Detection of repeated trigrams: Evidence of all-or-none learning

ARNOLD GLASS
Rutgers University

ARILD LIAN
University of Oslo

JAMES LAU
Watchung Hills Regional High School

Two experiments investigated trigram detection in a continuous recognition task. In Experiment 1 consonant trigrams were presented visually, one at a time, with occasional repetition of a trigram after an interval of 0, 2, 4, 8, or 16 other trigrams. Subjects were told to respond with a button press every time they saw a repeated trigram. If a subject responded to a repeated trigram, it was not repeated again. However, if a subject did not respond to a repeated trigram, it was repeated again at the same interval for up to 3 repetitions. For all intervals greater than 0, the probability of noticing a repeated trigram did not increase with the number of repetitions. In Experiment 2 meaningless shape trigrams were presented, and occasionally a trigram was repeated after an interval of 0, 1, or 2 trigrams. For both intervals greater than 0, the probability of noticing a repeated trigram did not increase with the number of repetitions. The results demonstrate that a repeated input does not necessarily leave a permanent trace in memory.

If an event is repeated in the experience of a person, that person comes to recognize the event as a repetition. For example, if a person sees a sequence of consonant trigrams and a particular trigram is repeated, then after a certain number of repetitions the person comes to notice that the trigram has been seen before.

This simple account leaves many questions unanswered about the shape of the learning curve and the factors that influence it. For example, consider a consonant trigram that has been repeated several times at a constant interval of *m* trials. First consider the probability that it is recognized as a repetition the first time it is repeated, p_1. Similarly, p_2 is the probability that it is recognized as a repetition the second time it is repeated, given that it is not recognized as a repetition the first time it is repeated. Finally, p_3 is the probability that it is recognized as a repetition the third time it is repeated, given that it is not recognized as a repeti-

AMERICAN JOURNAL OF PSYCHOLOGY
Winter 2004, Vol. 117, No. 4, pp. 497–515

tion the first two times it is repeated. The question this experiment addresses is whether $p_1 = p_2 = p_3$ or $p_1 < p_2 < p_3$.

Pre-theoretically, either result is possible. It could be the case that a trace of the consonant trigram is encoded into short-term memory and decays after a certain number of seconds if it is not refreshed by an external representation or internal rehearsal, either of which might count as a repetition. Furthermore, each repetition may transfer the trigram to long-term memory with some probability. According to this view, if a repetition of a consonant trigram went unnoticed this would indicate that it was absent from short-term memory and had not been matched in long-term memory. Hence, the effect of the presentation would be the effect for a "new" item uncontaminated by any previous memory trace. So it would be the case that $p_1 = p_2 = p_3$.

Alternatively, it could be the case that each repetition of the trigram is matched to its trace in long-term memory, and a match between a memory trace and an input causes the activation level of the memory trace to increase. So each repetition strengthens the trace such that the next time the consonant trigram is presented, its activation is higher. When the activation produced by the match of an input with its memory trace exceeds some criterion, the input is recognized as a repetition. Such assumptions are consistent with Morton's (1969) classic logogen model and its successors. According to this view every repetition increases the baseline activation level of its memory trace whether or not the consonant trigram is recognized. So it would be the case that $p_1 < p_2 < p_3$.

The experimental literature does not answer the question of the effect of an unnoticed repetition because in previous experiments the subjects were aware of the repetitions (Hebb, 1961; Glass, Krejeci, & Goldman, 1989). It was the purpose of this experiment to do so. In this experiment consonant trigrams were presented visually, one at a time. Subjects were told to respond with a button press every time a repeated trigram was seen. If a subject responded to a repeated trigram, it was not repeated again. However, if a subject did not respond to a trigram, then it was repeated again, for up to three repetitions. Finally, trigrams were repeated at intervals of 0, 2, 4, 8, or 16 items. If a trigram was not detected, it was repeated at the same interval. For example, if a target trigram was presented and then after an interval of four distractors it was repeated but not detected, it was repeated again after another four distractors, up to a maximum of three repetitions. A 1,300-ms study interval was used so that there would be no question that the subject had sufficient time to encode the study item. Consonant trigrams rather than words were used so that the dominant effect would not be prior familiarity with the study item.

The two hypotheses just described may be characterized as all-or-none learning and incremental learning. Initially, evidence of all-or-none learning was sought in studies of paired-associate learning. Consider a paired-associate learning task of alternating study and test trials. On a study trial 12 A–B study pairs were presented and on a test trial the 12 A terms were presented, and the subject had to respond with the B term. The study–test sequence was repeated until all 12 B response terms were correctly reported. Rock (1957; Rock & Heimer, 1959) tested the all-or-none learning hypothesis by modifying this control procedure. In the replacement condition, if the B response was not given on the test trial, then the A–B study pair was replaced with an all-new C–D study pair on the next presentation of the study list. The results were that the mean number of trials needed to achieve the criterion of all 12 responses correctly reported did not differ between the control and replacement procedures. Therefore, Rock argued that the subjects must not have developed any associative strength between the stimulus and response terms for the items missed. Had such associative strength developed, the elimination of the missed pairs should have retarded overall learning because for the new pairs the associative strength would have to be built up from scratch. The fact that the insertion of new items did not slow down learning was taken to mean that no strength was built up for the missed pairs and consequently the association must be an all-or-none affair. In his first experiment he used letter–number pairs, and in his second experiment he used nonsense syllables.

Unfortunately, Rock's (1957) procedure had two flaws, which were demonstrated by Underwood, Rehula, and Keppel (1962). New items in the replacement condition were easier and received more rehearsals, allowing subjects to learn the list with replacement items as fast as the control list. So Rock's studies do not provide evidence of all-or-none learning. This report reopens the question with a new experimental paradigm. The item selection effect that undermined Rock's conclusion is explicitly addressed in this article.

