The Memory Laws of Jost

Lars H. Alin
Department of Psychology
Göteborg University

Alin, L. H. The memory laws of Jost. *Göteborg Psychological Reports,* 1997, *27,* No. 1. The memory laws of Jost were analysed together with their experimental conditions. The design of Jost and his followers like Winz, Youtz, and Burdis resulted in confounded data which could not prove the correctness of the laws. The present author, however, extracted non-confounded data from the tables of Youtz (1941) which supported the laws of Jost. Three experiments were also presented which were in favor of the laws. The superiority of distributed over massed learning was discussed, especially its long-term effect. Finally it was pointed out that the laws of Jost, the "permastore" data of Bahrick, and the fact that the retention of single stimuli increases with time, are three independent indications of a consolidation of stored information in memory.

Key words: The memory laws of Jost, confounded and non-confounded data, distributed versus massed learning, the "permastore" of Bahrick, the retention of single stimuli.

Jost (1897) drew the following two conclusions from his experiments: "1) Given two associations of the same strength, but of different ages, the older one will get a greater value on a new repetition; 2) Given two associations of the same strength, but of different ages, the older one will fall off less rapidly in a given length of time."

Jost thought he had demonstrated the correctness of the first law by: 1) the positive effect of distributed learning as compared with massed one; 2) the different results he got by the method of "Treffen" (a kind of anticipation) as compared with the method of relearning (savings). The second proposition was based on Ebbinghaus's experiments with repeated learning.

Jost referred to Ebbinghaus (1885) who had found that 68 repetitions of 12 nonsense syllables in one day, the next day required 7 repetitions to be relearned. Thirty-eight repetitions, however, distributed across 3 days, required only 6 repetitions on the fourth day to be learned to the same criterion.

Jost's first experiments confirmed the results of Ebbinghaus. Thirty repetitions of 12 nonsense syllables on one day required 9 trials to be relearned against 7.6 trials, when the learning was distributed with 10 trials on each of 3 days. These results were the means of 2 subjects who had got some practice before the main experiment. Each subject learned and relearned 4 lists with distributed and 4 lists with massed learning. The different lists were mixed with each other so that 4 lists with distributed and 2 lists with massed learning were mastered in each of 4 days. Each subject thus accumulated a lot of intra-experimental interference of proactive and retroactive kind. This may have influenced the results (cf. Underwood, 1957).

Since Jost did not publish the results of the different lists, we can not know the exact influence of this interference.

Jost did not use any statistical methods to test his results. He found his data convincing, because the distributed learning, as he saw it, meant possibilities of forgetting after 24, 48, and 72 hours respectively, while the massed learning only admitted forgetting after 24 hours. In spite of this difference in retention intervals, the distributed learning was most effective.

In the next experiment only one subject took part, namely Jost's teacher, the well-known professor G. E. Mueller. He had to serve during 50 days. Mueller and Schuman (1894) had criticized Ebbinghaus's results of massed learning for being influenced by fatigue. Jost tried to avoid this factor by mixing distributed and massed learning according to a special schedule. The results (in relearning trials to the criterion) were again in favor of the distributed learning 5.3 trials for massed against 4.6 trials for distributed learning). It seems that Jost with this schedule succeeded in equalizing the effects of fatigue between the two kinds of learning. The interference effects between the different lists, however, must have influenced even these results.

Jost then concentrated on the method of "Treffen". He constructed an electric apparatus (Lippenschluessel) by means of which he could measure the latency time for the subject's response and thus the strength of association of the learned lists. In this experiment he used 3 subjects. The data of one, however, got lost. The goal was to compare 3 different distributions of learning with each other (3,8; 6,4; and 12,2). The first figure refers to the number of days, the second one to the number of repetitions used in each day. The total numbers of repetitions was 24. Each subject took part in all conditions and learned many lists. The conditions were mixed in about the same way as in the preceding experiment. Table 1 gives the results of this experiment.

Table 1
Mean Numbers of Correct Anticipations and Latency Times

	Distributions of Learning		
	3;8	6;4	12;2
Number correct	12.5	35.0	54.0
Total latency times in seconds	2462.5	1891.3	1841.0

The more the learning was spread out, the better were the results both in number of correct anticipations and in latency times. As 4 repetitions only took 4 seconds, the effect of a variation of attention and interest could reasonably be excluded.

What is the main difference between distributed and massed learning? Jost thought it to be the ages of the associations. May be that "older" associations behave in another way than the "younger ones, even if they are of the same strength? This reasoning resulted in the following hypothesis: "If two associations are of the same strength, but of different ages, the older one will benefit more from a new repetition." The hypothesis was tested in a new experiment. The design aimed at a comparison between old and new associations. (When Jost talked about "associations", he meant the average association strength of a list as measured by a retention test). The retention was measured after one minute and 24 hours respectively. Two methods were used: anticipation and relearning. In order to get an equal association (retention) strength of the two lists, they were given different acquisition trials. The "young" lists were read 4 times, the "old" ones 30 times. Table 2 shows the data.

Table 2
Mean Results of "Old" and "Young" Lists

	Numbers of correct anticipations	Total time for anticipation in seconds	Relearning trials
"Old" lists	0.9	4503	5.85
"Young" lists	2.77	1725	9.60

The results confirmed the hypothesis. The numbers of correct anticipations demonstrated a greater strength of the "young" lists. In spite of that the "old" lists were easier to relearn. Jost thought that his results showed two different functions of the memory. It is possible to know much of a certain material, but still need relatively many trials to relearn it ("young" lists). On the other hand one can know relatively little of something, but still relearn it in a relatively short time ("old" lists).

This experiment can be criticized on the same grounds as the previous ones. The possibilities of intra-experimental interference were great. The "young" lists were especially exposed for proactive as well as retroactive interference from the "old" lists. Jost, however, meant that the fact that they still were superior with reference to the numbers of correct anticipations, would strengthen his conclusions.

