis of covariance, adjusting for differences in the original learning, revealed no differences

Table 10. Analysis of variance of number of words recalled for mixed lists.

Source	df	M.S.	F
Between-Subjects	47		
Sex	1	2.01	.67
error (b)	46	2.98	
Within-Subjects	96		
Similarity (Sim)	2	.44	.25
Sex x Sim	2	1.05	.61
error (w)	92	1.71	
Total	143		

Table 11. A factorial analysis of variance of words recalled for unmixed lists.

Source	df	M.S.	F
Between	5		
Similarity (Sim)	2	16.14	2.58
Sex	1	70.09	11.21**
Sex x Sim	2	10.40	1.66
Within	42	6.25	
Total	47		

* .05 level of significance

** .01 level of significance

in number of words recalled. This is contrary to Bugelski and Cadwallader's results, which found significant differences between each condition.

Analysis of variance (Table 11) of words recalled found females recalling significantly more words than males, but not significant differences between the similarity conditions.

DISCUSSION

Osgood reported in his 1946 study, and also stated in his second empirical law, that when stimuli are identical and responses varied, negative transfer and retroactive interference decrease as similarity between responses increases. Bugelski and Cadwallader however, asserted that the law should be modified to read, "... the magnitude of both increasing and then decreasing as similarity between the responses increases." (1956, p. 356)

The results of the present study did not confirm the effect of response similarity variation reported by either Osgood or Bugelski. Our recall measure showed no significant differences. If one merely inspects the recall data for trends, once again there is a lack of agreement. The mixed list groups suggest the least retroactive interference at the neutral condition, and no difference between similar and opposed conditions while the unmixed list groups suggest the most retroactive interference at neutral and the least at opposed. Bugelski found the most retroactive interference at the similar point while Osgood found the least interference at this point.

Bugelski obtained significant differences in his first recall trial and used these exclusively for his analysis. Osgood analyzed both his recall measures

and his IL measures. He found no significant differences on his first recall trial and only a weak trend using the combined results of the 3 RL trials.

Therefore, it is worth inspecting Osgood's transfer results for agreement with ours. Osgood did find a significant difference at the .05 level, using trials to the IL criterion between the similar and opposed items. But, the results for the first and first two IL trials were far more sensitive indicators of transfer. We also found number of correct responses early in IL to be a more sensitive measure of transfer than number of trials to the IL criterion. Thus, we are in agreement with Osgood as to where transfer can be best measured. However, he found the similar condition to have the least negative transfer, and neutral and opposed to show about the same amount. We found similar and opposed to show about the same amount and neutral to show the most negative transfer.

The failure to replicate either the results of Osgood or Bugelski may be due to procedural differences. Osgood and Bugelski both used the method of adjusted or equated learning. This procedure attempts to equate the amount of learning on any given word, by removing words from the list that have met a criterion of anticipation. The present study kept lists intact until the criterion was reached.

Generally, results obtained from adjusted learning procedures have been similar in kind to those obtained with unadjusted procedures. However, the particular criterion of learning we chose may have accentuated the difference between adjusted and unadjusted procedures. Both Osgood and Bugelski had used a criterion of two successive correct anticipations for each stimulus-response pair. In attempting to replicate their studies as possible we used the same criterion of two successive correct anticipations, but applied this requirement to the entire list rather than to individual stimulus-response pairs. The mean number of trials to criterion reported by both Osgood and Bugelski is far less than the mean number of trials to criterion in our study. The mean number of trials to the IL criterion for Osgood's most difficult condition was 4.8. The mean for our most difficult mixed IL condition was 15.1 trials. The mean for Bugelski's most difficult IL task was 6.9 trials. The mean for our most difficult unmixed IL task was 19.8 trials. It is possible that such a difference in learning rate may be due to population differences in verbal ability. The size of the difference seems to point more directly to the interaction between type of learning procedure (adjusted or unadjusted) and the criterion of learning used. If this interpretation is correct, our choice

of criterion increased the amount of OL and IL practice. There is some reason to believe that transfer and retroaction vary as the amount of OL and IL practice is increased. (McGeoch and Irion, 1951). Available information is not sufficiently complete to indicate the manner in which this variable would interact with a similarity dimension.