EXPERIMENT 1

The purpose of this experiment was to determine the effect of unnoticed repetitions on the probability of detecting a target and the interaction of this variable with the interval between repetitions of the target. Presumably, target detection would be a negative function of repetition interval because the greater the number of intervening distractors, the greater the interference with the target's representation. Within the framework of the incremental hypothesis, hit rate should be positively correlated with the number of target repetitions (and negatively correlated with repetition interval if forgetting occurs between target repetitions). In contrast, the all-or-none hypothesis predicts a slope of zero at all repetition intervals.

METHOD

Subjects

Eight subjects (five men and three women) between the ages of 21 and 65 participated in the experiment. Five were native speakers of Norwegian and three were native speakers of English.

Design

During a block of trials a single target trigram was repeated in a sequence of distractor trigrams until a response to it was made or it had been repeated three times. To provide sufficient power to reliably determine the shape of the repetition detection function, each subject saw 540 blocks. Given the large number of trigrams per block it was not possible for a subject to perform the entire task in a single day. The experiment took place over a 1-week period. For 6 days of the week each subject participated in two sessions a day, one usually before noon and the other in the late afternoon. During a session a subject was shown three-consonant trigrams at the rate of one trigram per 1,300 ms on a computer monitor. The subject was instructed to press a button on the keyboard of the computer every time the trigram was a repetition of one that had appeared earlier. During a session a subject saw nine block sequences, separated by pauses of at least 30 s.

The trigrams were constructed as follows. Each block sequence consisted of five blocks of trials. For each block of trials a different retention interval was used. The retention intervals used were 0, 2, 4, 8, and 16 trigrams. At the beginning of the session the order in which the retention intervals were presented was randomly generated. From the subject's point of view the task was a continuous recognition task. However, each block of trials actually consisted of a study sequence followed by up to three test sequences. At the beginning of the study sequence, a number from 1 to 10 was randomly selected, and that number of randomly generated consonant trigrams was presented. Then the target trigram was randomly generated and presented. The presentation of the target trigram concluded the study sequence. Next, for an interval of m, m randomly generated consonant trigrams were presented and then the target trigram was re-presented. For example, if $m = 2$, then two randomly generated consonant trigrams were presented before the target trigram was repeated a first time. If the subject responded with a button press indicating that the target trigram was a repetition, this ended the block. Otherwise, if the subject did not respond, another m randomly generated consonant trigrams followed by the target trigram were presented. Again, if the subject responded with a button press indicating that the target trigram was a repetition, this ended the block. Otherwise, another m randomly generated consonant trigrams were presented, followed by another repetition of the target. This time the block ended whether or not the subject responded to the target.

Trigrams were generated without replacement. That is, except for target repetitions, each trigram appeared no more than once during a session. In generating a trigram, the consonants were selected without replacement, so a trigram consisted of three different consonants. Also, the completed trigram was compared with a list of familiar consonant trigrams (e.g., BMW), and if it was on list then the trigram was discarded and another trigram was presented instead.

Procedure

A subject began each block sequence by typing a code number that identified the subject and the block sequence. Next, the instructions appeared. Each subject was instructed to read the directions before his or her first experimental session. When the subject pressed any key, the experiment immediately began; the first consonant trigram was presented for 800 ms, followed by three underline symbols in the locations of the trigrams for 500 ms, to bring the total stimulus onset asynchrony (SOA) to 1,300 ms. Then the next trigram was presented. All subjects responded to a repetition by pressing the spacebar. Because the size of each study sequence and the order of the intervals were randomly selected, a subject could not predict when a target repetition would occur.

RESULTS

A p level of .05 was adopted for the analyses. The hit rate and false alarm rate for each subject were computed. The false alarm rate was .13 and did not vary across repetition interval. A repeated-measures analysis of variance (ANOVA) was performed in which the independent measure was repetition interval (0, 2, 4, 8, or 16) and the dependent measure was hit rate for a single repetition. Hit rate declined as repetition interval increased, $F(1, 7) = 56$, $p = 0$, and hit rate for the zero-item interval was greater than for the other intervals when tested by an orthogonal contrast, $t(7) = 11$, $p = 0$.

Also, a repeated-measures ANOVA was performed in which the factors were repetition interval (0, 2, 4, 8, or 16) and response (hit or false alarm), and the dependent measure was response time (RT). There was a significant effect of repetition interval, $F(4, 28) = 14$, $p = 0$, and a significant interaction between repetition interval and response, $F(4, 28) = 9$, $p = 0$, because hits were faster than false alarms only for the zero-item interval, 581 ms (95% confidence interval [CI], 547–616) versus 694 ms (CI, 630–718) for zero interval, 678 ms (CI, 631–721) versus 697 ms (CI, 639–752) for other intervals.

The zero-item interval was quite different from the other intervals. It was conceptually different in that a target was immediately repeated without any intervening items. It was perceptually different in that the immediate repetition of the target gave an on–off–on effect that did not occur for any other interval. It was empirically different in that hit rates were much higher and hit RT was much shorter. The hit rate for one repetition was 100% for four of eight subjects. To be certain that the zero interval was not the sole cause of significant differences between conditions, it was excluded from subsequent analyses.

