Journal of Experimental Psychology:
Human Learning and Memory
1979, Vol. 5, No. 3, 241-252

Credibility of Repeated Statements: Memory for Trivia

Frederick T. Bacon
McMaster University, Hamilton, Canada

Previous research has shown that repeated statements are rated as more true than new ones. Experiment 1 found that the repetition effect depends on subjects' detection of the fact that a statement is repeated; that is, statements that are judged to be repeated are rated as truer than statements judged to be new, regardless of the actual status of the statements. Experiment 2 determined that repeated statements increment in credibility even if subjects are informed that they are repeated. It was further determined that statements that contradict early ones are rated as relatively true if misclassified as repetitions but that statements judged to be changed are rated as relatively false. That is, subjects are predisposed to believe statements that seem to reaffirm existing knowledge and to disbelieve statements that contradict existing knowledge.

Although much conventional wisdom fails to survive direct experimental tests, some widely held views fare quite well. One common opinion is the rather cynical belief that

> if statements are made again and again in a confident manner, then their hearers will tend to believe them quite independently of their soundness and of the presence or absence of evidence for their truth. (Thouless, 1974, p. 111)

The tendency to believe statements on the basis of repetition rather than evidence is not logically founded; indeed, Wittgenstein ridiculed the tendency by comparing it to buying a second newspaper to determine whether the first one was correct (Kenny, 1973). On the other hand, as Russell (1973) observed, it is often difficult, if not impossible, to obtain direct evidence regarding the truth of referential statements. In such cases, testimony from others is a reasonable source of information regarding probable truth, although such testimony is not direct evidence in itself. It is not unreasonable, therefore, for hearers to place more stock in statements that they have encountered previously than in statements that they are encountering for the first time. The important point for present purposes is that the assumed relation between repetition and belief stands up to experimental scrutiny.

In a recent experiment, Hasher, Goldstein, and Toppino (1977) presented sets of plausible but probably unfamiliar statements to subjects for truth ratings. Some statements were repeated from session to session, along with other statements that had not been presented previously. Two results supported the conclusion that repetition influences belief. First, within each test session, repeated statements were rated as more true than new statements. Second, across test sessions, repeated statements received higher truth ratings on later sessions than on earlier sessions. Further, the positive effects of repetition on belief were obtained for statements that were actually false as well as for statements that were true.

This research was supported by Grant A8122 from the National Research Council of Canada to Ian Begg. F. T. Bacon is a research group whose members, Ian Begg, Grant Harris, John Mitterer, and Douglas Upfold, collaborated equally in the design, analysis, and writing of this research. This research was presented at the Midwestern Psychological Association meeting of May, 1978. We thank Carolyn Hopkins for her help and encouragement.

Requests for reprints should be sent to F. T. Bacon, c/o Ian Begg, Department of Psychology, McMaster University, Hamilton, Ontario, Canada L8S 4K1.

 0096-1515/79/0503-0241$00.75

The purpose of the present article is to offer a theoretical account of the repetition effect and to assess the adequacy of the account against several alternative accounts. It is important to get some theoretical grasp of the phenomenon, since it offers a bridge between social psychologists, who have long been concerned with factors influencing credibility (e.g., Asch,1940, 1948; Hovland, Janis, & Kelley, 1953), and cognitive psychologists, who have been equally concerned with the effects of repetition (e.g., Begg, 1974; Howell, 1973). Let us begin by considering the account offered by Hasher et al. (1977) to explain their findings. They posited that probable truth is inferred directly from event frequency, which serves as a "criterion of certitude" (p. 112). Frequency, according to Underwood's (1969) attribute theory, is directly encoded in a frequency attribute that is specifically sensitive to event frequency and that accrues automatically with repetition (cf. Hasher & Chromiak, 1977). Accordingly, as a repeated statement accrues frequency, its probable truth also increases in parallel fashion.

Our account is in agreement with Hasher et al.'s (1977) account in postulating that truth is inferred from some memorially based criterion of certitude. In the present account, however, that critical function is served directly by the subject's memory judgment concerning the test statement being rated, without the mediated and redundant process of estimating the frequency with which the statement has occurred. Whereas Underwood (1969) postulated the existence of a frequency attribute that directly encodes event frequency, several other theories assume that frequency is not directly encoded but may be inferred from other memory information accessed by the test item, such as the number of traces (Hintzman & Block, 1971), the number of list markers (Anderson & Bower, 1972), or the availability of any trace (Tversky & Kahneman, 1973). For example, Tversky and Kahneman reported that subjects are likely to judge incorrectly that there are more English words in which the first letter is K than in which the third letter is K because they can more readily generate instances of the former than the latter. If frequency is itself an inference from memorability of a test statement, it would be inefficient for subjects to estimate frequency as a basis for assigning truth ratings. Rather, probable truth, like frequency, can be directly inferred from the recognition judgment regarding the statement being rated. If repeated statements are more likely than new ones to be judged as old, they should be rated as more true. Similarly, if recognizability increases with frequency of exposure, truth ratings should increment with frequency of exposure.