In order to get further confirmation Jost made a new experiment which principally was designed as the preceeding one, but with some small modifications. One subject took part and read the "old" lists 20 times in stead of 30. The "young" lists were read 6 times in stead of 4. The results were in favor of the hypothesis mentioned above. "Old" lists were easier to relearn (13.6 trials) than "young" ones (17.85 trials) in spite of the fact that the latter ones were of greater strength (2.1 correct anticipations against 0.2).

In judging the latter two experiments one must consider the fact that the lists which were compared, got different numbers of acquisition trials. The "young" lists got 4 or 6 readings, the "old" ones, 30 or 20 readings. Since the retention after one and the same interval of time has been shown to be a linear function of the number of learning trials (Ebbinghaus, 1885), the results of Jost could be explained with reference to the differences in learning. The effect of the differences in learning would be lesser when the intervals of time are different than when they are equal. It can not, however, be excluded.

Jost then discussed the possibility that "old" lists could differ from the "young" ones with reference to the degree of attention. Thirty readings mean a great lot of overlearning. This may result in an equalizing of the strengths of the different associations in the list. It is then probable that next day the strengths of associations of the "old" lists are more equal than those of the "young" lists. This factor could possibly have a differential effect on the two methods used. The anticipation method could favor lists with different strengths of associations (the "young" lists), the relearning method could favor lists with more equal strengths of associations (the "old" lists). Jost tried to test this hypothesis in a specific experiment. This experiment seems to have been badly planned and gave an undecisive result. Jost, however, pointed out that his Experiment VIII decreased the possible differences in strengths of association between "old" and "young" lists from 30 and 4 readings to 20 and 6 respectively, and still gave the same principal results. Jost then used some data of Ebbinghaus in order to formulate a second law. Ebbinghaus, 1885, p. 110-114) reported an experiment in which he learned 12 nonsense syllables to the criterion of one perfect recitation and then tried to see, how many relearning trials he needed to reach the same criterion after 1, 2, 3, 4, 5, and 6 days respectively. The data of Ebbinghaus are seen in Table 3.

Table 3
Mean relearning Trials of Ebbinghaus

	Day					
	1	2	3	4	5	6
Trials	16.5	11.0	7.5	5.0	3.0	2.5

Jost discussed these data in a questionable way. On the assumption that the association strength of the list was the same after the last reading each day, he wondered, why the savings of repetitions increased after each day. His assumptions can, however, be questioned. The strengths of associations could scarcely be equal, since they were the effects of different numbers of distributed trials. Anyhow, Jost tried to solve his incorrectly stated problem by referring to the age of the associations. Finally and after some diffuse reasoning, he tried to explain the results of Ebbinghaus by what has been known as the second law of Jost: "If two associations are of the same strength, but of different ages, the older one will decline more slowly." This proposition has no other support than the above mentioned experiment of Ebbinghaus, a support which can be questioned (the associations were not of the same strength). Strangely enough Jost did not base his second law on the form of Ebbinghaus's memory curve, from which the law can be directly deduced.

The last experiment of Jost (Experiment IX, op. cit. p. 469) was aimed at a confirmation of his first law. He wanted to find support for this law in a third form namely by means of direct learning after different forms of distribution. The question to be answered was what kind of learning would lead to the goal with the least numbers of repetitions. Jost himself was the only subject in this experiment. During 31 days he learned and relearned 24 lists of two kinds. Each day he learned 6 lists by means of 4 repetitions (R-lists) and 6 lists by means of 2 repetitions (R'-lists). The two kind of lists were mixed so that every other was an R-list and every other an R'-list. Next day he relearned the lists to the criterion of one perfect recitation. The lists were relearned in a reverse order to the learning schedule. The mean number of trials to the criterion was 18.5 for the R-lists (4 repetitions in learning) and 17.9 for the R'-lists (2 repetitions in learning). We do not know, if the difference is statistically significant. Anyhow, the results show that 2 repetitions in learning were about as effective as 4, when the lists were relearned after 24 hours. The conclusion drawn by Jost was that even at such low numbers of repetitions, the factor of distribution is effective. The fact that the difference (18.5 against 17.9) was so small, he explained to be the result of retroactive inhibition, although he did not use this very term.

Jost benefitted from the work of Ebbinghaus. His first law was based on the findings of Ebbinghaus. His own experiments specified the conditions of distributed learning and demonstrated empirically based differences between "old" and "young" associations. His second law was based on false asumptions. The experiments of Jost suffer from the fact that only one or two subjects were used. Proactive and retroactive interference has influenced most of his data in an indeterminate way. His use of different acquisition trials for "young" and "old" lists have further confounded his data. Jost had a clear style of writing and performed his experiments carefully. Sometimes, however, he lost the grip and his writing became difficult to understand, especially in the last two parts of his study.

Winz (1931) tested the laws of Jost in a series of experiments with nonsense syllables and substitution tasks. Four-teen students were subjects, 6 women and 8 men. (In his data, however, Winz sometimes used the means of 16, sometimes of 17 individuals). The subjects took part in a lot of experiments with both kinds of materials. The experiments of Jost with massed and distributed learning were replicated. In the massed condition the subjects red a list with nonsense syllables 12 times in the same session. The distributed learning meant that they red a list 4 times on 3 consecutive days. In the massed condition, however, the subjects got a free recall test after each 4th reading. This test was preceeded by a pause during which the subjects had to write their names on a piece of paper. This pause was inserted to eliminate the "echo" of the last syllables. The pause and the free recall test, however, meant a break in the massed learning.