To assess whether targets could be discriminated from distractors and the possible effects of criterion shifts across conditions, a repeated-measures ANOVA was performed in which day (six levels), interval (2, 4, 8, or 16), and response (hit versus false alarm) were the factors and response rate was the dependent factor. There was a significant effect of response, $F(1, 7) = 41$, $p = 0$, because the hit rate was higher than the false alarm rate. There was also a significant effect of interval, $F(3, 21) = 32$, $p = 0$, and a significant interaction between interval and response, $F(3, 21) = 19$, $p = 0$, because hit rate declined as repetition interval increased, but false alarm rate was unchanged. (The Pearson correlation between hit rate and d' was $r = .72$, and the rank order correlation was 1.0. When analyses reported for hit rate were also performed on d', the results were indistinguishable, so only the results for hit rate are reported.)

Finally, the variable of interest, the probability of a hit as a function of the number of repetitions of an unrecognized target, was assessed. A repeated-measures ANOVA was performed in which interval (2, 4, 8, or 16) and number of repetitions (1, 2, or 3) were the factors, and hit rate was the dependent variable. The main effect of interval, $F(3, 21) = 25$, $p = 0$, was significant, but the effect of repetition, $F(2, 14) = 1.3$, $p = .31$, and their interaction, $F(6, 42) = 1.0$, $p = .38$, were not significant. Figure 1 shows the probability of a hit plotted as a function of the number of repetitions for repetition intervals of 2, 4, 8, and 16. Hit rate decreased as a function of the interval between repetitions of the target. In fact, there was a significant linear trend for interval, $F(1, 7) = 47$, $p = 0$, in the analysis. As is shown in Figure 1, hit rate does not increase monotonically as a function of repetition. The hit rates for one, two, and three repetitions were .46, .52, and .50, respectively. There was not a significant linear trend for repetition, $F(1, 7) = 0.6$, $p = .46$. It is clear that the results are inconsistent with the prediction of the incremental hypothesis. The slopes of the repetition functions are not positive values that decrease as function of response interval. Instead, all four slopes are not far different from zero. In contrast, the results were generally consistent with the prediction of the all-or-none hypothesis. At the 95% confidence level, for the repetition intervals of 2, 4, 8, and 16, slopes of 0.085, 0.065, 0.135, and 0.07 could be detected. So the failure to find a monotonic increase in the probability of recall as a function of repetition cannot be attributed to a lack of power.

Another way to examine the data is to fit linear trend lines to each condition for each subject and look for stationarity around a slope of zero. A

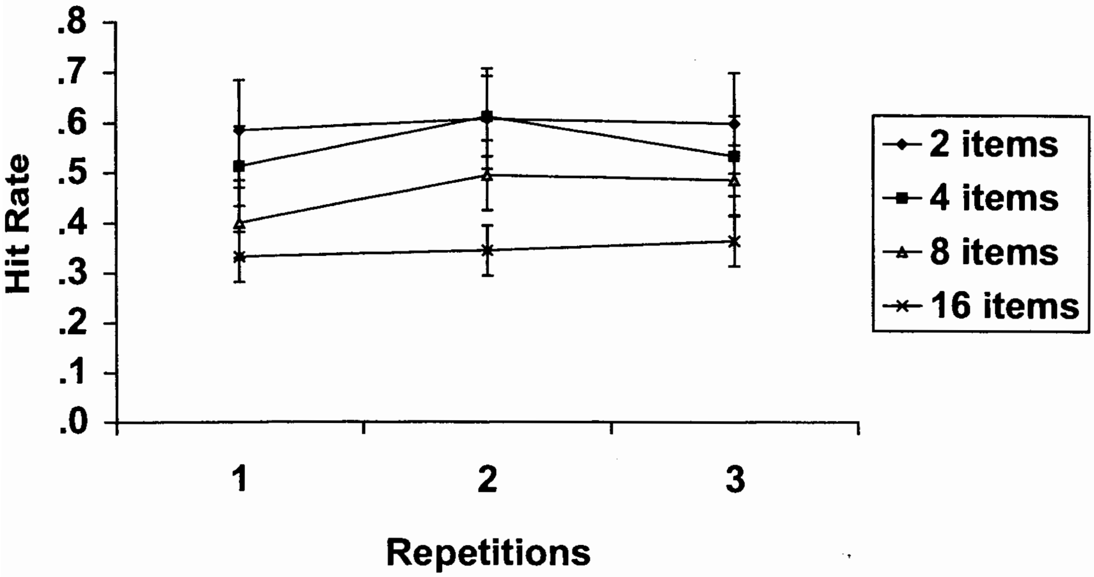

Figure 1. Probability of a hit as a function of the number of repetitions and repetition interval in Experiment 1; bars show 95% confidence intervals

detailed analysis of the data is provided by Table 1, which shows the slopes of all the functions for the first 2, middle 2, and last 2 days. As shown in Table 1, the slopes of the functions decrease with practice and on the last 2 days of the experiment show stationarity around zero. Thus, the weight of the evidence is consistent only with the all-or-none hypothesis.

DISCUSSION

The main result of this experiment was that the probability of a hit did not increase for a repeated, unrecognized trigram. The main effect of repetition was not significant, and there were no significant interactions between repetition and any other variable. When the data were tabulated separately for each subject, after 4 days of practice the median slope across subjects was 0 for interval 2, –0.03 for interval 4, 0.04 for interval 8, and –0.07 for interval 16. The only result inconsistent with the all-or-none hypothesis was the significant contrast between one and two repetitions even though there was not a significant contrast between one and three repetitions. Perhaps it should be emphasized how remarkable it is that a consonant trigram can be repeated three times at an intertarget interval of as little as 2.6 s, and yet the probability of detection of the target at three repetitions is no greater than for one repetition. With the exception of the results just cited, such a result is at variance with what has been assumed to be true throughout the modern history of psychology (Estes, 1960).