Both Hasher et al.'s (1977) account that truth ratings are mediated by a frequency attribute and the present account that truth ratings are inferred from recognition decisions are capable of accounting for the basic repetition effect. The accounts, however, lead to different expectations concerning the influence of recognition on truth ratings. According to the frequency account, the frequency attribute accrues automatically with repetition. Truth ratings, therefore, should be more sensitive to an item's actual status as old or new than to the subject's judgment that the item is old or new. On the other hand, if truth ratings are inferred from recognition decisions, judged rather than actual oldness should be associated with increments in belief. Repeated statements should only become more credible with repetition if that repetition is detected; therefore, repeated statements judged to be new should not be rated as more true than on a prior rating, whereas repetitions judged to be old should increment in rated truth. Equivalently, if new statements are incorrectly judged to be old, they should be rated as more true than new statements correctly judged to be new. Therefore, by conditionalizing truth ratings on recognition judgments, it should be possible to contrast the merits of the different theories.

A possibility not accounted for by either the attribute theory or the recognition theory is that subjects might not change their rating for repeated statements that they recognize as being merely repeated, as implied in Thouless' (1974) comment that an orator

might be "using the method of repetition although he repeats himself in different words . . . partly to conceal the method used" (p. 113). Given the basic repetition effect, then, the locus of the effect should be restricted to repeated statements that are incorrectly judged to be new. Thus, if repeated statements convey information that is frequent or familiar, they will increment in rated truth, but if the statement is recognized as being a repetition of a statement given earlier, the truth ratings will not change. Such an account raises questions regarding whether the repetition effect is explainable on the basis of memory for prior ratings or on the basis of subjects' familiarity with the factual information conveyed by the statements. The former possibility seems unlikely based on Hasher et al.'s (1977) results, since the repeated statements were rated systematically truer as frequency increased. However, if Thouless' (1974) comment is founded, it could be the case that subjects remember and use again the prior ratings for recognized statements, with inflated truth estimates being applied to the repeated statements not recognized as repetitions. The critical test of the hypothesis pertains to new statements, since they have no prior ratings to retain; if such statements show differential ratings for those judged to be new versus those judged to be old, the hypothesis can be rejected.

The latter possibility that truth ratings are inferred from familiarity with the content of the statements rather than from recognition decisions regarding the oldness of the statements can be quite plausible. Suppose that in the initial session, the information conveyed by the statements has some probability of being incorporated into a subject's knowledge (with a higher probability for more plausible statements). Repeated statements, then, would be more likely than new ones to convey known information, acquired during the initial session, and would accordingly be rated as relatively true. In agreement with the present account, such a semantic account would lead us to expect different truth ratings for repeated statements judged to be new versus old. However, since the recognition decisions and truth ratings depend on familiarity with content and since that familiarity depends on initial plausibility, recognized and missed repetitions should differ in initial plausibility as well as in later ratings of truth. Again the theoretical difference leads to a direct experimental test.

In all, there are four viable accounts of the repetition effect, each of which leads to testable outcomes regarding the relation between recognition judgments and truth ratings. If ratings are inferred from recognition judgments, then statements that are judged to be old should be rated as more true than those that are judged to be new, regardless of the correctness of the judgment. If ratings are inferred from a frequency attribute, repeated statements should be rated as relatively true regardless of whether the statements are recognized or missed. If ratings are inferred from familiarity with semantic content, recognized and missed statements should receive different truth ratings, but the difference should reflect initial differences in plausibility rather than differences in recognition decisions or repetition per se. Finally, if subjects are skeptical, then statements that are recognized as being merely repetitions will not increment in rated truth, with the entire repetition effect stemming from increased ratings for repetitions incorrectly judged to be new. Experiment 1 assesses the merits of these alternative accounts.

Experiment 1 [1]

The purpose of the first experiment was to distinguish among the alternative explanations of the repetition effect. The basic procedure consisted of having subjects rate sentences for truth on two occasions, 3 weeks apart. Half of the statements on each occasion were true, half of them were false. Half of the statements rated in the second session were repetitions of statements presented earlier, whereas half were new statements. On that second test, some subjects made

[1] The research in this article is presented in a two-experiment format for clarity of exposition, although all results were collected for a single class on 2 test days.

recognition judgments prior to assigning truth ratings. The critical results of interest concern the relation between recognition and rated truth. Before reiterating the predictions of interest, several procedural differences between the present experiment and the experiment described by Hasher et al. (1977) will be mentioned.

