

# RECALL OF INTERRUPTED TASKS UNDER STRESS: A PHENOMENON OF MEMORY OR OF LEARNING?<sup>1</sup>

ALBERT J. CARON AND MICHAEL A. WALLACH

*Harvard University*

PERHAPS no experimental approach to repression has received more attention than the attempt to administer Zeigarnik's interrupted-task procedure (22) under ego-involving conditions. The rationale underlying this use of the Zeigarnik technique held that ego involvement would cause incompleting and completed tasks to be viewed as failures and successes respectively, and hence, in accordance with repression theory, would lead to the expulsion of incompleting items from consciousness. The prediction followed, therefore, that the typical "Zeigarnik" recall pattern (superior recall of incompleting tasks) would be replaced by a "repressive" recall pattern (superior recall of completed tasks).

Since this view was first advanced, much evidence has accumulated (1, 2, 3, 7, 8, 18, 19) which indicates, not surprisingly, that such a reversal of the Zeigarnik ratio is not a function of stress alone, but depends, in addition, on personality factors—some individuals recalling a preponderance of successes and others yielding extreme Zeigarnik ratios under ego-involving conditions. While the isolation of the particular personality dimensions associated with each of these stress recall patterns offers a fertile field for personality research, the mechanisms assumed to underlie each pattern—namely, repression of failures in the case of selective recall of successes (S-recall) and superior trace stabilization of failures in the case of selective recall of failures (F-recall)—are still open to question. As others (15, 20) have observed, finding a selective recall preference for success items under stress is not definitive proof that failure items were repressed. An alternative explanation, selective *learning* in favor of success items, has not been ruled out. Likewise, a selective recall preference for failure items under stress may arise not from enhanced stabilization of their traces but rather from their being better learned in the

first place. The obvious experimental solution here—to equate successes and failures for original learning—is hard to achieve, for if failure items are really threatening, then the same factors which theoretically lead to their repression (or trace stabilization) may also be expected to impair (or improve) their learning.

The aim in the present study was to provide a more feasible test of the memory versus learning explanations of both these recall patterns. The approach employed follows from an analysis of the assumptions underlying interpretation of each pattern as a memory or learning effect. Consider first the alternative assumptions with respect to S-recall under stress.

As has been noted (4, 23), the essential postulate of repression theory is that threatening events are not actually forgotten but persist in an unconscious state, continuously striving to regain consciousness. Moreover, should the threatening character of repressed events be allayed (via psychotherapy or other means), they should *re-emerge* into consciousness. The selective learning position, on the other hand, implies no such restoration for forgotten items: it assumes that decreased recall results entirely from a deficiency in original registration. The two interpretations thus offer different predictions for the recall of events whose threatening nature is allayed after learning—the repression view implying enhanced recall, the selective learning view implying no change in recall, for such items. In the present study, two groups (both known to exhibit S-recall under stress) were subjected to a stressful interrupted-task procedure. For one group, however, recall was preceded by a quasi-therapeutic session in which the test was exposed as an experimental hoax. If the S-recall pattern is indeed due to repression, such exposure, by reducing associated anxiety, should restore the recall of incompleting items, and hence cause this group to shift toward a Zeigarnik pattern. If, on the other hand, selective learning underlies S-recall, then no such shift should occur.

With respect to the F-recall pattern, ex-

<sup>1</sup> This research was conducted while the senior author held a U.S.P.H.S. postdoctoral research fellowship at Harvard University. Grateful acknowledgment is made to Dr. Jerome S. Bruner for making available the facilities of his laboratory.

planation in terms of selective learning is much like that offered above for S-recall. Recall is said to be totally dependent on learning and, whatever the reason may be, failure items have registered more strongly than success items. The interpretation in terms of selective remembering, on the other hand, is based, according to Rosenzweig's recent formulation (17), on the task-tension theory of Lewin (12) and Zeigarnik (22). According to this view, the traces of failure items are better stabilized than those of success items, because the former are laid down in a region of high psychic tension—tension aroused by instructions to perform a task and dischargeable only by its completion. Rosenzweig has called this a "need-persistent" reaction to failure as contrasted with the "ego-defensive" reaction of repression. This position predicts that anything done after learning to reduce task tension should bring about a reduction in the recall of failure items. The selective learning view, on the other hand, implies no recall change as a consequence of such post-learning manipulation. To test these predictions, two other groups—both known to exhibit F-recall under stress—were subjected to the same stressful situation described above, and recall for one group again followed exposure of the experimental hoax. If the F-recall pattern is indeed due to superior trace stabilization of failures, the revelation that incompletable tasks were actually impossible to solve should serve to discharge any residual tension associated with incompletable and thereby tend to reduce the recall of failures. No such reduction in recall of failures should occur, on the other hand, if this effect is due to differential learning.