Table 1. Slopes of the linear trend lines for individual subjects in Experiment 1

		Subject								
Day	Interval	1	2	3	4	5	6	7	8	Median
1–2	2	.095	.125	–.06	.145	.150	.000	.035	.11	.065
1–2	4	.140	.145	–.095	.000	.140	.030	.06	–.2	.045
1–2	8	.055	.140	.020	.085	.130	.105	.035	.11	.095
1–2	16	.050	.005	–.11	.030	.055	.055	.06	.07	.052
3–4	2	.065	.010	.000	.020	.110	–.055	–.035	–.14	.005
3–4	4	.020	.045	.000	.090	.030	.000	.035	–.075	.025
3–4	8	.045	.045	.070	.235	.135	–.035	.1	.05	.048
3–4	16	.070	.050	.065	.115	.000	.085	.15	.035	.078
5–6	2	.205	–.125	.000	.000	.180	.000	–.035	.02	.000
5–6	4	.090	.000	–.07	.220	–.085	–.17	.1	.055	–.03
5–6	8	.055	.030	.000	.030	.115	.055	.0	–.035	.03
5–6	16	.075	–.005	–.10	–.10	–.11	.155	.035	–.025	–.015
Median		.067	.037	.000	.057	.082	.015	.035	.028	

Nevertheless, there are two possible objections to these results. The first objection is to the quality of the result. After all, the main result was the acceptance of the null hypothesis. Perhaps the difference between one and two repetitions would be significant with more observations. Perhaps a significant difference would emerge with more power or smaller confidence intervals. Perhaps the zero slope observed on days 5 and 6 would not remain at zero if the experiment had been continued longer. The second objection is to the interpretation of the results. Perhaps the zero slopes are the result of item selection effects. The easiest items are learned with one repetition, the next easiest items with two repetitions, and so on. A second experiment was done to respond to both of these objections. To respond to the first objection, subjects were given much practice in the second experiment. To respond to the second objection, trigrams were constructed from shapes that were all entirely novel and of equal complexity. So there was no a priori reason for any one trigram to be any more learnable than any other.

EXPERIMENT 2

Experiment 2 was again a repetition detection task. This time the items composing the trigrams were 20 novel patterns designed to have the same number of line segments as letters but to be entirely novel; they are shown in Figure 2. Notice that the patterns also are of equal complexity. They were designed to be of the same complexity as the consonants used in Experiment 1 without suggesting any particular consonant. As can be seen

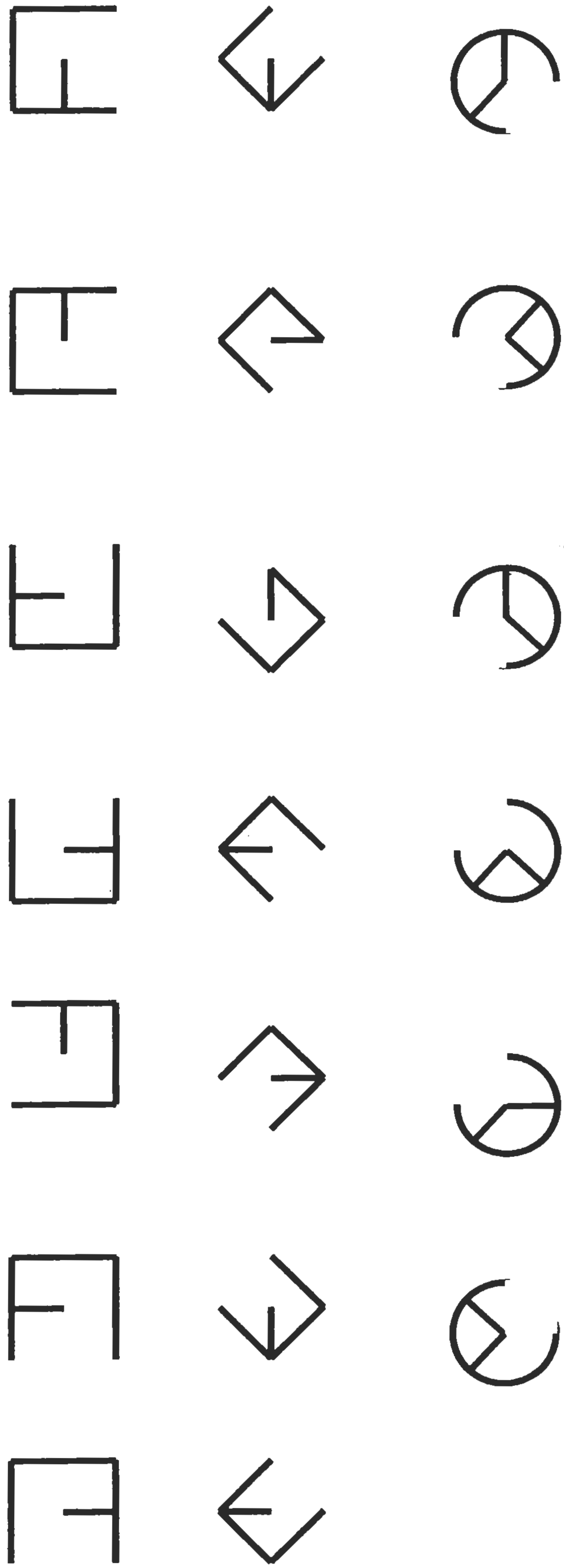

Figure 2. Novel shapes used to construct trigrams in Experiment 2

from the figure, initially a circle, diamond, or square was selected for generating the shape. Then one side was randomly deleted from the figure. Then a vertical or horizontal line was added from the center to a side for the square and diamond shapes. For the circle shape two lines were added from the center to the sides. This was because when the shapes were pretested, circles with only a single center line were found to be slightly more memorable than the squares and diamonds, but circles with the second line were not.