First, Hasher et al. chose statements that they judged to be plausible but unfamiliar, and they constructed a complementary set of false statements. The rationale for such selection was to allow study of the effects of repetition uncontaminated by individual differences in knowledge or by variation in the plausibility of the statements. Their judgment of plausibility was reasonable, since the statements were, on the average, rated as moderately plausible. Nonetheless, we felt that a more objective assessment of the plausibility of both true and false statements was desirable, along with a more objective basis for determining whether the statements were already familiar to the subject population. Accordingly, we initially chose statements from a trivia book (McKenzie, 1976) and generated a false version of each statement. Subjects then sorted subsets of the statements into true and false categories. Statements were only retained as experimental material if two or three of five raters classified the true and false versions as true. Thus, we have some reason to believe that the statements are neither totally plausible nor implausible and that the factual content of the statements is not widely known. Additionally, both the true and false versions are guaranteed to be about equally plausible.

Second, Hasher et al. (1977) included 60 statements in each session, 20 of which were repeated. In our first session, 90 statements were presented, of which the middle 60 were repeated later along with 60 new statements that were also randomly chosen from the pool. Thus, in the present experiment, half rather than a third of later statements were in fact repeated statements. Third, Hasher et al. had three sessions with 2 weeks between sessions, whereas the present experiment only involved two sessions separated by 3 weeks. Fourth, the truth ratings collected by Hasher et al. were on a 7-point scale, with 7 being true; in the present case, the 7-point scale was reversed, with 1 being true. In sum, all of the procedural variations are minor but add to the generality of the repetition effect if indeed it obtains under present conditions.

One condition in the present experiment, a control condition, simply involved having subjects provide truth ratings on each test occasion. In this condition, we expected that the basic repetition effect reported by Hasher et al. would be replicated. That is, the 60 statements that are presented on each test should be rated as more true on the second occasion than on the first and also as more true on the second test than the 60 new statements on the same test. The critical condition in the present experiment, in which recognition decisions precede truth ratings, should also replicate the basic effect, if it is analyzed without regard to recognition judgments.

The most important contrasts in the experiment concern truth ratings that are conditionalized on recognition judgments. Several theoretically relevant outcomes are possible. First, if the difference in rated truth is larger between actually repeated versus new statements than between statements judged to be old versus those judged to be new, the results would support Hasher et al.'s (1977) interpretation that a frequency attribute serves as a criterion of certitude. Second, given the opposite difference, the results would support the present account that truth is inferred from the recognition judgment. Especially important are new statements: if new statements that are judged to be old are rated as more true than those that are judged to be new, that difference cannot be attributed to a frequency attribute, since the statements have no prior occurrences in the task. Third, if repeated statements that are correctly identified as old are rated as no more true than on their first rating, the results would support a social "suspicious subject" explanation. The final contrast of interest concerns the possibility that both recognition and truth ratings depend on initial plausibility, in which event

recognized and missed repetitions would differ not only in final truth ratings but also in initial ratings.

Method

Subjects

Ninety-eight students who were enrolled in a second-year course in personality participated in the experiment, with 35 in the replication condition, 32 in one recognition condition, and 31 in another recognition condition.

Materials

Item selection. The goal of item selection was to produce a reasonably large pool of pairs of plausible statements, with one member of each pair true and the other false. Initially, about 1,000 statements were selected from a published source of unrelated facts (McKenzie, 1976). For each statement, a false version was prepared by altering some detail of the original such that the two statements could not both be true. Each version of each statement was recorded on an index card. The cards were randomly sorted into 10 decks of 200 statements, with the restrictions that each deck contain an equal number of true and false statements and that no deck contain both versions of any particular statement. Each deck was given to five volunteers who sorted the statements into two groups, true and false. On the basis of the sorting, statement pairs were retained if both the true and false versions were judged to be true by two or three of the five sorters. In this fashion, a set of 195 pairs of statements was selected, with both members about equally plausible but with no statements extremely believable or unbelievable.

Final refinement of the pool consisted of sorting the statements into five very general categories (animal kingdom, general science, U.S. facts, inventions and firsts, and people). We were able to reach consensus on at least 24 statements from each category. By discarding at random, the final pool was reached, consisting of 120 statement pairs with 24 statements in each of the five categories. In each list to be described, equal numbers of true and false statements from each of the five categories were included. However, since the categories were equivalent in all performance measures, categorization will not be discussed further, except to mention that our results cannot simply reflect selection artifacts.