A final question examined in the present study may now be mentioned. If our results should indicate that both the S- and F-recall patterns are a function of learning rather than retention, does this mean that there is no evidence in the present experiment for the operation of mnemonic processes? This question is prompted by consideration of an aspect of the present kind of procedure which distinguishes it from the original Zeigarnik technique, a difference heretofore neglected by most writers. To construct a situation sufficiently threatening for college students, the experimental tasks have usually been presented in the guise of an intelligence test and indeed are so presented here. In so doing, however,

the procedure is given a unity and integration which is not present in the typical Zeigarnik situation, where *S* is merely asked to perform a series of disparate and ostensibly unrelated tasks. Under intelligence-test conditions, then, it is entirely possible that *S*'s defensive reaction is oriented more toward the test as a whole, i.e., toward the single experience of examination failure, rather than toward particular items within the test, and hence selective recall of individual items may not be a measure optimally sensitive to *S*'s invoked defense. Memory for the entire test situation would seem more pertinent, but for obvious reasons such a measure is unfeasible. To determine whether there was indeed a mnemonic reaction to the total test situation in this study, an alternative indicator in the area of perception was employed. The rationale for its use is as follows: If *S* tends to repress an experience, aspects of that experience should be relatively inaccessible as hypotheses for perceptual recognition. Likewise, if the trace of an experience is well stabilized for *S*, aspects of it should be hyperaccessible as recognition hypotheses. Thus, following the stress situation, words related to it in a general way were presented tachistoscopically (along with matched neutral words). A repressive reaction to the entire situation should result in heightened recognition thresholds for such words; enhanced trace stabilization, on the other hand, in lowered thresholds. Moreover, if these effects are truly memorial, and not due to registration processes, they should cease to occur when exposure of the experimental hoax precedes the perceptual test.

#### METHOD

##### *Subjects*

Since none of the personality or performance measures previously shown to relate to selective recall under stress have been conclusively validated, it was not considered appropriate to use any one or combination of them as a basis for preselecting the S-recallers and F-recallers in our various treatment groups. As an alternative to preselection, it was decided to administer a large battery of promising measures<sup>2</sup> to randomly selected *Ss* already tested in our experimental conditions, to factor analyze these measures, and, from the correlations between the resulting factors and the

<sup>2</sup> The measures used included tests related rationally as well as experimentally to each of three personality areas: achievement motivation, hysteria, and ego strength.

selective recall scores of our stress group, to isolate a stable predictor that could be used *post hoc* to subdivide all treatment groups. This analysis yielded four orthogonal factors, only one of which was at all related to recall direction in the stress condition. The measure with the highest loading in this factor (and, incidentally, with negligible loading in the other factors) was the dichotomous variable of private prep school versus public high school attendance, and indeed this single variable was more strongly associated with our recall measure<sup>3</sup> than any other variable or combination of variables defining this factor (see Footnote 4 for the major variables). Thus, the public-private dichotomy involved fewer dimensions than one might at first have supposed, and also accounted for most of the factorial relationship with recall direction. These considerations, coupled with considerable evidence suggesting that the reactions of Harvard public and private school Ss should be quite different in an intelligence-test situation,<sup>4</sup> led us to use this variable to subdivide our groups into S-recaller and F-recaller components.

The Ss were chosen in the following way: Invitations to participate in the experiment were mailed out to every tenth name in the list of Harvard freshmen. The list was gone through twice in this manner, starting at a different name each time. The letter attractively described the experiment as a study of "symbolic processes," and also emphasized the remuneration for participation. About half of the men to whom letters were mailed volunteered—125 in all, of whom 117 completed the experiment—and they were randomly assigned to the several experimental conditions. The number of public and private school Ss in each condition was roughly equivalent.

### Sequence of Procedures

The order of procedures for all Ss was as follows: (a) initial digit recall test (4 minutes);

<sup>3</sup> Chi square was significant at less than the .025 level with a one-tail test—use of a one-tail test being justified since recall directions were predicted from the findings discussed in Footnote 4.

<sup>4</sup> McArthur's studies (13, 14) of the personalities of Harvard public and private school boys, for instance, indicate that the former tend to have a stronger drive for achievement and to be more sure of their personal and intellectual worth than the latter. Research by Atkinson (3) provides further support for this view. Ss with high need for achievement recalled more failures under stress and more successes under neutral conditions, while Ss with low need for achievement did the opposite—results which directly parallel our recall findings for public and private school Ss respectively. Finally, we may consider the performance of public and private school boys on some of our independent measures. The public school students made significantly higher grades at Harvard than those from private schools ( $p < .02$  with a two-tail test), although the two groups did not differ in intelligence. Moreover, on the factor mentioned above, public school attendance was associated with such achievement variables as level of aspiration, while private school attendance was related to such indications of personal insecurity as the Knutson Personal Security Inventory (11) and the Social Introversion scale of the Guilford STDCR battery (10).