The failure to obtain significant differences when response similarity was varied suggested that the similarity relation itself should be checked. More than fifteen years have elapsed since Osgood scaled his words and his Yale sophomore judges probably were drawn from a population with more verbal ability than our Ss.

Therefore, a group of 20 (Michigan State University undergraduate) judges, similar to our learning Ss, was given a list of the response words for the OL list and a list of all of Osgood's similar words. They were asked to pair each of the original response words with the word most similar to it from the list of similar words. The same procedure was followed with another group of 20 judges for the opposed words with instructions to match up in terms of an opposed relation. This procedure is closer to the actual learning procedure than Osgood's own technique of judging similarity, since it allows the Ss to see all the words in

relationship to each other, whereas Osgood's Ss made similarity judgments with reference to only one key word at a time.

Only four words in either the similar or the opposed lists were paired up by at least ten of our judges in agreement with Osgood's relations. This lack of agreement was not due to the unreliability of the judgments since at least sixteen judges agreed on pairing three words which were not paired by Osgood.

The extent of disagreement between the two methods of judging similarity relations (plus possible population differences) would seem to argue that studies on transfer and retroaction with verbal material must first make certain that the methods for judging similarity are defensible and that they are in fact cross-situational. Otherwise, studies reporting to vary similarity relations may not be varying this dimension, or varying it in a manner quite different from that intended.

One possible way of avoiding the uncertainties of the judged similarity approach would be a procedure for building up different levels of strength between pairs of words in the laboratory prior to the test for transfer. This could be done by varying the number of joint exposures of pairs of words or by setting levels of anticipation or recall of a word in the pair when the other was presented.

It was hypothesized that the difference between the Osgood and Bugelski studies might be a function of the mixed and unmixed lists design. Instead, the results of this study support the finding of Twedt and Underwood (1959) that there is no difference between mixed lists design. The single significant difference that occurred in one block of four words of the neutral relation seems due to an interaction between list design and the nature of the particular four words involved. This interaction could not be tested directly due to the lack of an appropriate statistical test.

It is still possible that the use of the two designs might affect transfer and retroaction if different meaningful relations are used, such as highly similar responses (synonyms) and antagonistic or opposite responses.

An interesting finding in the present study in the sex differences occurring on each of the variables. Women did consistently significantly better than men in all analyses of unmixed lists and in trials to criterion of mixed lists.

This variable has been neglected in previous studies on verbal transfer and retroaction with adult subjects.

SUMMARY

The purpose of the present study was to study transfer and retroaction as a function of mixed and unmixed list design, using a paired associate verbal task, varying response similarity along the dimensions used by Osgood (1946) while holding stimuli constant. Osgood (1946), using a mixed list design, reported that varying response similarity resulted in negative transfer and retroactive interference which increased as response similarity decreased. Bugelski and Cadwallader (1956), using an unmixed list design, and using the same response similarity dimension, reported that this response variation resulted in retroaction which decreases as similarity decreases. Recently, Twedt and Underwood (1959) found no significant differences in amount of transfer between mixed and unmixed list procedures, but failed to fully check all the response similarity variations of Osgood.

Ninety-six Ss in six experimental groups learned an original list and a transfer list, and were given one recall trial on the original list. The transfer task for the mixed list groups contained all three responses variations - similar, neutral and opposed. The transfer list for each unmixed list group represented only one response variation.

The performance measures used were number of trials to criterion on the IL task, number of correct responses during the first six IL trials, and the number of words recalled during the recall trial. Statistical analyses were performed on the results for mixed and unmixed lists together, and on each separately.

No significant differences were found between the various experimental groups on number of trials to criterion during the IL task or number of words recalled on the recall trial.