List construction. Four lists were prepared, consisting of one initial list common to all conditions and three different test lists. The initial list contained 90 statements, 30 of which were fillers; the first 14 and last 16 statements were fillers with half of each of these blocks consisting of true statements. The 60 critical items, half of which were true, occupied the middle 60 positions in the list. All analyses concerned the critical items only, with fillers ignored.

The three test lists all consisted of the same 120 statements; 60 were repetitions of the 60 critical items and 60 were new, again with half of the statements being true and with equal numbers from each general category. The set of 120 statements was used to generate three test conditions. Condition 1 was the replication condition, with the 120 statements typed in random order with a blank space to the left of each one. Condition 2 was a recognition condition, identical to Condition 1 except that the words *yes* and *no* were typed beside each blank. Condition 3 was another recognition condition, differing from Condition 2 only in that the 24 test items from each category appeared in a block, preceded by the category label (As mentioned above, categorization produced no effects; therefore, Condition 3 can simply be treated as a replication of Condition 2.)

Procedure

The 98 students rated the validity of each statement in the initial list, using a 7-point scale ranging from certainly true (1) to certainly false (7). The statements were read aloud, to assure a standard acquisition session, at the rate of about three statements/min, with students writing a numeral from 1 to 7 next to the appropriate statement number on an answer sheet. The second test was administered 22 days later, again to an intact group. Booklets for the three test conditions were distributed. Each booklet contained the statements and instructions that informed the students of the number of true and false statements, reminded them of the nature of the 7-point scale, and instructed those who were in the recognition conditions to circle yes if they felt a statement was repeated from the initial test and no if it was new. Testing was self-paced.

Results and Discussion [2]

In preliminary analyses, it turned out that the actually true and false statements received virtually identical truth ratings, which should be expected because of the manner in which the statements were selected in the first place. It also turned out that the five general categories were equally plausible. Accordingly, in all analyses that follow, performance measures were averaged across the categories and across the actual truth of the

[2] Throughout the article, alpha was set equal to .05 for all inferential tests.

Table 1
Mean Truth Ratings of New and Repeated Statements

Statement	Condition 1	Condition 2	Condition 3
Initial test	3.80	3.75	3.64
Second test			
Repeated	3.63	3.63	3.46
New	3.98	3.90	3.92
M	3.81	3.76	3.69

Note. Lower ratings are more true than higher ratings; standard deviations range from .27 to .54.

statements. It should be noted, however, that each effect obtained in each condition was present in all 10 sets of items.

The repetition effect. The first analysis was simply to determine whether Hasher et al.'s (1977) basic repetition effect replicated. Mean truth ratings are presented in Table 1 with the replication condition in the left-hand column. The mean score for the initial test pertains to the 60 critical items. Those items were rated as more true on their second occurrence ($M = 3.63$) than on their first ($M = 3.80$), $t(34) = 3.88$, and also received higher ratings than the 60 new items present on the same test, ($M = 3.98$) $t(34) = 6.33$. Thus, the basic repetition effect was replicated. Note also that the same pattern was obtained in Conditions 2 and 3, with the 60 repeated statements on the second test rated as more true than either the new statements on that test or the same critical items on the first test. Thus, interpolating a recognition judgment between the presentation of items on the second test and ratings of their truth did not remove the basic effect of interest. The final point to note is that despite some differences in mean ratings for the three different groups of subjects, the overall mean ratings for the 60 critical items on the initial test were virtually identical to the mean ratings for all 120 items on the second test, indicating that no change in subjects' bias to believe or doubt the statements occurred over the test interval.

Recognition. The point of the second analysis was simply to determine how accurately subjects judged items to be either repeated or new. Recognition judgments were generally quite accurate. Repeated statements were correctly judged as old with hit rates of .733 ($SD = .09$) and .698 ($SD = .12$) for the two conditions. New statements were correctly classified as new with correct rejection rates of .855 ($SD = .11$) and .854 ($SD = .07$). Despite the reasonably good memory shown, enough new statements were judged to be both new and old, and enough repeated statements were judged to be new and old to allow the analysis of truth ratings conditionalized on recognition judgments that is necessary to test the various hypotheses.