The substitution meant to learn a certain code.In the massed conditions a pause and a free recall test were also inserted after each fourth trial. The free recall test was here a written one. The means of the massed learning conditions, however, exceeded those of the distributed ones in both kinds of materials. For nonsense syllables an immediate retention test gave a score of 9.0 (massed) vs. 7.1 (distributed). For substitution the massed conditions gave 9.7 correct against 7.8 in the distributed conditions. With the help of the free recall tests Winz analysed the trends of acquisition in the data. He found the learning process to be more consequent in the massed conditions than in the distributed ones. The subjects thus more often showed increasingly positive results in the three consecutive tests than they did in the distributed conditions. As Magne (1942) pointed out, the results of

Winz are doubtful, since the repeated free recall tests favoured the method of massed learning. The subjects' knowledge of their results must have been better, when they were aware of them during one learning session than, when there was a time interval of 24 hours between the recall test and the new learning session. Another factor may also have contributed to the failure of Winz to replicate the results of Jost. Winz let his subjects make their final retention test immediately after their last learning trial. Jost gave the test 24 hours later.

Winz then performed some new experiments which he tried to make equivalent to those of Jost. Four subjects took part in the first experiment. The massed condition meant 30 consecutive trials in the same session. The distributed condition split the trials in 10 trials a day during three consecutive days. In each condition the retention was tested by the method of savings 24 hours after the last learning trial. The results now coincided with those of Jost. The distributed learning required only 4.83 trials to the criterion against 6.40 for the massed condition.

The next experiment reduced the number of learning trials to 12. The conditions of massed, respectively distributed learning were the same as in the preceeding experiment. The retention was also tested after 24 hours. The means of 4 subjects were 6.07 trials to the criterion (distributed) against 7.67 (massed). The two last mentioned experiments were then replicated but with free recall tests inserted as in the first described experiments. When 30 trials were used, the means of 5 subjects were now 10.9 correct (distributed) against 9.3 (massed). When 12 trials were used, the distributed learning gave somewhat inferior result than the massed one, 7.4 aginst 7.7. Winz could not explain these results. He interviewed his subjects and tried to explain his data with reference to differences in attention and fatigue. Since his subjects took part in many conditions and thus learned a lot of the same type of tasks, his analyses are not clear-cut or free from objections.

In summary, the results of Winz were in line with those of Jost, when he copied the methods of the latter. Distributed learning then gave better results after 24 hours than massed one. This phenomenom seems not to appear immediately after the learning. The results of Winz are in accord with those of Austin (1921) who tested the laws of Jost with "logical material". She let her only subject read a longer, scientific text 5 times on the same day. The distributed learning meant to read a text once a day on five consecutive days. A test after one day gave practically no difference between the two methods. After 2 and 4 days respectively, the retention after distributed learning was about thrice the value after the massed one. In a second experiment Austin used 32-53 college students. One group got 3 consecutive readings without a pause. The other group got one reading at each of 3 consecutive days. No significant difference was found in immediate recall.

Winz as well as Austin did not directly test the laws of Jost. They used the superiority of distributed learning over massed one as a demonstration of the first law of Jost, but did not directly test any differences between "old" and "young" associations.

Youtz (1941) made an experimental evaluation of the laws of Jost. She pointed out that these laws had an accredited place in the body of psychological knowledge.The basis for their acceptance, however, lied in the often seen superiority of distributed over massed learning and in the form of the retention curves. She also pointed out the inadequacies of the early studies of memory, namely the use of one or two subjects and the lack of any measures of reliability.

The fact that Jost was able to formulate his laws at all, she meant, was an evidence of the uniformity of the results in this field of investigations. Youtz did not only critize the experimental methods of Jost, but also his basic assumptions. The age of the associations is not an active process, but only a frame of reference she meant. She referred to Hull (1935) and Hovland (1938) who used the notion of conditioned responses in rote learning. The dissipation of the secondary inhibition which accumulates when a list is learned, would most probably be the active process which is correlated with the age of the associations. Youtz used the term "inhibition" as a psychological term, not as a physiological one. Rote learning of a list results probably in "competitive interference", due to backward and forward associations, that is intralist-interference.

Youtz let her 15 subjects learn 12 nonsense syllables to three different grades of acquisition: 12, 7, and 4.5 correct anticipations. They got retention tests after 6 seconds, 10, 20, 40, and 60 minutes, 2, and 24 hours respectively. After 6 practice sessions, each subject learned and relearned 36 different lists. The experimental conditions were counterbalanced. Each subject learned one list a day.

Five measures of retention were used: aided recall, numbers of relearning trials, total errors in relearning, saving scores, and second recall scores (from the second relearning trial). Youtz constructed memory curves by means of the five different retention measures. She connected points of different ages and of the same strength of association with each other. These points were situated on different curves, that is from two of the three acquisition grades. The correlation ratio and product-moment correlations were used to test the hypotheses.

All retention measures confirmed the first law of Jost. A formula was presented for the effect of a new repetition, $y = K \log t + c$ (y = the increment caused by a new repetition, t = time, K and c are constants). Youtz summarized her results to the effect that older associations generally demand fewer repetitions in relearning than the younger ones, whatever measure was used. The second law of Jost was also confirmed, at least in one comparison, when the time interval was 20 minutes.