(b) scrambled sentence tasks (50 minutes); (c) second digit recall test (4 minutes); (d) interim discussion with *E* (8–10 minutes); (e) initial recall of scrambled sentences (5 minutes); (f) determination of perceptual recognition thresholds (60 minutes); (g) second recall of scrambled sentences (5 minutes); (h) third recall of scrambled sentences, two days later (5 minutes).

### Tests

*Digit recall.* Half of the Wechsler-Bellevue Digit Recall Test (21) was administered before, and half after, the scrambled sentence tasks, one instance of each forward and backward couplet being presented in each administration. All responses were written from left to right, and writing never began until *E* finished reading the digits. The number of digit sequences completely correct provided the metric. Scores for the first half alone constituted a check on the equality of memory ability for the several experimental groups; the difference between scores for the first and second halves provided a measure of the anxiety induced by the intervening work on scrambled sentences.<sup>5</sup>

*Scrambled sentences.* The experimental tasks were sixteen 20-word sentences similar to ones used by Alper (1), Eriksen (6), and Caron (5). Each sentence was divided into 8 to 10 two- and three-word phrases and was presented in scrambled form, *S*'s job being to rearrange them into a meaningful sentence. Half of the sentences, randomly distributed in the series of sixteen, were unsolvable. Contents of the sentences were diversified so that recall of one would be no aid to recall of another, and the completed and incomplete sentences were equated for recall difficulty in a previous study [see Caron (5)]. To facilitate group administration, the sentences were presented in booklet form, one sentence to a page. A 2¼-min. time limit was set for each solution and all Ss were required to start a new sentence at the same time. As an aid to recall, each sentence was named by selecting one phrase that best represented its meaning and placing this phrase above the scrambled sentence in the test booklet. *E* called out this "name" phrase at the start of each problem. For both initial and later recall tests, *S* had no prior knowledge that recall would be requested. In recalling sentences, *S* was required to write the name of a sentence or any other key phrase or word sufficient to identify it without question, no more than five minutes being allowed for each recall session. The measure of recall direction was the number of completed minus the number of incomplete tasks recalled (CR-IR).

*Perceptual materials.* Immediately after initial recall, perceptual recognition thresholds were determined for words pertaining in a general way to the scrambled sentence test and for neutral words of matched structure and of matched frequency (according to the Thorndike-Lorge L count). The former group included the following words: *sentence, booklet, phrases, examination, reasoning, scrambled, and tested*; the corresponding neutral words were *sunshine, cookies, sketches, agriculture,*

<sup>5</sup> For the use of digit recall as a measure of anxiety, see Rapaport *et al.* (16).

*leisurely, channeled, and regard.* Ss, tested in groups of four, were seated five feet from a ground glass screen in a room illuminated solely by a 40-watt fluorescent ceiling bulb one foot behind S. Each word, no more than 14 inches in length, was flashed at  $\frac{1}{100}$ -second, beginning at subthreshold brightness levels and rising in two-volt steps until all Ss recognized the word. Threshold was taken as the voltage level at which correct report first occurred. The order of threatening and neutral words was random, with matched words always being seven positions apart. To prevent communication between Ss, responses were written, E observing these and telling each S to stop and turn over his paper when he had written the correct word.

### Conditions

*Stress.* The stress condition was structured with two objectives in mind: (a) to convince S that he was taking a highly valid and discriminating intelligence test, and (b) to make him feel that he had done very poorly on it. With respect to the former objective, a number of techniques were used: In the first place, the scrambled sentence task was introduced as a test recently developed by the "Stanford people" in connection with a top level government project, a brief comment being added on the relationship of the Stanford people to the Stanford-Binet test. Second, E implied that the government project involved careful selection of personnel for critical intelligence work, and hence that the present test had been designed to provide precise discrimination in the average to superior range of adult intelligence. Third, bogus validity data were presented specifying academic and professional success as criteria. Finally, the title "Stanford Reasoning Test" was printed on the cover of the test booklet, and beneath that were printed "name," "age," and "IQ." (IQ was to be filled in, if known. Accomplices, discussed below, always filled in high IQs.) That an intelligence test would be administered in a study ostensibly concerned with "symbolic processes" was rationalized on the ground that E wished to compare the symbolic performances of average and highly intelligent students.