Conditional analyses. The data of interest are presented in Table 2, which displays mean truth ratings for new statements from the second test (column 1), for repeated statements from the second test (column 2), and for those same statements on the first test (column 3); means are partitioned by condition and also by the recognition judgment made. There are several points to note here pertaining to the hypotheses being considered. First, focusing on the mean ratings for repeated statements, one point is that repetitions that were judged correctly to be repeated were no different in initial plausibility than those that were judged incorrectly to be new; the respective means, 3.78 and 3.73, for Condition 2 and 3.65 and 3.65, for Condition 3 allowed us to reject any

Table 2
Truth Ratings According to a Partition Based on Recognition

Recognition judgment	Test 2 new statements	Repeated statements	
		Test 2	Test 1
	Condition 2		
New	4.02	3.81	3.78
Repeated	3.41	3.38	3.73
	Condition 3		
New	3.94	3.75	3.65
Repeated	3.45	3.36	3.65

Note. Standard deviations range from .38 to .68.

hypothesis which asserts that later recognition and truth ratings reflect initial differences in the plausibility of the statements.

Still with reference to the repeated statements, it was possible to determine whether increments in rated truth from session to session depended in any way on recognition judgments. Repeated statements that were judged to be new changed little over occasions, and the change that did occur was in the direction of decreased truth ratings: the change was from 3.78 to 3.81 for Condition 2, and there was a comparable change from 3.65 to 3.75 for Condition 3. In contrast, the repeated statements that were correctly classified as repeated were rated as more true on the second test than on the first, $t(31) = 3.28$, for Condition 2, and $t(30) = 3.65$, for Condition 3. Accordingly, the hypothesis that subjects would only show increments in belief for repeated statements that they failed to recognize as mere repetitions could safely be rejected.

The most important analysis contrasted actual repetition with recognition judgments as correlates of rated truth. As the entries in columns 1 and 2 show, the truth ratings were much more sensitive to judged status than to actual status of the statements. The observations were confirmed by separate 2×2 analyses of variance for the two conditions. Thus, statements that were judged to be old were rated as more true than those judged to be new, $F(1, 31) = 15.1$, and $F(1, 30) = 20.8$, respectively. There was also some difference in truth ratings as a function of whether the statements were actually repeated or new, but the respective Fs of 4.04 and 3.80 were only marginally reliable ($.05 < p < .10$). Unambiguously, then, differences in rated truth are more sensitive to recognition judgments than to actual repetition of the statements being rated. The interaction between variables was not reliable in either condition (Fs < 1), with respective MS_e values of .72 and .19.

Since there were several crucial findings, let us review the findings and discuss their theoretical import before introducing Experiment 2. The basic finding of a positive effect of repetition on belief, initially reported by Hasher et al. (1977), was obtained here in three conditions for each of 10 subsets of items. Accordingly, there is little doubt regarding the robustness or replicability of the repetition effect. However, the main purpose of the experiment was to contrast various theoretical accounts of why the effect occurs. Explanations that attribute the increments in belief to subjects' acquisition of the informational content of plausible statements can be rejected because repeated statements that were judged to be old were rated as more true than repeated statements that were judged to be new, but the difference in rating was only present in the second test. Since remembered and forgotten sentences were of equal initial plausibility, that initial plausibility could not be evoked as an explanation of the different truth ratings on the second test. It is also hard to imagine how such an account could explain the different ratings for new statements that were judged to be repeated or new. Accordingly, such hypotheses will not be considered further.

A second unsatisfactory account, derived from Thouless (1974), assumed that subjects were skeptical, and they would therefore not increase rated truth for statements recognized to be merely repetitions. The major failing of the account is that the only statements that accrued in credibility were those judged to be old, precisely the statements that should not have changed at all. In fairness to Thouless, it should be noted that his account pertains specifically to repetitions close enough together in time for the initial rating to be remembered, whereas the present finding involved a 3-week separation between repetitions. It is possible that Thouless' hypothesis would be appropriate for temporally close repetitions, in which case it would be interesting to determine the length of interval necessary for the repetition effect to occur. It is clear that the present results cannot be accounted for on the basis of memory for previous ratings because the new statements, which had no previous ratings to remember, were as sensitive to recognition judgments as were the repeated statements. Thus, explanations in terms of sub-

jects being suspicious of mere repetition or of retaining prior ratings are inadequate to account for the results obtained and will not be discussed further.

The third unsatisfactory account, offered by Hasher et al. (1977), explains the repetition effect on the basis of accrual in frequency attributes corresponding to repeated statements. Since such attributes increment automatically with repetition, rated truth should depend more on actual than judged oldness. Such a finding was not obtained, although there was some marginal effect of mere repetition in both experimental conditions. If the basic theory were modified to depend not on actual repetition but rather on judged repetition as a prerequisite for accrual of the frequency attribute, the results could be accounted for. However, such a modification would make a frequency attribute an unnecessary and redundant construct. On balance, then, the attribute theory must be rejected as an explanation of the effect of repetition on belief.