As mentioned above, Youtz explained her results in terms of the notion of dissipation of secondary inhibition. According to the theory of Hull, secondary inhibition would develop in the course of learning a verbal list. These inhibitory tendencies were thought to be the effect of associative bonds between the different items in the list and would accumulate in the middle of the list or just after. They could also be called competitive interference effects, because they were erroneous responses which compete with the correct ones. They should decrease in strength faster than the correct ones and then dissipate. Youtz tried to prove her claim by means of the distributions of errors in the three different degrees or grades of learning, she used. She found that the distribution of errors in position 7 - 8 and 9 - 10 were different in the older measures of retention than in the younger ones. Many of these differences, however, are relatively small, especially between Grades II and III. Furthermore, she ought to have made comparisons between different points on the same curve. The different grades of learning (12, 7, and 4.5 correct) gave different numbers of errors, for what reason her comparisons are misleading. Youtz tried to explain the first law of Jost by reference to the hypothesis that in the older lists, the inhibitory tendencies should have dissipated more than in the younger ones. Therefore a new repetition would give a better result for an old list than for a young one. She thus reduced the problem to be an artifact of method which

especially would concern the method of anticipation. One has further to consider the
competition between correct and incorrect responses during the learning and the result also is dependent on the number of reinforcements which the correct and incorrect responses will get during the learning (Alin, 1964; 1982). The middlemost items or those immediately after the middle will give most errors and thus get the least number of positive reinforcements (correct responses). The erroneous respones have been shown to have a long life-time, over weeks, months, and years (Alin, 1964; 1982; 1986; 1989), and therefore the theory of dissipation of secondary inhibition may be doubted. Furthermore, when Youtz tested the differences betwen "old" and "young" lists she compared lists learned by different amounts of learning with each other. Her results then could be an effect of the different numbers of learning trials and not of the age of the lists. A third possibility would be a combination of the two factors.

A closer inspection of Youtz' data, however, gave the surprising result that she, without realizing it, also had published results from "old" and "young" lists which got the same amounts of learning. In her Tables IV - XII (op.cit.p.18-30) she has recorded some of the increments from the first to the second recall tests. These data show the benefits of a new repetition for "young" and "old" associations in all three degrees of learning. The present author has put them together in Table 4.

The data show that the first law of Jost is supported in all three degrees of learning. In Learning I (12 correct of 12 nonsense syllables) the "young" associations after 2 hours increased by 67.2 % against 228.6 for the "old" ones after 24 hours. In Learning III (4.5 correct of 12) the increment was 9.8 % for "young" lists after 6 seconds against 31.9 % for "old" lists after 10 minutes. The data from Learning II (7 correct of 12) were not quite unambigous. The per cent increments did rise from 15.5 (6 seconds) over 24.4 (20 minutes), 32.8 (40 minutes) and 51.2 (1 hour). There was one exception, however. After 10 minutes the increment was 35.5 % a figure which breaks off the otherwise steady increase with time. The special learning conditions may explain this exception. Each subject took part in all 15 conditions two times. The 15 different lists were learned in the same order by each subject, but the conditions were presented in counter-balanced orders and the second cycle was the reverse of the first. Generally, each subject learned one list a day. Exceptions were the 10 minutes interval. The subjects had to learn this list 30 minutes after having relearned another list from the 24 hours interval. Youtz (op. cit. p. 8) says: "It should be noted that the 10 minute retention session after Learning II of the second cycle came, quite consistently in this possibly disadvantage position."

Table 4

Mean Number of Nonsense Syllables in Recall I and II Together With The Increment Caused by The First Repetition (After Youtz, 1941)

	Learning I			Learning II			Learning III		
Learning of Time	*Rec. I*	*Rec. II*	*% I.*	*Rec. I*	*Rec. II*	*% I.*	*Rec. I*	*Rec. II*	*% I.*
6 sec	9.80			8.39	9.69	15.5	6.45	7.08	9.8
10 min	9.47			5.07	6.87	35.5	4.17	5.50	31.9
20 "	8.90			5.20	6.47	24.4	4.23		
40 "	7.63			4.67	6.20	32.8	3.50		
1 h	6.07			4.10	6.20	51.2	3.53		
2 h	3.87	6.47	67.2						
24 h	1.40	4.60	228.6						

a) Learning I = 12 correct of 12 nonsense syllables; Learning II = 7 correct of 12; Learning III = 4.5 correct of 12; Rec.I = Recall I (the first anticipation trial in the relearning), Rec. II = Recall II (the second anticipation trial).
% I. = Per cent increment from Recall I to Recall II.

Anyhow, the increments from Recall I to Recall II are in 8 of 9 cases in support of Jost's first law. Even when "young" and "old" lists got an equal amount of learning the "old" ones benefitted most by a new repetition.

Rohracher (1963) discussed the so called "paradox of Jost", that is, that "old" associations would benefit more by a new repetition than "young" ones. He referred to Burdis (1958) who replicated some of the experiments of Jost. Burdis got the same results as Jost. He found a significant difference between "old" and "young" lists when the later ones were of an age of 5 minutes. When the interval of time between learning and relearning the "young" lists was increased to 10 minutes, the difference decreased and disappeared after 20 or 60 minutes respectively. Only two subjects took part in the experiments of Burdis. They learned two new lists each day during 20 days. One can not exclude that interlist interferences may have influenced the results. Besides that, the "old" lists were read 20 times, the "young" ones only 6 times. Furthermore, in this experiments there was a confusion between age and the number of learning trials.

In order to test the first law of Jost the present author has made some experiments. For the purpose of minimizing the effects of intra-trial interferences, two conditions were laid down: 1) a subject should only take part in one experimental condition; 2) the method used in learning and relearning should minimize the probability of making erroneous responses.

Experiment I

Method

Subjects. Eighteen students in a first psychology course served as subjects. They were randomly allocated to two groups. Group 1 (Mean age = 22.5 years) consisted of 6 female and 3 female students. In Group 2 (Mean age = 21.O years) there were 5 female and 4 male students.

Design, material and procedure. The learning material consisted of the 10 figures seen in Figure 1. They were presented on a white screen with the help of a drum projector.

Figure 1. The learning material.