An SOA of 2,800 ms was used in Experiment 2, compared with 1,300 ms in Experiment 1, because the shape trigrams in Experiment 2 were more novel and therefore took more time to encode. In Experiment 2 only repetition intervals of 0, 1, and 2 were used. The same effects had already been shown at longer repetition intervals in Experiment 1. The use of long repetition intervals makes the task an extremely tedious vigilance task, necessitating a high degree of attention. It did not seem necessary to make the subjects suffer through long repetition intervals in this task as well. They would suffer in another way. Subjects in Experiment 2 were required to perform the task four times a day, 5 days a week, for 6 weeks. They had more practice on repetition detection at a particular interval at the end of 1 week than the subjects in Experiment 1 had in the entire experiment.

METHOD

Subjects

Four subjects (two men and two women) between the ages of 18 and 21 participated in the experiment. Three were native speakers of English, and one was an early (age 5) speaker of English who had lived since then in the United States and used it as her primary language. The subjects were volunteers taking an undergraduate honors course in memory. Nevertheless, they were naïve to the predictions of the incremental and all-or-one hypotheses before and during (and, in some cases, after) their participation in the experiment.

Design

During a block of trials a single target trigram was repeated in a sequence of distractor trigrams until a response to it was made. Each subject saw 1,800 blocks. Given the large number of trigrams per block, it was not possible for a subject to perform the entire task in a single day. The experiment took place over a 6-week period. For 5 days of the week each subject participated in four sessions a day, two usually before noon and the other two in the late afternoon or evening. During a session a subject was shown three-shape trigrams at the rate of one trigram per 2,800 ms on a computer monitor. The subject was instructed to press a button on the keyboard of the computer every time the trigram was a repetition of one that had appeared earlier. During a session a subject saw five block sequences that were presented as a continuous sequence.

The trigrams were constructed as follows. Each block sequence consisted of 15 blocks of trials. Retention intervals of 0, 1, and 2 intervening items were each used for five blocks in the block sequence. Each subsequence of three blocks contained each of the three retention intervals, but the order in which the retention intervals occurred was random. From the subject's point of view the task was a continuous recognition task. However, each block of trials actually consisted of a study sequence followed by as many test sequences as were necessary for the subject to recognize the target. At the beginning of the study sequence, one to five randomly generated shape trigrams were presented. Then the target trigram was randomly generated and presented. The presentation of the target trigram concluded the study sequence. Next, for an interval of m, m randomly generated shape trigrams were presented, and then the target trigram was re-presented. For example, if $m = 2$, then two randomly generated shape trigrams were presented before the target trigram was repeated the first time. If the subject responded with a button press indicating that the target trigram was a repetition, this ended the block. Otherwise, if the subject did not respond, another m randomly generated shape trigrams followed by the target trigram were presented. Again, if the subject responded with a button press indicating that the target trigram was a repetition, this ended the block. Otherwise, another m randomly generated shape trigrams were presented, followed by another repetition of the target. As mentioned earlier, m could be 0, 1, or 2.

Trigrams were generated without replacement. That is, except for target repetitions, each trigram appeared no more than once during a session. In generating a trigram, the shapes were selected without replacement, so a trigram consisted of three different shapes.

Procedure

A subject began each block sequence by typing a code number that identified the subject and the block sequence. Next, the instructions appeared. Each subject was instructed to read the directions before his or her first time in the experiment. When the subject pressed any key, the experiment immediately began; the first shape trigram was presented for 2,600 ms, followed by three underline symbols in the locations of the shapes for 200 ms, bringing the total SOA to 2,800 ms. Then the next trigram was presented. All subjects responded to a repetition by pressing the spacebar. Because the size of each study sequence and the order of the intervals were randomly selected, a subject could not predict when a target repetition would occur.

RESULTS

The hit rate and false alarm rate for each subject were computed. The false alarm rate was .15 and did not vary across repetition interval. A repeated-measures ANOVA was performed in which the independent measure was repetition interval (0, 1, or 2) and the dependent measure was hit rate for a single repetition. Hit rate declined as repetition interval increased, $F(2, 6) = 93$, $p = 0$, and hit rate for the zero-item interval was greater than for the other intervals when tested by orthogonal contrasts, $t(3) = 23$, $p = 0$.

Also, a repeated-measures ANOVA was performed in which the factors were repetition interval (0, 1, or 2) and response (hit or false alarm), and the dependent measure was RT. There was a significant effect of repetition interval, $F(2, 6) = 75$, $p = 0$, and a significant interaction between repetition interval and response, $F(2, 6) = 26$, $p = .001$, because hits were significantly faster than false alarms only for the zero-item interval, 901 ms (CI, 534–1,267) versus 1,550 ms (CI, 1,355–1,744) for zero interval, 1,405 ms (CI, 1,232–1,580) versus 1,566 ms (CI, 1,372–1,760) for other intervals).

The zero-item interval was very different from the other intervals. It was conceptually different in that a target was immediately repeated without any intervening items. It was perceptually different in that the immediate repetition of the target gave an on–off–on effect that did not occur for any other interval. It was empirically different in that hit rates were much higher and hit RT was much shorter. To be certain that the zero interval was not the sole cause of significant differences between conditions, it was excluded from subsequent analyses.

To assess whether targets could be discriminated from distractors and the possible effects of criterion shifts across conditions, a repeated-measures ANOVA was performed in which week (six levels), interval (1 or 2), and response (hit versus false alarm) were the factors, and response rate was the dependent factor. There was a significant effect of response, $F(1, 3) = 16$, $p = .027$, because the hit rate was higher than the false alarm rate.