The present account, which is certainly the simplest one considered, predictively fares well. Thus, truth ratings are inferred directly from recognition decisions regarding whether or not the test statement is old. Statements judged to be old are also judged to be relatively true, independently of whether the recognition decision is correct. The remainder of this article will address further expectations derived from the recognition account.

Experiment 2

The intent of the second experiment was to determine whether the influence of recognition decisions on truth ratings is modifiable by experimenter-provided information and whether that influence extends to contradictions of the original statements. Recall that the item pool consists of statements and contradictions of those statements, since the true and false version of any statement cannot both be true. Since the present hypothesis asserts that truth ratings depend on recognition decisions, it follows that both verbatim repetitions and contradictions, if judged to be old, will be rated as more true on the second test than on the first. It also follows that the repetition effect should not be removed if subjects are told that repeated statements are repetitions. Common sense might argue that if subjects are told that a set of statements, half of which are actually false, consists of repetitions of earlier statements, then the ratings of truth should not change over occasions. Nonetheless, we predicted a change. Thus, the repetition effect should characterize any statements judged to be old, including contradictions, and should also characterize correctly labeled repetitions.

One question remains regarding the effect on credibility of a subjective judgment that a test statement contradicts an earlier statement, or of experimenter-provided information to the same effect. Consider a pair of contradictory statements, namely, "Scorpions are not immune to their own poison, so they can commit suicide if cornered," as opposed to "Scorpions are immune to their own poison, so they cannot commit suicide if cornered." In a sense, the second statement is new, since its meaning is different from the meaning of the first; in another sense, however, the second statement is a repetition, since there is an obvious similarity in topic. Because contradictions are intermediate between new and repeated statements with respect to content, it could be the case that truth ratings would be intermediate as well. However, our account leads us to expect that contradictions will be judged to be relatively false by the following line of reasoning: A statement that is simply repeated has the same truth status on each occurrence. Since statements judged to be repeated are rated as relatively true, there is an implication that subjects have a bias toward believing the information that they possess or there would be no reason for judged repetitions to change in rated truth. It follows that if statements that reaffirm an earlier piece of information become more true, then statements that contradict an earlier piece of information should become more false. Therefore, it is predicted that test sentences judged to be opposite in truth value to originals should become more false, as should statements correctly classified by the experimenter as contradictions.

To assess the various predictions, subjects were presented with an initial series of statements for truth ratings. Three weeks later, a test series was presented, consisting of statements that either were repeated from the first test, were new, or were contradictory with respect to statements on the first test; half of each set was comprised of true statements. Subjects first classified each test statement as being repeated, new, or changed and then rated its truth. The same test statements were also used for a second condition. In that condition, the repeated statements were blocked together and appropriately labeled, as were both the new and changed statements. However, subjects were not asked to make recognition judgments in this informed condition. Such a request would allow a comparison between subjective decisions and experimenter-provided information as determinants of rated truth, thereby making closer contact with the social literature discussed earlier. However, our intent here is to examine cognitive determinants.

Method

Subjects

Sixty-four students from the same population as in Experiment 1 served as subjects, with 32 in each of two conditions.

Materials and Procedure

An initial list containing 60 critical statements and 30 fillers was presented orally at the rate of three statements/min, during which time subjects rated validity on a 7-point scale from certainly true (1) to certainly false (7); the list was identical to that used in Experiment 1. The second test, presented 22 days later, consisted of 90 statements for each of two conditions. Thirty of the statements were new and 30 were identical to critical items, with half of them true in each case. The remaining 30 test items were alternate versions of critical items; 15 test items were true versions of statements that were false initially and 15 were false versions of initially presented true statements. In an uninformed condition, the 90 test sentences were randomly ordered, with a blank space and the letters *N, I,* and *C* next to each statement. Subjects were informed that 30 statements were verbatim repetitions of earlier statements, 30 were new statements and 30 were changed from earlier statements so that initially true statements were now false and initially false statements were now true. They were instructed to circle N if the test statement was new, I if it was identical to an original, and C if it was changed from an original; following classification, each statement was rated for truth.

In the second condition (the informed condition), the 30 new statements were presented in a block preceded by the statement "These are new statements that were not presented last time," which was followed by the repeated statements headed by "All of these statements are identical to statements presented last time," and, finally, followed by the contradictions headed by "These statements have been changed from last time; they have been rewritten so that if they were true last time they are false now, and vice versa." Subjects simply rated the truth of each statement in a self-paced presentation, using the same 7-point scale as all other conditions.