Each figure was presented during 5 seconds. The interval of time between the figures was 5 seconds, and between the different presentations one minute. The subjects were instructed to learn the figures as best they could. Group 1 got one presentation, Group 2 got two presentations. A relearning test was given to Group 1 after 24 hours, to Group 2 after one week. The learning and the relearning were carried out in groups. The relearning was performed as a written recall test. Each subject was given a small booklet with 10 empty squares of the same type as those in Figure 1. At a given signal they had to fill in the empty squares with the correct signs. The time allowed was one minute. The series were then presented on the screen in the same way as in the learning, and the subjects had to control their results. If these were totally correct they had to leave the room quietly. If not, they had to turn the page and try to fill in the empty squares on the new leaf. The relearning went on in that like this, until the subject reached the criterion of one perfect performance. Three measures were used: 1) the difference between the first two relearning trials, 2) this difference expressed as increases in per cent of the corresponding scores in the first trial, and 3) the numbers of relearning trials to reach the criterion.

Results

The means and the standard deviations of the first two relearning trials together with their respective differences are given in Table 5.

Table 5
Means and Standard Deviations of The First Two Relearning Trials and Their Respective Differences

	Group 1			Group 2		
	Rel.1	Rel.2	Diff.[a]	Rel.1	Rel.2	Diff.
Mean	2.67	4.78	2.11	2.77	7.22	4.45
s	1.50	2.95	2.76	2.63	2.86	3.13

[a] Rel. = Relearning Trial; Diff. = Difference.

The retention strength of the two groups was nearly the same. Group 2 benefitted most from a new repetition. A t-test for the differences of the increases between the first and the second relearning trials gave a value of $t(16) = 1.68$, $.10 > p > .05$. When the increases were transformed to percentages, the mean values were 104.1 for Group 1 and 358.8 for Group 2. This difference was statistically significant, $t(16) = 2.0$, $p < .05$ All tests in this and the following experiments were one-sided, since the hypothesis put forward meant that the older memories in Group 2 would benefit from of a new presentation.

The numbers of relearning trials to the criterion of one perfect performance were also tested. Group 2 needed a lesser number of trials than Group 1. The mean values were 3.56 (Group 2) against 4.78 (Group 1). A t-test gave the result of $t(16) = 1.412$, $.10 > p > .05$.

In sum, the results were in support of the first law of Jost. The two groups, however, got different numbers of learning presentations. The results thus could be an effect of this difference or a combined effect of the learning differences and the age of the responses.

Experiment II

Method

Subjects. Twenty-one students in a first course of psychology served as subjects. They were randomly allocated to two groups. Group 1 (Mean age = 24.6 years) consisted of 6 female and 4 male students. In Group 2 (Mean age = 24.9 years) 6 female and 5 male students took part.

Design, material and procedure. The same type of figures as in Experiment I were used as learning material. The procedure was the same with the exception that no memory test was given before the first relearning trial. Group 1 got one presentation and relearned after 24 hours. Group 2 got three presentations and relearned after 14 days. Due to unexpected events Group 2 had to relearn in an other room than that in which the learning took place. This fact may have decreased the retention values of these subjects. Furthermore six subjects in this group relearned after 14 days, five after 15 days. The mean interval of time for Group 2 was 14.46 days.

Results

Table 6 gives the means and standard deviations of the first two relearning trials.

Table 6
Means and Standard Deviations of The First Two Relearning Trials

	Group 1			Group 2		
	Rel. 1	Rel. 2	Diff.	Rel. 1	Rel. 2	Diff.
Mean	3.30	6.40	3.10	2.36	5.81	3.45
s	2.11	2.68	1.52	1.75	1.33	1.97

A t-test for the difference scores between the groups gave a $t(19) = .452$, $.35 > p > .30$. When the increases from Relearning 1 to Relearning 2 were transformed to percentages, $t(19)$ was 1.782, $p < .05$. The mean numbers of trials to the criterion of one perfect result were 4.6 (Group 1) and 4.09 (Group 2): Their respective standard deviations were 1.9 (Group1) and .94 (Group2). A t-test showed no significant result, $t(19) = .763$, $.25 > p > .20$.

Also this experiment gave results in support of the first law of Jost. Like Experiment I, however, the two groups got different numbers of learning presentations. The unexpected division of Group 2 in two subgroups, however, meant an opportunity to test two groups with an equal amount of learning, but with different retention intervals. Group 2 was thus split in two sub-groups, 2a and 2b. All these subjects had got the same amount of learning, namely three presentations but relearned at different days, 2a after 14 and 2b after 15 days. The means and standard deviations of these sub-groups are seen in Table 7.

Table 7
Means and Standard Deviations of Subgroups 2a and 2b

	Group 2a			Group 2b		
	Rel. 1	Rel. 2	Diff.	Rel. 1	Rel. 2	Diff.
Mean	2.83	6.0	3.17	1.80	5.60	3.80
s	1.72	1.67	4.81	1.79	0.89	2.28

The divergence between the difference scores was tested by use of a t-test. The result was $t(9) = .424$, $.35 > p > .30$. When the difference scores were transformed to per cent increases, the relationship was 112.0 (Group 2a), against 211.1 (Group 2b). A statistical test gave a $t(9) = 1.495$, $.10 > p > .0.5$. The number of trials to the criterion were 4.17 (2a) and 4.0 (2b). This small difference was not statistically tested. The difference scores were numerically in support of the first law of Jost. Statistically reliable results, however, were not reached. A new experiment was therefore planned to investigate the results when two groups were given an equal amount of learning and when the retention test was given after different intervals of time.

Experiment III

Method

Subjects. Twenty students in a first psychology course served as subjects. They were randomly distributed in two groups. Group 1 (Mean age =24.3 years) was made up of 8 female and 2 male students. Group 2 (Mean age = 24.0 years) consisted of 6 female and 4 male subjects.
Design, material and procedure. The learning material and the procedure were the same as in Experiment II with the exception that a retention test was made immediately after the learning. All subjects got six learning presentations at the same session. They were then tested groupwise, Group 1 after 14 days, Group 2 after 21 days.