The following means were used to make S feel that he had done badly on the test: First, E stated that most Harvard students should perform very well on this test, since they were undoubtedly at the high end of the distribution of adult intelligence. Second, Ss were tested in groups of six, and each such group included two accomplices who pretended to solve all the tasks rapidly and exuded an air of confidence throughout the session. Finally, the four naive Ss received alternate forms of the test such that when any two of them failed to complete a task the other two Ss completed that same task, the result being that for any single task, four of the six testees achieved solution while two did not.

After completion of the exam, the second digit recall test was administered, and was immediately followed by a period of 8-10 minutes during the first part of which E made remarks intended to reactivate feelings of stress (referring to the exam's highly discriminative nature, the fact that scores would be made available, etc.), and during the remainder of which he scheduled Ss for subsequent sessions. Such scheduling was done at this point in order to make the interim period for the stress condition as long as that for the relief condition, and yet prevent Ss in the stress condition from discuss-

ing the exam among themselves. Half of the stress Ss were kept under stress for all three recall sessions, while the other half were relieved just prior to the second recall test.

*Relief.* This condition was exactly the same as the stress condition, except for the contents of the 8-10-minute interim period mentioned above. Rather than reactivate feelings of stress during this period, E attempted to relieve such feelings. He said the following: "Everybody sit back and relax. From the looks on your faces I gather that this was a pretty rough test—but perhaps you didn't do as badly as you think. You'll hate me when I tell you this, but believe me, it was all in the interest of science . . ." E then explained that the entire situation had been a hoax—that half the tasks had been unsolvable, that the two apparent geniuses in the group had really been E's accomplices, and so on. This revelation released a flood of affect—best described as a mixture of surprise and tremendous relief—which was inevitably followed by a round-robin of descriptions of one's feelings and reactions during the session. The subsequent airing of these feelings gave the session a definite therapeutic tone.

*Neutral.* In contrast to the stress condition, the aim of the neutral condition was to create as informal and relaxed an atmosphere as possible, and to focus attention on the tasks rather than on S's performance. To this end, the scrambled sentence tasks were introduced as a pretest of materials for a future experiment. The fact that the tasks were timed was rationalized by telling S that the purpose of the pretest was to eliminate tasks requiring too much time, E supposedly being interested in tasks that took no longer than about two minutes to solve. Moreover, whereas the timing for the stress condition was done in a test-like manner with E holding his stop watch conspicuously, the timing in the neutral condition was done unobtrusively and casually. The booklet of scrambled sentences, in addition, had no title on it, and S was told that he need not supply his name. Finally, no accomplices were present and all Ss had the same form of the test, so that everyone completed, or failed to complete, any given task. The 8-10-minute interim period between the second digit recall test and initial recall of the scrambled sentences was spent scheduling Ss for further sessions and having them fill out parts of some questionnaires.

### RESULTS

Before considering the results which bear directly on our major questions, let us examine the data pertaining to a basic assumption of the study; namely, that the cognitive effects obtained in our stress condition were due to the ego-involving operations themselves, rather than to something inherent in the test materials or to subject differences alone. Comparison of the stress and neutral conditions for each of the subject groups indicates that this assumption is entirely tenable: Private and public school groups yielded significant differences ( $p < .01$  in each case) between stress and neutral conditions on both the recall difference (Table 1, Column 2) and the per-

TABLE 1  
MEAN INITIAL DIGIT RECALL (D), SCRAMBLED SENTENCE RECALL SCORES, AND PERCEPTUAL RECOGNITION SCORES FOR S-RECALLERS (PRIVATE SCHOOL Ss) AND F-RECALLERS (PUBLIC SCHOOL Ss) UNDER THREE EXPERIMENTAL CONDITIONS<sup>a</sup>

Condition	D	CR-IR	TR	CR	IR	P <sup>b</sup>
S-recallers (Private School Ss)						
Stress	9.5 (N = 19)	+1.1	5.5	3.3	2.2	+6.0
Relief	9.3 (N = 19)	+0.6	4.4	2.5	1.9	+0.9
Neutral	8.6 (N = 15)	-0.2	6.4	3.1	3.3	-1.7
F-recallers (Public School Ss)						
Stress	9.0 (N = 25)	+0.1	5.3	2.7	2.6	-2.6
Relief	10.2 (N = 23)	+0.1	5.7	2.9	2.8	-0.7
Neutral	9.4 (N = 18)	+1.2	6.8	4.0	2.8	+3.8

<sup>a</sup> The scrambled sentence recall and perceptual recognition measures include mean difference between number of completed and incompleting tasks recalled (CR-IR), mean total number of tasks recalled (TR), mean number of completed (CR) and incompleting (IR) tasks recalled, and mean perceptual threshold difference for threat-related and matched neutral words (P).