Finally, the variable of interest, the probability of a hit as a function of the number of repetitions of an unrecognized target, was assessed. A repeated-measures ANOVA was performed in which interval (1 or 2) and number of repetitions (1, 2, 3, 4, or 5) were the factors, and hit rate was the dependent variable. The main effects of interval, $F(1, 3) = 2.3$, $p = .223$, and the effect of repetition, $F(4, 12) = 0.6$, $p = .677$, were not significant. Figure 3 shows the probability of a hit plotted as a function of repetition interval and the number of repetitions. As shown in Figure 3, hit rate does not increase monotonically as a function of repetition. There was not a significant linear trend for repetition, $F(1, 3) = 0.5$, $p = .525$. It is clear that the results are inconsistent with the prediction of the incremental hypothesis. The slopes of the repetition functions are not positive values that decrease as function of response interval. Both slopes are near zero. In contrast, the results were generally consistent with the prediction of the all-or-none hypothesis. Because there were only four subjects but each subject was tested for many days, a power analysis was done for each subject. At the 95% confidence level, for each of the four sub-

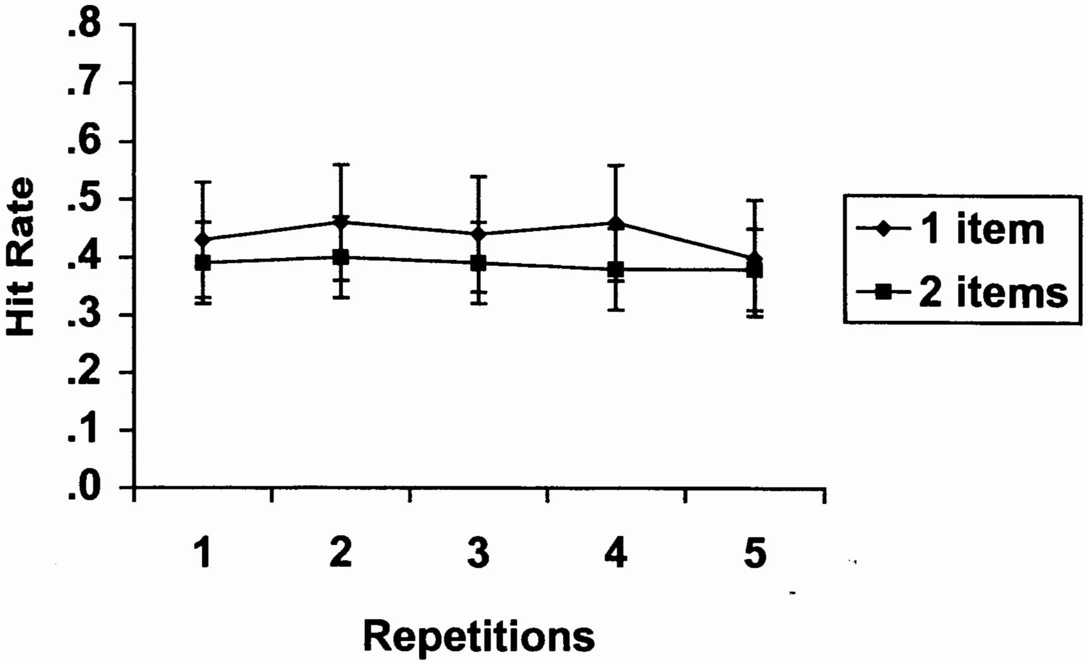

Figure 3. Probability of a hit as a function of the number of repetitions and repetition interval in Experiment 2

jects, at repetition intervals of one and two, the slopes that could be detected are .064 and .063, .053 and .011, .012 and .07, and .011 and .048, respectively.

Another way to examine the data is to fit linear trend lines to each condition for each subject and look for stationarity around a slope of zero. As shown in Table 2, the slopes of the functions for the four subjects at both intervals are trivially different from zero. Thus, the weight of the evidence is consistent only with the all-or-none hypothesis.

DISCUSSION

The main result of this experiment was that the probability of a hit did not increase for a repeated, unrecognized trigram. The main effect of repetition was not significant, and there were no significant interactions between repetition and any other variable. Again it should be emphasized

Table 2. Slopes of the linear trend lines for individual subjects in Experiment 2

	Subject				
Interval	1	2	3	4	Median
1	.002	–.019	.025	–.029	–.085
2	–.003	–.022	.145	.01	.035

how remarkable it is that a shape trigram can be repeated five times at an intertarget interval of as little as 2.8 s and yet the probability of detection of the target at five repetitions is no greater than for one repetition.

The item selection explanation for repetition detection failure

The zero slope for the probability of repetition detection has been interpreted as support for the hypothesis that no trace of the target remained from the preceding trial when detection failed. There is an alternative hypothesis that predicts this result. Suppose that the trigrams differed in learnability so that the most learnable required only one presentation to encode, the next most learnable required two presentations to encode, and so on. If the distribution of learnability across study items was such as to produce a zero function, then the zero slope could be attributed to the effect of learnability and the concept of incremental learning could be preserved. Here, the history of the learnability hypothesis is reviewed. It will be seen that although it was a valid criticism of the experiments to which it was originally applied, it cannot be plausibly extended to subsequent experimental results, including the results reported here.

The relationship of the all-or-none hypothesis to the incremental hypothesis is of a special case to a more general model. The all-or-none hypothesis is that on a study trial an item is learned or it is not. The degree of learning can be represented by a one or a zero. In comparison, the incremental hypothesis is that on a study trial the degree of learning can be represented by a continuous variable, that is, a number between one and zero. Because the all-or-none hypothesis is a special case of the incremental hypothesis, any data described by the all-or-none hypothesis can also be described by the incremental hypothesis. On the other hand, if all data can be described by the all-or-none hypothesis, then there is no basis for assuming that the more complicated incremental hypothesis is true.