Results

The test immediately after the learning gave an insignificant difference between the two groups, 9.8 correct (Group 1) against 10.0 (Group 2). The means and standard deviations of the first two relearning trials and their respective differences are given in Table 8.

Table 8
Means and Standard Deviations of The First Two Relearning Trials and Their respective Differences

	Group 1			Group 2		
	Rel. 1	Rel. 2	Diff.	Rel. 1	Rel. 2	Diff.
Mean	5.50	8.50	2.60	2.30	5.90	3.70
s	3.78	2.95	2.99	1.06	2.39	2.85

Group 1 needed 2.3 trials to the criterion of one perfect series in the relearning (s = 1.06). The corresponding values for Group 2 were 3.70 (s = 1.42). A t-test for the difference gave a $t(18) = 2.5$, $p < .025$. This difference was the reverse to be expected from the data of Jost.

A t-test for the group difference of the increase from Relearning 1 to Relearning 2 gave a $t(18) = 0.35$, $.40 > p > .35$. The increase scores were transformed to per cent values. The data of two subjects in Group 1 with the maximum score of 10 had to be eliminated. Otherwise this group incorrectly would get too low scores. The difference scores were 128.08 % (Group 2) against 94.01 % (Group 1). A t-test gave a value of $t(16) = .658$, $.30 > p > .25$.

The group difference of increase from the first to the second relearning trial was numerically in favor of Jost's first law, but not statistically reliable.

The results of the relearning to the criterion were contrary to those of Jost. This is not astonishing. According to the saving formula of Ebbinghaus and the usual form of memory curves, an older memory task should take more time or trials to be relearned than a younger one. If both tasks got the same amount of learning, the older one should not need a lesser number of trials in the relearning than the younger one. If so, the per cent savings and other measures of retention would steadily increase with time. After a sufficient interval of time there can be a standstill between the two points on the memory curve when it has flattened out. The number of relearning trials, however, will possibly be the same in the corresponding groups but not decrease after the longest interval of time. A non-significant difference in favor of the oldest group may randomly occur. The superiority of Group 2 in Experiment I, however, can be assigned to its greater amount of original learning.

This fact leaves us in a dilemma. In order to meet the requirements of the first law of Jost, one has to compare groups which relearn the same task from the same starting point. In order to bring about that, previous researcher inclusive Jost have given the groups different amounts of learning. Since the relearning data are positively correlated with the number of of learning trials, we can not know, if the difference between the two groups is caused by the different number of acquisition trials or by some intrinsic factor in memory. The last part of Experiment II and Experiment III, however, fulfill the demand of an equal amount of learning. The group which was tested after the longest interval of time benefitted most by a new repetition. The probabilities of these two, independent experiments were then combined according to the method of adding probabilities (Rosenthal, 1978). Edgington`s P (Edgington, 1972) was found to be .06.

An experiment with a prose text. A further support of the first law of Jost can be seen in an experiment with a prose text of 26 words which the present author has performed (Alin, 1964, Experiment XI, pp. 142-154). The text was learned to the criterion of one perfect trial with the help of the method of anticipation. Four groups of students served as subjects, each of 19 individuals. The different groups relearned the text after 1 (Group 1), 7 (Group 2), 14 (Group 3), and 42 days (Group 4) respectively. The first relearning trial was taken as a memory test. The differences between the first and the second relearning trials were taken as measures of the benefit of a new repetition. The mean difference scores were: 1.84 or 8.65 % (Group 1), 3.37 or 18.61 % (Group 2), 5.53 or 43.8 % (Group 3), and 8.10 or 76.53 % (Group 4). Since the variances were very much heterogeneous in raw scores as well as in per cent values, transformations to log (1+x) were used. The transformations reduced the heterogeneity to acceptable values. Fmax (Hartley, 1950) was 1.63, $p > .10$ for the raw scores and 1.70, $p > .20$ for the per cent values. Statistically significant values were found in both ANOVAs. $F(3, 72)$ was 9.99, $p < .01$ (raw scores) and 9.74, $p < .01$ (per cent values). Some facts, however, make these results discussable. In the first relearning trial, five subjects in Group 1 reached the maximum score of 26, four in Group 2, two in Group 3, and none in Group 4. These facts would put the first three retention groups at a disadvantage when all groups were compared with respect to the benefits of the first repetition. Two different methods were used in order to neutralize these effects. The differences of zero were either replaced by the respective group means of increase from the firsty to the second relearning trial or simply excluded. Table 9 gives the means and standard deviations of the per cent increases when the zero scores were replaced by the mean group increases.

Table 9
Means and Standard Deviations of the Per Cent Increases from the First to the Second Relearning Trial

	Group 1	Group 2	Group 3	Group 4
Mean	11.73	23.83	49.67	76.53
s	11.54	19.00	44.23	110.08

An ANOVA resulted in an $F(3, 72) = 3.89$, $p <.05$. Since the variances were heterogeneous the percentages were transformed to log (1+X). Fmax was reduced to 1.77, $p > .05$. $F(3, 72)$ was 7.5, $p < .01$.

The group means and standard deviations when subjects with zero scores were excluded are seen in Table 10.

Table 10
Group Means and Standard Deviations with Zero Scores Excluded

	Group 1	Group 2	Group 3	Group 4
Means	13.70	23.57	48.95	76.53
s	11.54	19.00	44.23	110.08

An ANOVA resulted in an F (3, 62) = 3.73, p < .05. Since there was a good deal of heterogeneity of variance (Fmax was 88.8, p <.01), the raw scores were transformed to log (1+X). Fmax was now reduced to 2.88, p >.05. F (3, 62) was 5.99, p<.01. The results mean that the ”older” memories benefitted most by a new repetition. That was especially true of the differences between Groups 4 and 2, Groups 4 and 1, Groups 3 and 1 which were significantly reliable as measured by t-tests.