<sup>b</sup> A positive difference means higher thresholds for threat-related than for neutral words. The brightness difference in volts is twice the size of each entry.

ceptual threshold scores (Table 1, Column 6). Moreover, the directions across recall and perception were completely consistent for each group, private school Ss exhibiting S-recall as well as "repressive" perceptual patterns under stress but just the opposite tendencies under neutral conditions, public school Ss doing exactly the reverse. (The interaction *F*s are significant beyond the .01 level for each variable. See Table 2.)<sup>6,7</sup>

<sup>6</sup> Lack of a strong recall tendency in favor of incompleting tasks in any of our subgroups indicates the presence of a general recall bias in favor of completed tasks. Thus, when the term "F-recall pattern" is employed in the present study, it is being used in a relative, not an absolute, sense. That this term is nevertheless being used appropriately, i.e., to specify a type of cognitive reaction, is indicated by the finding, noted earlier, that our group differences directly parallel those of Atkinson (3) for high and low need achievers tested under similar conditions.

<sup>7</sup> In all our analyses of variance, the form of the test,

For both private and public school Ss, total recall drops significantly from neutral to stress conditions (Table 1, Column 3). However, while for private school Ss this is due entirely to a decline in recall of failures (Column 5)<sup>8</sup> for public school Ss it is completely determined by a decline in recall of successes (Column 4). These losses assume added import when one compares the groups on our measure of recall ability—digit-retention. Column 1 of Table 1 shows the mean pretest digit-retention scores for the subgroups. The analysis of variance of these scores yielded no significant or near significant *F*s (Table 2). Consider now the findings that indicate whether these losses are due to mnemonic or learning processes.

1. *Is the S-recall pattern a function of repression of failures or selective learning in favor of successes?* The critical datum here is the degree of shift in the recall difference score from stress to relief conditions for private school boys (Table 1, Column 2). Although a slight shift occurred, it was far from significant, *t* being less than 1. Recall of incompleting tasks (Table 1, Column 5) shows a nonsignificant decrease from stress to relief conditions, whereas, if a repression effect were operating and given the mnemonic comparability of our groups, one would expect, if anything, an increase (memory for failures being restored by the relief treatment). Recall of completed tasks (Column 4) and total recall (Column 3) also exhibit nonsignificant declines. In sum, then, no immediate recall shift appeared in the relief group. To determine, on the other hand, whether a delayed shift occurred for this group or for other groups; namely, for stress Ss who were subsequently relieved, and for nonrelieved stress Ss (the last as a function of time lapse alone), we turn to the data of Table 3 (Columns 1-6). None of the recall differences recorded there (for CR-IR, CR alone, or IR

i.e., which particular tasks were completed and which were incompleting, was found to be a nonsignificant source, and hence its variance was added to the within-cells variance.

<sup>8</sup> These results are similar to those of Glixman (9). Whereas Glixman, however, regards such findings as indicative of a repression process, we are suggesting that they do not point unequivocally to such a conclusion.

TABLE 2

ANALYSES OF VARIANCE OF INITIAL DIGIT RECALL (D), SCRAMBLED SENTENCE RECALL SCORES, AND PERCEPTUAL RECOGNITION SCORES FOR S-RECALLERS (PRIVATE SCHOOL Ss) AND F-RECALLERS (PUBLIC SCHOOL Ss) UNDER THREE EXPERIMENTAL CONDITIONS

Source	D			CR-IR		TR		CR		IR		P	
	df	ms	F	ms	F	ms	F	ms	F	ms	F	ms	F
Type of Ss	1	3.86		0.76		5.80		0.56		2.62	1.19	111.34	2.15
Conditions	2	5.78	1.16	0.96		23.25	3.82*	7.46	3.50*	4.84	2.20	13.12	
Ss × Conditions	2	6.32	1.27	13.57	5.19*	5.50		5.38	2.53	4.55	2.07	458.50	8.84*
Error	110	4.99		2.61		6.09		2.13		2.20		51.87	

\* For 2 and 110 df, an F of 3.09 is significant at the .05 level; an F of 4.81 is significant at the .01 level.