The relationship between the all-or-none hypothesis and the learnability hypothesis as explanations for an experiment in which a zero learning slope is found is that between two logically possible alternative hypotheses that both predict the same set of data. Because both hypotheses predict the same data, these data cannot be used to choose between the hypotheses. Rather, all data must be considered that render one or the other of the hypotheses more plausible in that context. Therefore, let us consider the results of other studies on all-or-none learning and the principal objection to them.

Underwood et al. (1962) concluded first that learnability effects occurred in Rock's experiments and second that his results were artifactual and did not provide evidence of all-or-none learning. On both these points their work is conclusive, and neither of these conclusions will be disputed here. However, notice that it does not follow logically from these two points that learnability effects were the artifact that produced the equivalence of the experimental (replacement) condition with the control condition. This final point will be considered here. To this end, the experiments of Underwood et al. that used Rock's materials will be considered.

To demonstrate learnability effects, Underwood et al. (1962) had two control groups learn letter–number pairs with the control procedure. Control One (C_1) was identical to Rock's (1957) control group. It was given the same list on the first block as the experimental group (E) who learned with the replacement procedure. In contrast, Control Two (C_2) was tested after the other groups. Each subject in C_2 was given a list composed of the 12 paired associates eventually learned to criterion by a subject in E. Presumably, this list would contain easier items if easier-to-learn items tended to replace more difficult ones during list learning with the replacement procedure. If Rock's results were caused by learnability effects, then Underwood et al. (1962) should have replicated Rock's finding that the replacement list was learned as fast as the control list ($E = C_1$) but should have found that both the replacement list and first control list took longer to learn than the second control list made up of easier items ($E = C_1 > C_2$). In fact, in their first replication of Rock (Experiment IIA), Underwood et al. found E to be nonsignificantly different from C_1 and C_1 to be nonsignificantly different from C_2. These results replicated Rock's and were consistent with the all-or-none hypothesis. Furthermore, the failure to find a significant difference between C_1 and C_2 was inconsistent with a learnability effect. Of course, Underwood et al. noticed and acknowledged the problem (p. 362): "The puzzle that remained was why, if there is consistent item-selectivity, the Ss in C_2 did not learn more rapidly than those in C_1 in Experiment IIA."

They provided an answer to their own question by noting the second flaw in Rock's procedure as well as in their first replication (p. 362): "The slow rate of presentation and the 30 sec between successive trials during which time most Ss obviously rehearsed, are probably responsible."

In their remaining three experiments they controlled rehearsal through faster presentation and the use of a distractor task. The results are shown in Table 3, which shows the relevant experiments of Rock (1957), Rock and Heimer (1959), Underwood et al. (1962), and two additional similar experiments by Postman (1962), who also insightfully suggested that controlling rehearsal inhibited item selection. In the four experiments in which rehearsal was not controlled, $E = C_1$, but in the six experiments in which rehearsal was controlled, $E > C_1$. These results can be explained by assuming that when rehearsal was uncontrolled subjects

Table 3. Results of Rock (1957) (R), Rock and Heimer (1959) (R & H), Underwood et al. (1962) (U), and Postman (1962) (P)

Experiment	Rehearsal controlled	E vs. C_1	C_1 vs. C_2
R, 1	No	$E = C_1$	
R, 2	No	$E = C_1$	
R & H, 1	No	$E = C_1$	
U, 2A	No	$E = C_1$	$C_1 = C_2$
U, 3	No	$E = C_1$	$C_1 = C_2$
U, 4	Yes	$E > C_1$	$C_1 = C_2$
U, 5	Yes	$E > C_1$	$C_1 = C_2$
P, 1	Yes	$E > C_1$	$C_1 > C_2$
P, 2	Yes	$E > C_1$	$C_1 = C_2$

C_1 = control group 1; C_2 = control group 2; E = experimental group.

devoted their rehearsals to the new items in the replacement procedure, thus making learning in the replacement procedure equivalent to that of the control procedure. However, when rehearsal was controlled so that more rehearsals could not be devoted to the new items in the replacement procedure, learning was significantly slower in the replacement procedure than in the control procedure. The difference in learning between the replacement procedure and the control procedure demonstrates that Rock's results were an artifact of the subjects' ability to selectively rehearse new items in the replacement procedure. It also demonstrates that the item selection artifact that produced Rock's results was not an effect of replacing more difficult with easier items but rather the result of allowing the subjects to devote most of their rehearsals to the new items. Further evidence of the limited role of learnability effects is that in five of six experiments the difference between C_1 (with easy items) and C_2 (with difficult items) was not significant (although Underwood et al. found a significant difference when they pooled experiments in a meta-analysis). Recent research confirms that when rehearsal is controlled, learnability effects are eliminated. Hirshman, Fisher, Henthorn, Arndt, and Passannante (2002) eliminated the normally higher hit rate for low-frequency compared with high-frequency words in recognition by increasing the rate of presentation.

Three subsequent studies of all-or-none learning (besides this one) all addressed both item selection flaws in the original studies by controlling rehearsal and using items of equal learnability. Nelson and Batchelder (1969) made the replacement–repetition comparison in a distractor task (Peterson & Peterson, 1959), in which subjects had to recall consonant trigrams. The associative strength of the trigrams was held constant. Furthermore, Conrad (1964) and Wickelgren (1966) showed that encoding of consonant trigrams is primarily phonemic, and there was no evidence that phonemic differences between trigrams influenced their recall in their study. Of course, any more than a single rehearsal of the target was suppressed by the distractor task. Nelson (1971) made the replacement–repetition comparison for individual digits in Hebb's (1961) repeated string task. All the digit strings they used were of equal difficulty, and again rehearsal was controlled. Nevertheless, the studies found all-or-none learning for the study materials.