The first law of Jost also states that the retention values of the groups to be compared should be equal. An ANOVA for the mean results of the first relearning trial gave an F (3, 62) of only 0.643, p >.05.

The non-parametric analysis of variance of Kruskal and Wallis (1952) was also used on the original raw scores. The resulting H was 15.15. With 3 degrees of freedom this gives a p <.01.

The results of Jost, Winz, Burdis, and Youtz were based on groups with different amounts of learning. This was also the case with the first two experiments of the present author. As has been pointed out above, such data are confounded. Three experiments of the present author, however, equalized the amounts of learning between the groups, and were still in support of the first law of Jost. To that must be added the data of Youtz which the present author has gathered in Table 4. With one exception they confirm the conclusion that memories are in some way strengthened with time.

The Problemacy of Massed and Distributed Learning

Jost (1897) chiefly based his first memory law on the superiority of distributed to massed learning. He was inspired to that by the experiments of Ebbinghaus (1885). Another experiment of Ebbinghaus is seldom cited but of great interest. Ebbinghaus (1919 p. 722) learned some stanzas of Byron by massed learning and relearned them after 22 years with a saving of 7 per cent as compared with the learning of new ones. He learned other stanzas of the same quantity to the criterion of one perfect performance during 4 consecutive days. After 17 years the relearning resulted in a saving of 20 per cent. Even with consideration to the difference in time interval of four years and that the latter learning was cumulative, this indicates a remarkable superiority of distributed learning after such a long time.

The problems of massed versus distributed learning have been studied for more than hundred years. An over-whelming mass of experiments have repeatedly demonstrated the superiority of distributed learning over massed one. A convincing explanation of the superiority of the spacing of learning, however, is still missing. In an excellent series of experiments Bahrick (1979; 1984; Bahrick & Phelps, 1987) has explored these problems. He invented a new technique, "the method of drop-out". The first learning session started with a presentation of a series of paired associates. There followed alternating test and presentation trials. Only the items which failed on a test trial were given on the next presentation trial, the rest were dropped out. A session ended with the first test trial on which all remaining items were passed. At this point the subject had made a single correct response to each target item. The only difference between the original learning and the relearning was that the relearning started with a test trial. In some of these experiments Bahrick used 50 English-Spanish word-pairs. Different groups learned the word-pairs with different time intervals between the original learning and the relearning sessions, namely 0, 1, and 30 days. The learning was cumulative, that is in all sessions the subjects learned and relearned the task to the criterion of one perfect performance. Bahrick & Phelps (1987) studied the retention of these 50 word-pairs after 8 years. Forty-eight of the 64 subjects who had served in the original study agreed to participate in a follow-up study. Thirteen of them were eliminated because they had been exposed to the Spanish language during the 8 years or because they failed to return the new test. The remaining 35 subjects got a recall test and later on a recognition test with all words in English which the subject had failed to recall.The recognition test was of multiple-choice type. In Table 12 are seen the percentages of words recalled and recognized after 8 years.

Table 12

Percentage of Words Recalled, Recognized but Not Recalled and Failed as a Function of the Primary Intersession Interval (Bahrick & Phelps, 1987, Table 2, p. 346)

	% Correct		
Primary intersession interval	Recall test	Recognition test	% Failed on both tests
30 days	15	83	14
1 day	8	80	18
0	6	71	27
control	1	62	37

There is an over-whelming superiority for the group which learned with 30 days intersession intervals against the groups with 1 or 0 days interval. These data are, however, confounded by the variable number of learning sessions used 8 years earlier.

The cumulative relearning to the same criterion meant that the different groups started from different points of the learning curve. The longer the intersession interval was, the more relearning trials were needed to reach the same criterion. Bahrick, however, had noted all numbers of presentation for each word and each subject during the first learning. He could thus compute the probability of recall and recognition after 8 years as a function of the intersession interval and the number of presentations during acquisition. They are reproduced here in Table 13.

Table 13

Probability of Retention as a Function of the Intersession Interval and the Number of Presentations During Acquisition (Bahrick & Phelps, 1987, Table 4, p. 347)

Intersession interval (in days)	Acquisition presentations 1-2	3-4	5-6	7+
	Recall			
0	.09	.03	.02	.00
1	.15	.06	.03	.01
30	.23	.15	.14	.07
	Recognition			
0	.78	.66	.62	.54
1	.84	.80	.76	.72
30	.92	.92	.78	.71

The data of Table 13 indicate that an intersession interval of 30 days even after 8 years was much superior to the intervals of 1 or 0 days independently of the number of acquisition presentations. Bahrick (1984) showed that Spanish vocabulary acquired in high school or college courses was recallable for more than 25 years, if it not was lost during the first 5 years following learning. He used the term "permastore" to designate unrehearsed memory content with a life span of about 25 years. The 1984 data indicated that the conditions of acquisition were decisive for the longevity of individual target items and not the conditions during the retention interval. In the 1987 study he found that 8-year retention probability was greatly enhanced for words which were well encoded in one or two presentations and subsequently accessed several times at intervals of 30 days. Bahrick meant that encoding variability was not critical for permastore memory as suggested by Thompson, Wener, and Bartling (1978).