TABLE 3

MEAN DIFFERENCE BETWEEN FIRST AND SECOND RECALL SESSIONS, AND BETWEEN FIRST AND THIRD RECALL SESSIONS, FOR CR-IR, CR, AND IR

Condition of Initial Recall	Condition of Delayed Recall	S-recallers (Private School Ss)						F-recallers (Public School Ss)					
		2nd-1st sessions			3rd-1st sessions			2nd-1st sessions			3rd-1st sessions		
		CR-IR	CR	IR	CR-IR	CR	IR	CR-IR	CR	IR	CR-IR	CR	IR
stress	stress	-0.1 (N = 13)	0.2	0.3	-0.5 (N = 12)	0.0	0.5	-0.7 (N = 10)	0.2	0.9	0.1 (N = 10)	0.2	0.1
stress	relief	0.0 (N = 6)	0.0	0.0	-1.2 (N = 6)	-1.0	0.2	0.0 (N = 13)	0.3	0.3	0.4 (N = 13)	0.3	-0.1
relief	relief	0.1 (N = 19)	0.3	0.2	-0.2 (N = 16)	-0.2	0.0	-0.3 (N = 23)	0.0	0.3	-0.3 (N = 23)	-0.2	0.1
neutral	neutral	0.2 (N = 15)	0.4	0.2	0.2 (N = 15)	0.1	-0.1	-0.5 (N = 18)	0.0	0.5	-0.6 (N = 18)	-0.2	0.4

alone) are significantly different from zero,<sup>9</sup> hence implying no delayed restoration of forgotten items. The present findings thus do not support predictions derived from a repression interpretation of the S-recall pattern, but are consistent with those derived from a selective learning interpretation.

2. *Is the F-recall pattern a function of selective remembering or selective learning in favor of failures?* To answer this question, we must examine the degree of shift in the recall difference score from stress to relief conditions for public school Ss (Table 1, Column 2). No such shift occurred, the two means being identical. Recall of incompleting tasks alone, moreover, shows a nonsignificant increase, whereas the task-tension theory would predict, if anything, a decrease. Recall of completed tasks and total

recall also show negligible increases. Here again, then, no immediate recall shift occurred in the relief condition. Further, there was no delayed shift for any of the subgroups. As indicated in the last six columns of Table 3, none of the differences for CR-IR, CR alone, or IR alone, were significantly different from zero. Our findings, therefore, do not confirm predictions derived from a task-tension interpretation of the F-recall pattern, but once more are consistent with those implied by a selective learning view.

3. *Is there a mnemonic reaction to the test as a whole?* The data relevant to this question are the threshold shifts from stress to relief conditions for test-related words (compared to matched neutral words). According to Table 1, Column 6, private school Ss in the relief group show a significant return ( $p < .05$ , two-tail test) to the level obtained under neutral conditions, whereas public school Ss do not shift significantly. That the perceptual

<sup>9</sup> As can be seen from Columns 5 and 6, the CR-IR difference of -1.2 ( $p < .075$ ) is a function of decreased recall of completed tasks rather than increased recall of incompleting tasks.

defense pattern dissipated for private school Ss when they were relieved, suggests that a genuinely repressive reaction—a memory rather than a learning phenomenon—occurred for these Ss with regard to the threatening situation as a whole. Since the perceptual vigilance reaction of the public school Ss, on the other hand, did not dissipate in the relief condition, it would seem that this reaction is due to a learning rather than a memorial process.

#### DISCUSSION

Three propositions summarize the present results. 1. *Both the S- and F-recall tendencies in the present study are due to a selective learning rather than a selective remembering mechanism.* The pertinent datum here is the following: Among the boys from private schools, i.e., those exhibiting an S-recall pattern, the tendency to recall failures was no greater for relieved than for nonrelieved Ss; likewise, among the boys from public schools, i.e., those exhibiting an F-recall pattern, the tendency to recall failures was no less for relieved than for nonrelieved Ss. 2. *A repression mechanism was found for S-recallers, but with respect to the stress situation as a whole rather than for specifically failed items.* The evidence here: Nonrelieved private school boys manifested perceptual defense for words related in a general way to the examination situation, whereas those who were relieved showed no such perceptual defense. 3. *Instead of a comparable process of enhanced retention of the total stress situation, F-recallers showed enhanced registration for the total situation.* Thus, although public school boys under stress were perceptually vigilant for test-related words (relative to public school boys under neutral conditions), this effect did not dissipate in the relief condition.

Before we can accept these statements, two critical procedural issues require reexamination: namely, (a) was the stress situation sufficiently threatening; and (b) if so, was the relief treatment really therapeutic? Consider first the question of threat.