Table 3
Mean Truth Ratings for Repeated, Contradictory, and New Statements

Statement	Condition	
	Informed	Uninformed
Initial test	3.82	3.61
Second test		
Repeated	3.62	3.41
New	3.73	3.76
Contradictory	4.01	3.80
M	3.78	3.66

Note. Standard deviations range from .30 to .55.

Results and Discussion

Truth ratings: Marginal analysis. Table 3 presents the mean truth ratings for the 60 statements on the initial test as well as for the 30 repeated statements, 30 contradictions, and 30 new statements on the second test. Means for the uninformed condition were without regard to the recognition decisions. For each condition, the repeated statements were rated as more true on the second test than on the first test, $ts(31) > 3.39$, whereas the contradictions of earlier statements were rated as more false on the second test than on the first test, $ts(31) > 2.92$. As in Experiment 1, the overall mean rating for the second test was similar to the overall mean of the first test. The informed condition was in line with expectations, since repetitions

Table 4
Recognition Judgments for New, Repeated, and Contradictory Statements

Judgment	Statement history		
	New	Repeated	Contradiction
New	.778	.232	.355
Identical	.105	.557	.338
Changed	.117	.209	.306

Note. Standard deviations range from .09 to .18.

correctly labeled by the experimenter did indeed accrue in rated truth, just as did the statements judged to be repetitions in Experiment 1. Therefore, the repetition effect occurred even if subjects were told that the statements were merely being repeated. It was also the case that contradictions, on the average, became more false; this point is addressed further following the results of the conditional analysis.

Recognition decisions. The basic judgment data are shown in Table 4. New statements were most likely to be called *new* (.78), and misclassified new statements were equally likely to be called *identical* (.11) or *changed* (.12). Repeated statements were also reasonably well classified, with .56 called identical and with misclassified repetitions about equally likely to be called new (.23) or changed (.21). However, contradictions were about equally likely to be called changed (.31), identical (.34), or new (.36). It is clear that subjects showed a response bias, since test statements were likely to be called new .46 of the time, identical .33 of the time, and changed only .21 of the time. Since the conditional analysis to follow concerns judgment as much as it concerns actual history, it is important to determine the accuracy of the different classifications, given that they were made. Thus, if the new judgment was given, the test statement was most likely to be new (.57), then contradictory (.26), followed by repeated (.17); that is, the contradictions were intermediate between new statements and repetitions. If the identical judgment was given, the test statement was most likely to be a repetition (.56), then a contradiction (.34), then a new one (.11); that is, the identical and new judgments were mirror images of each other, with each being most likely to be applied correctly and least likely to be applied to the alternate set of statements, with contradictions intermediate between the two. Although the changed response was not used as often as the other responses, when it was used, it was applied about as accurately, with the test statement actually being a contradiction .48 of the time, a repetition .33 of the time, and a new statement only .19 of the time.

Conditional analysis. The truth ratings for the uninformed group were analyzed by a 3×3 analysis, with item type (new, repeated, contradiction) and judgment (new, identical, changed) as the factors. Six of the 32 subjects were deleted, since at least one of the nine cells was empty; the means for the remaining 26 subjects were approximately equal to the means presented in Table 3. The means for item type and judgment type are shown in Table 5, along with mean values for the nine cells; the cell values should not be taken as seriously as the marginal means, since standard deviations were quite large. As in Experiment 1, neither the effect of the true history of the items, $F(2, 50) = 1.77$, nor the interaction was reliable, $F(2, 50) = 1.30$. The overwhelming effect was due to judgment type, $F(2, 50) = 16.9$. Accordingly, the positive effect of repetition on belief required a judgment that the test statements were indeed repeated, and this effect extended to cases in which the judgments were wrong. The negative effect of contradiction on belief likewise required that

Table 5
Mean Truth Ratings by Judged and True History

Judgment	Statement type		
	Repeated	New	Contradiction
Identical	3.21	2.88	3.45
New	3.62	3.88	4.00
Changed	3.90	4.11	4.10

Note. Standard deviations range from .50 to 1.42.

the test statement be judged to be contradictory and it again extended to wrong judgments.

Once again, repetition was positively associated with belief, and it was the subjects' decision that a statement was repeated that produced the effect. It would be interesting to compare subjects' judgments of statement history with experimenter-provided information by requesting recognition judgments in an informed condition, to determine the relative importance of the two for truth ratings.

Finally, reversals of truth value of statements over the course of repetition yielded differences in rated truth, with the statements judged to be changed as well as statements classified by the experimenter as changed rated as relatively false.