The present author want to make two additions to the works of Bahrick. The fact that the more presentations a target item needed, the worse was the retention after 8 years, can be explained by reference to the competition between correct and wrong responses. It has been shown that even the erroneous responses have a long life-time, over weeks, months and years (Howe, 1970; Alin, 1964; 1982; 1986; 1989). Since the drop-out method means that a correct response can be retained over 8 years, (cf. p. 31-32), the erroneous responses, especially if they are repeated, may survive just as long. This can explain the negative correlation between retention values and number of presentations in the learning. Another comment to the impressive results of Bahrick is that they fit well into a theory of consolidation. The results of Bahrick and the data supporting the first law of Jost may explain the fact that distributed learning in general is superior to massed one. If stored information is strengthened with time, relearning after a longer interval of time will work upon a stronger memory than after a shorter one. This contention is also supported by a series of experiments with single stimuli (King, 1963a; 1963b; 1963c; and 1965; Alin et al. 1980a; 1980b; 1980c; and 1980d). Single stimuli as the length of a line, the loudness of a tone or the brightness of a light give retention values which increase as an approximate logarithmic function of time. Proactive or retroactive stimuli of the same kind as the target interfere with the memory of the target and result in forgetting.

In sum, the first law of Jost, Bahrick`s findings about the permastore and the increasing strength of simple stimuli with time all lead to the conclusion that memories increase in strength with time.

References

Alin, L. H. (1964). *Experimental studies in verbal versus figural learning and retention*. Stockholm: Almqvist & Wiksell.

Alin, L. H. (1982). *The retention of erroneous responses made in verbal learning.* (Göteborg. Psychological Reports, 12, No. 8.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H. (1986). *Some facts and artifacts in psychological memory curves.* (Göteborg Psychological Reports, 16, No. 1.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H. (1989). *The retention of connected discourse after 27 years*. (Göteborg Psychological reports, 19, No. 5.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H., Forslund, G., Hamberger, K., Svensson, L., & Söderström, G. (1980a). *The retention of a single event I: Remembering the length of a line.* (Göteborg Psychological Reports, 10, No. 5.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H., Lorentzon, B., Nobel, A.-M., Ribenfors, E., & Wik, G. (1980b). *The retention of a single event II: Remembering the loudness of a single tone with and without proactive and retroactive interference*. (Göteborg Psychological Reports, 10. No. 7.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H., Rosen, M., Spångberg, A., & Möller, M. (1980c). *The retention of a single event III: The retention of the loudness of a tone with and without proactive and retroactive interference.* (Göteborg Psychological Reports, 10, No. 8.) Sweden: Göteborg University, Department of Psychology.

Alin, L. H., Gustafsson-Lundqvist, C., Larsson, E., & Wetterholm, P. (1980d). *The retention of a single event IV: The retention of the loudness of a tone with and without proactive and retroactive interference. A complementary study with a* *change of method*. (Göteborg. Psychological Reports, 10, No. 9.) Sweden: Göteborg University, Department of Psychology.

Austin, S. D. M. (1921). A study in logical memory. *American Journal of Psychology, 32,* 370-403.

Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General, 108,* 296-308.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1-29.

Bahrick, H. P. (1987). Retention of spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 344-349.

Burdis, H. (1967). Experimentellen Untersuchungen zum I. Jostschen Satz. *Wiener Zeitschrift fuer Philosophie, Psychologie und Pädagogik. VIII.*

Ebbinghaus, H. (1885). *Ueber das Gedächtnis.* Leipzig: Verlag von Duncker & Humbolt.

Ebbinghaus, H. *(1919). Grundzuege der Psychologie* (4th rev. ed.) I, p. 722. Leipzig: Verlag von Veit & Comp.

Edgington, E. S. A. (1972). A normal curve method for combining probability values from independent experiments. *Journal of Psychology, 82,* 85-89.

Hartley, H. O. (1950). The maximum F-ratio as a short-cut test for heterogeneity of variance. *Biometrika, 37,* 308-312.

Hovland, C. I. (1938). Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and distributed practice. *Journal of Experimental Psychology, 22*, 201-224.

Howe, M. J. A. (1979). Repeated presentation and recall of meaningful prose. *Journal of Educational Psychology, 61,* 214-219.

Hull, C. L. (1935). The conflicting psychologies of learning - a way out. *Psychological Review, 42,* 491-516.

Jost, A. (1897). Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. *Zeitschrift fuer Psychologie und Physiologie der Sinnesorgane, 14,* 436-472.

King, H. E. (1963a). The retention of sensory experience I. Intensity. *The Journal of Psychology, 56,* 283-290.

King, H. E. (1963b). The retention of sensory experience II. Frequency. *The Journal of Psychology, 56*, 291-298.

King, H. E. (1963c). The retention of sensory experience III. Duration. *The Journal of Psychology, 56,* 299-306.

King, H. E. (1965). The retention of sensory experience IV. Short delays versus long-delay intervals. *The Journal of Psychology, 60,* 101-113.

Magne, O. (1942). Inlärning och behållning med utgångspunkt från allmänna minneslagar. *Tidskrift för Psykologi och Pedagogik. Göteborgs Högskolas Psykologiska och Pedagogiska Institution. 1. No. 2.* Sweden: Göteborg.

Mueller, G. E. & Schumann, F. (1894). Experimentelle Beiträge zur Untersuchung des Gedächtnisses. *Zeitschrift fuer Psychologie, 6*, 81-190.

Rohracher, H. (1963). *Gedächtnis und Lernen.* Separatdruck aus Meili, R. & Rohracher, H. *Lehrbuch der Experimentellen Psychologie.* Bern und Stuttgart: Hans Huber.

Rosenthal, R. (1978). Combining results from independent studies. *Psychological Bulletin, 85,* 185-193.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49-60.

Winz, A. (1931). Neue Versuche ueber Lernen in Häufung und Verteilung. *Psychotechnische Zeitschrift, 6,* 129-168.

Youtz, A. C. (1941). An experimental evaluation of Jost`s laws. *Psychological Monographs, 53, (*No.1. Whole No. 238).