Needless to say, the present data would have no bearing on a theory of repression if the stress situation had not induced intense ego threat. That such threat did indeed occur may be inferred from several lines of evidence—

observational, introspective, and experimental. In the first place, Ss in the stress condition showed many overt signs of emotional disturbance (little or none of which appeared among the neutral Ss). When stumped by a problem, they would squirm nervously or shake their heads in disgust (sometimes swearing simultaneously) or would stare at the page with a very pained expression. Further, there were many attempts to conceal one's difficulty from others, either by pretending to have finished problems not actually completed, or, when able to solve a problem, by trying to make one's success very obvious to competitors (e.g., putting one's pencil down with a bang or yawning loudly while stretching). Again, depressive and hostile reactions were quite prevalent following the exam, many Ss sitting morosely in their chairs, others openly challenging the worth of the test. Lastly, expressions of great relief—deep sighs, postural relaxation, laughter—almost always appeared when the ruse was finally explained. As to introspective evidence: First, most Ss voluntarily reported that they had been extremely concerned prior to, and tremendously relieved following, exposure of the hoax. They attributed emotions to themselves ranging all the way from intense feelings of inadequacy ("I said to myself, 'See, you really don't belong at Harvard' ") to violent hostility ("I wanted to get that supercilious bastard sitting beside me.") Secondly, continued questioning revealed that although some thought the test a poor one, no one had doubted its reality. With regard to experimental data, finally, consider the results for our indicator of anxiety: digit recall. Whereas recall of digits showed a marked pretest to posttest improvement for Ss in the neutral condition, Ss under stress showed no such improvement, the *F* between conditions being significant well beyond the .01 level. The most plausible explanation for the latter Ss' deficiency would seem to be the disrupting effects of anxiety generated during the scrambled sentence test. These data, in sum, would seem to support the assumption that intense ego threat was aroused in the present study.

Let us turn now to the second question: Can exposure of the hoax really be considered a therapeutic, i.e., anxiety-reducing, technique?

Even if it were conceded that anxiety had been aroused in our stress condition, the present findings could be legitimately challenged on the ground that this anxiety was not effectively reduced in the relief condition. Such a challenge might proceed as follows: In the typical clinical setting, the therapist's efforts to uncover repressed events meet with strong defensive resistance. Hence, recovery of repressed material requires a great deal of time and involves a special kind of intimacy between patient and therapist, neither of which occurred in the present relief treatment. The point is well taken, but it overlooks two important differences between repression as it ordinarily occurs and repression as precipitated in the present experiment. First, in ordinary circumstances, the fearful event is unknown to the therapist, having occurred before therapist and patient began their relationship. Thus, as long as the therapist is ignorant of the actual situation and cannot confront the patient with reality, the patient may safely continue to hide behind his defense and deny the event. In the present situation, however, the individual is faced with an omniscient therapist—one who produced the threatening event and who is as much an authority on its nature as the patient. To continue repressing in the face of such omniscience is to render one's contact with reality precarious at best. Second, and perhaps more critical, there would seem to be but two ways in which a threatening event may be made less fearful: (a) by effecting a radical change in the patient's motivational structure (e.g., eliminating his fear of intellectual incompetence) so that the event (e.g., exam failure) no longer poses a threat; or (b) by altering the significance of the event so that it is no longer motivationally relevant. The former is the proper goal of the therapist in the usual clinical setting. The latter is what was obviously involved in the revelation of the experimental hoax to our Ss: the failure experience had initially been created by deception and was now being restored to its true meaning as a psychological gimmick. *E*'s revelation does not make *S* any less of a repressor or eliminate his problem with respect to intellectual competence (nor was it intended to), but merely renders this particular event irrelevant to that problem and hence no

longer something to be defended against. In the light of both these points, then, the assumption that effective therapy occurred in the present experiment seems entirely tenable.

Two general conclusions follow from the present findings. One is methodological, and touches on the question of criteria for a suitable test of repression. The other concerns the particular mechanisms that underlie the avoidant and adient cognitive reactions obtained here. Consider each of these in turn.

1. In attempting to provide an experimental demonstration of repression, the general practice has been to compare the recall of certain "nonthreatening" aspects of the experimental situation with that for certain threatening aspects. In particular, it has been common to examine the recall of failures in contrast to successes on a bogus intelligence test. The present results impute this approach, for they suggest that repression embraces only the most general aspects of the stress situation, whereas recall of particular threatening or "nonthreatening" aspects of that situation is a function of learning. This finding seems entirely reasonable, for when a person has failed half the items on a test, what is likely to be threatening for him is not that he has failed those particular items but that he has failed the *test*. Nor can the items he has passed be entirely free from threat if the total situation in which they are embedded arouses anxiety. It would thus seem necessary, in experimental research on repression, to determine precisely what category of materials is relevant to the subject's conflicts and fears. While recent experiments in this area have been careful to satisfy the previously neglected criterion of motivational relevance of experimental materials, the present findings strongly suggest it would also be well to consider the cognitive relevance of these materials.