Significant differences were found when comparing number of correct responses during the first 6 IL trials. Neutral responses gave significantly more negative transfer than similar or opposed responses which did not differ from each other. This result indicated: as response similarity decreases, the negative transfer increases, then decreases, then increases to about the same level as the similar condition.

A check on the adequacy of the scale values of Osgood's response similarity dimension with 40 Michigan State University student judges found only slight agreement with Osgood's categories. It is possible that the contradictory results may be due to a failure to adequately control response similarity.

The difference in transfer results may also be attributable to the greater number of CL and IL trials

required by our Ss to meet criterion compared to those of Osgood and Bugelski.

The mixed and unmixed procedures were not significantly different - confirming the results of Twedt and Underwood.

REFERENCES

1. Bugelski, B.R. and Cadwallader, T.C. A reappraisal of the transfer and retroaction surface. J. exp. Psychol., 1956, 52, 360-366.

2. Gibson, E.J. Retroactive inhibition as a function of degree of generalization between tasks. J. exp. Psychol., 1941, 28, 93-115.

3. Gibson, E.J. A systematic application of the concepts of generalization and differentiation to verbal learning. Psychol. Rev., 1940, 47, 196-229.

4. Hamilton, R.J. Retroactive facilitation as a function of degree of generalization between tasks. J. exp. Psychol., 1948, 32, 363-376.

5. McGeoch, J.A. and Irion, A.L. The Psychology of Human Learning, New York, Longmans, Green, 1952.

6. McGeoch, J.A. and McGeoch, G.O. Studies in retroactive inhibition: X. The influence of similarity of meaning between lists of paired associates. J.exp. Psychol., 1937, 21, 320-329.

7. McGeoch, J.A. and McDonald, W.T. Meaningful relation and retroactive inhibition. Amer. J. Psychol., 1931, 43, 579-588.

8. Lindquist, E.F. Design and Analysis of Experiments in Psychology and Education, Houghton Mifflin Company, Boston, 1953.

9. Müller, G.E. and Pilzecker, A. Experimentelle beiträge zur lehre vom gedachtniss. Z. Psychol., 1900, Ergbd 1, 1-288.

10. Osgood, C.E. Meaningful similarity and interference in learning. J. exp.Psychol., 1946, 36, 277-301.

11. Osgood, C.E. The similarity paradox in human learning. Psychol. Rev., 1949, 56, 132-143.

12. Robinson, E.S. The similarity factor in retroaction. Amer. J. Psychol., 1927, 39, 297-312.

13. Slamecka, N.J. and Ceraso, J. Retroactive and proactive inhibition of verbal learning. Psychol. Bull., 1960, 57, 449-475.

14. Twedt, H.M. and Underwood, B.J. Mixed vs. unmixed lists in transfer studies. J. exp. Psychol., 1959, 58, 111-116.

APPENDIX A

TASK INSTRUCTIONS

Original Learning

This experiment is designed to test how rapidly you can learn twelve paired items. The first items are two letters such as l.z., and they will appear here in the window. (POINT TO IT) The other items are words, and they will appear here in the window. (POINT TO IT)

This is how the pairs will appear. (SHOW CARD WITH "l.z. - Idea" ON IT) First you will see the letters, then the machine will move, and you will see the letters again with the word they are paired with on the right. The first time through the list, pronounce the letters when you first see them and the words when you first see them. When you see these stars again, you will be going through the list for the second time. Starting then, and on through the rest of the lists, try to anticipate out loud, when you see the letters, the word it is paired with, before it appears in the window. You have two seconds to anticipate. If you say the wrong word, or can't think of any word at all, say the correct word when it appears. The lists are not in any special order, so don't try to memorize that, just the pairs. It is all right to guess. Any questions ?

Interpolated Learning

We are going to learn a different list now, in the same way, so the first time through the list, say the pairs out loud, and then start anticipating.

Recall

Now we will have one trial on the list you learned first. Please try to anticipate as many as you can this one time through the list.