General Discussion

Our intent has been to determine the cognitive basis of the fact that repeated statements become more credible. Although there are plausible reasons to expect that statements judged to be repetitions would not become more credible, the results are clear: Test statements judged to be repeated are rated as relatively true whether the statements are indeed repeated, whether they are new, or even whether they contradict initial statements. Further, an equivalent increment occurs even when subjects are informed that repeated statements are repetitions. Consequently, the repetition effect is not really a repetition effect after all but a recognition effect.

It also turned out that both statements that subjects judge as changed from originals and correctly-labeled contradictions are especially likely to be rated as false. It is reasonable that if truth-preserving statements become more credible, truth-altering statements should become less credible. In a sense, already known information has a privileged truth status or priority: Subjects believe the information they possess to be true, and they doubt information that contradicts it. It would be interesting to determine whether minor changes, such as changes in sentence structure, would, if detected, lead to decrements in credibility as well.

Another noteworthy effect is that in every comparison, the subjects' perception of an item's history was much more influential than the item's actual history in determining truth ratings. That is, in the recognition conditions, statements judged to be repeated were rated as relatively true, even if the statements were actually new or contradictions. Similarly, when subjects were informed that repetitions were indeed repeated, the statements were also rated as relatively true. A question remains regarding whether, for example, new statements would become more or less credible if subjects are misinformed that the statements are either repeated or contradictions.

The final finding of note is that mean truth ratings remained constant from test to test, although the items included in the tests were different. If this mean-preserving tendency is general, there are several implications. For example, if only repeated statements are presented for second ratings and if subjects are informed that they are repetitions, there should be little change in rated truth. Similarly, it should be possible to have the repeated statements become more false on the second test if that test contains obviously true new statements. At any rate, there are several obvious directions for further research.

As a final comment, it is worth bearing in mind that all of the statements used here are examples of trivia. We chose such statements because we wished to study repetition in a range in which one repetition makes a difference, and we wished to allow for measurable changes in rated truth to occur. We have no way of knowing whether personally relevant statements would obey the same laws. If a mother repeatedly calls her child good, will the child come to believe it? If a statement is prefaced by "I've told you a thousand times before," will it be more credible? Pending the outcome of appropriate experiments, it might be instructive to ask ourselves how many of the statements we believe to be true are believed on the basis of repeated encoun-

ters rather than on the basis of the accrual of evidence in their favor.

The present theory is simply that ratings of truth are dependent on the outcome of recognition decisions rather than on the basis of more elaborate attributes, on memory for prior ratings, on familiarity with content, or on other more complex cognitive procedures. Of course, complexity may be required in other domains or even in some extensions of the present research. However, until evidence forces additional complexity, it would be inefficient to assume it.

References

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review,* 1972, *79,* 97–123.

Asch, S. E. Studies in the principles of judgments and attitudes: II. Determination of judgments by group and ego standards. *Journal of Social Psychology,* 1940, *12,* 433–465.

Asch, S. E. The doctrine of suggestion, prestige and initiation in social psychology. *Psychological Review,* 1948, *55,* 250–276.

Begg, I. Estimation of word frequency in continuous and discrete tasks. *Journal of Experimental Psychology,* 1974, *102,* 1046–1052.

Hasher, L., & Chromiak, W. The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior,* 1977, *16,* 173–184.

Hasher, L., Goldstein, D., & Toppino, T. Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior,* 1977, *16,* 107–112.

Hintzman, D. L., & Block, R. A. Repetition and memory: Evidence for a multiple trace hypothesis. *Journal of Experimental Psychology,* 1971, *88,* 297–306.

Hovland, C. I., Janis, J. L., & Kelley, H. H. *Communication and persuasion.* New Haven, Conn.: Yale University Press, 1953.

Howell, W. C. Storage of events and event frequencies: A comparison of two paradigms in memory. *Journal of Experimental Psychology,* 1973, *98,* 260–263.

Kenny, A. *Wittgenstein.* London, England: Penguin Press, 1973.

McKenzie, E. C. *Salted peanuts.* New York: Signet, 1976.

Russell, B. *An inquiry into meaning and truth.* Harmondsworth, Middlesex, England: Penquin Books, 1973.

Thouless, R. H. *Straight and crooked thinking.* Suffolk, England: Richard Clay (Chaucer Press), 1974.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* 1973, *5,* 207–232.

Underwood, B. J. Attributes of memory. *Psychological Review,* 1969, *76,* 559–573.

Received August 23, 1978
Revision received November 29, 1978 ■