The logic of Murdock and Babick's (1961) experiment was similar to the one reported here. A study list containing a critical word was repeatedly presented until the critical word was recalled. The probability of recalling the critical word was constant as a function of number of presentations.

The critical words varied over a wide range of learnability, as indicated by the probability of recalling a word on its first presentation. When Murdock and Babick (1961) took account of the different levels of probability of initial recall in averaging over repeated presentations, their results did not change. In the two experiments reported here, rehearsal was controlled through the rapid presentation of study items and learnability was controlled through the use of similar, homogeneous, novel items of equal complexity.

The zero slope found by Murdock and Babick (1961) and us when retrieval is plotted as a function of repeated presentation puts a strong constraint on the learnability hypothesis. For the learnability hypothesis to predict a zero slope, when the proportion of items at a given level of learnability is plotted as a function of learnability, the function must be a decreasing exponential function described by the equation $y = e^{-cx}$, where x is a level of learnability, e is the base of a natural logarithm, c is a constant, and y is the proportion of items at that level of learnability. That is, the highest value of learnability is associated with the greatest number of items in the set. However, when we consider all the factors known to affect the learnability of words and nonsense syllables (e.g., frequency, imagery, meaningfulness, and concreteness), for none of them does the frequency distribution of the items follow a decreasing exponential function (Toglia & Battig, 1978). Most items are in the middle of the distribution, with modest differences between them. This distribution of learnability factors does not predict a zero slope.

Thus, it has not been empirically shown that demonstrations of all-or-none learning are item selection artifacts. Item selection is merely a logical alternative interpretation of evidence of all-or-none learning. Because the all-or-none hypothesis is a special case of the incremental hypothesis and item selection assumes incremental learning, there will always be a version of the item selection hypothesis that describes the results. A more general model can always mimic its special case. However, outside the fanciful world of memory theory it is usually the simpler hypothesis that is preferred.

GENERAL DISCUSSION

In a repetition detection task, a subject has no reason to rehearse one study item more than another. Furthermore, the rapid rate of presentation forces the subject to study each item as it appears. Under these conditions, there is no evidence of incremental learning.

The result reported here is similar to the repetition blindness effect observed at much shorter SOAs. In that paradigm (Kanwisher, 1987; Mozer, 1989; Bavelier & Potter, 1992) a sequence of letters or words is presented at the rate of 150 ms per item or higher, and the subject is asked to detect the repeated item. Subjects often fail to notice the second occurrence of the repeated item. Kanwisher attributed the recognition failures she observed to an exotic perceptual mechanism. However, just as much repetition blindness was observed here at an SOA of 2,800 ms, so *repetition blindness* is simply another name for ordinary recognition failure, whether as the result of encoding or retrieval failure. All that is necessary to achieve a high rate of recognition failure is the suppression of rehearsal.

Even if the all-or-none hypothesis is rejected, the data presented here in support of it remain to be explained. At a minimum they present a challenge to the incremental hypothesis.

Notes

This work was begun while the first author was a Fulbright Scholar at the University of Oslo and was supported by the University of Oslo.

Correspondence about this article should be addressed to Arnold L. Glass, Psychology Department, Rutgers University, New Brunswick, NJ 08903 (e-mail: aglass@rci.rutgers.edu). Received for publication May 7, 2003; revision received November 20, 2003.

References

Bavelier, D., & Potter, M. C. (1992). Visual and phonological codes in repetition blindness. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 134–147.

Conrad, R. (1964). Order errors in immediate recall of sequences. *British Journal of Psychology, 55,* 75–84.

Estes, W. K. (1960). Learning theory and the new "mental chemistry." *Psychological Review, 67,* 207–223.

Glass, A. L., Krejeci, J., & Goldman, J. (1989). The necessary and sufficient conditions for motor learning, recognition, and recall. *Journal of Memory and Language, 28,* 189–199.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning.* London: Blackwell.

Hirshman, E., Fisher, J., Henthorn, T., Arndt, J., & Passannante, A. (2002). Midazolam amnesia and dual-process models of the word-frequency mirror effect. *Journal of Memory and Language, 47,* 499–516.

Kanwisher, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition, 27,* 117–143.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

Mozer, M. (1989). Types and tokens in visual letter perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 287–303.

Murdock, B. B., & Babick, A. J. (1961). The effect of repetition on the retention of individual words. *American Journal of Psychology, 74,* 596–601.

Nelson, T. O. (1971). Transfer of serial information from short-term to long-term memory: Toward the locus of the repetition effect. *Psychonomic Science, 25,* 344–345.

Nelson, T. O., & Batchelder, W. H. (1969). Forgetting in short-term recall: All-or-none or decremental? *Journal of Experimental Psychology, 82,* 96–106.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Postman, L. (1962). Repetition and paired-associate learning. *American Journal of Psychology, 75,* 372–389.

Rock, I. (1957). The role of repetition in associative learning. *American Journal of Psychology, 70,* 186–193.

Rock, I., & Heimer, W. (1959). Further evidence of one-trial associative learning. *American Journal of Psychology, 72,* 1–16.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

Underwood, B. J., Rehula, R., & Keppel, G. (1962). Item-selection in paired-associate learning. *American Journal of Psychology, 75,* 353–371.

Wickelgren, W. A. (1966). Similarity and interference in short-term memory for single letters. *Journal of Experimental Psychology, 71,* 396–404.