2. That both public and private school Ss showed opposite and significant differences between stress and neutral conditions, and that the directions of these differences were completely consistent across recall and perception, indicate that two quite different defensive reactions to stress occurred in this study. One—an avoidant reaction by private school Ss—was indeed shown to be repressive as far as the total test situation was concerned, but, with regard to particular items, was a

matter of registration differences. The other—an adient reaction by public school Ss—was a case of registration differences at both levels. One may properly inquire at this point what the particular processes might be which determine the various registration differences found. Osgood (15), among others, has suggested that differential rehearsal may account for such learning superiorities. The possibility of covert review of failures by Harvard public school Ss indeed seems especially reasonable on the assumption that these Ss are high achievers. While such individuals are not unconcerned about failure, they undoubtedly have a history of repeated reward for efforts to counteract initial failure. Hence these Ss would be expected to ruminate over their difficulties as a cognitive precursor to constructive adjustment. In the case of private school Ss, covert review of successes might be expected as a kind of ego-bolstering reaction. Another possibility for these Ss is that anxiety over inability to solve items acts to disrupt their registration. These two mechanisms are, of course, by no means mutually exclusive.

#### SUMMARY

The present study proposed to answer three questions: 1. Is superior recall of successes in an intelligence test situation (S-recall) a function of selective forgetting (repression) of failures or selective learning in favor of successes? 2. Is superior recall of failures in this situation (F-recall) a function of selective remembering or selective learning in favor of failures? 3. Is there a mnemonic reaction to the test as a whole? The first two questions were examined by comparing the recall preference of nonrelieved S- and F-recallers on a bogus intelligence test (involving solvable and unsolvable scrambled sentences) with that for comparable relieved Ss, selective remembering theories predicting a shift from nonrelieved to relieved conditions for each recall group. The third question was examined by comparing the perceptual recognition thresholds of the above relieved and nonrelieved Ss on exam-related and matched neutral words. Forty-two Ss each were tested in the relief and nonrelief condition, with 33 more tested under neutral conditions. Prior attendance at a private

versus a public high school was used as the indicator of S- and F-recallers respectively.

The findings were as follows: 1. Both the S- and F-recall tendencies were due to a selective learning rather than a selective remembering mechanism. 2. A repression mechanism was demonstrated for S-recallers with regard to the stress situation as a whole. 3. A comparable process of enhanced retention of the total stress situation was not demonstrated for F-recallers; rather, the effect here was due to enhanced registration. Two conclusions were drawn from these findings: First, that repression was found for the total test situation rather than for specific items, suggests experimental tests of repression must fulfill a criterion of cognitive relevance of test materials. Second, the kinds of enhanced registration found may have been due to covert rehearsal, and possibly also to disruption of the registration of other items by anxiety.

#### REFERENCES

1. ALPER, THELMA G. Memory for completed and incompleting tasks as a function of personality: An analysis of group data. *J. abnorm. soc. Psychol.*, 1946, **41**, 403-420.
2. ALPER, THELMA G. Memory for completed and incompleting tasks as a function of personality: correlation between experimental and personality data. *J. Pers.*, 1948, **17**, 104-137.
3. ATKINSON, J. S. The achievement motive and recall of interrupted and completed tasks. *J. exp. Psychol.*, 1953, **46**, 381-390.
4. BLUM, G. S. *Psychoanalytic theories of personality*. New York: McGraw-Hill, 1953.
5. CARON, A. J. The effects of task, objective orientation, and personality factors on recall in an interrupted-task situation. Unpublished doctoral dissertation, Univer. of Chicago, 1954.
6. ERIKSEN, C. W. Defense against ego-threat in memory and perception. *J. abnorm. soc. Psychol.*, 1952, **47**, 230-235.
7. ERIKSEN, C. W. Individual differences in defensive forgetting. *J. exp. Psychol.*, 1952, **44**, 442-447.
8. ERIKSEN, C. W. Psychological defenses and "ego strength" in the recall of completed and incompleting tasks. *J. abnorm. soc. Psychol.*, 1954, **49**, 45-50.
9. GLIXMAN, A. F. Recall of completed and incompleting activities under varying degrees of stress. *J. exp. Psychol.*, 1949, **39**, 281-295.
10. GUILFORD, J. P. *An Inventory of Factors STDCR*. Beverly Hills, Cal.: Sheridan Supply, 1940.
11. KNUTSON, A. L. Personal security as related to station in life. *Psychol. Monogr.*, 1952, **66**, No. 336.
12. LEWIN, K. *A dynamic theory of personality*. New York: McGraw-Hill, 1935.

13. McARTHUR, C. Personalities of public and private school boys. *Harvard Educ. Rev.*, 1954, **24**, 256-262.
14. McARTHUR, C. Personality differences between middle and upper classes. *J. abnorm. soc. Psychol.*, 1955, **50**, 247-254.
15. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
16. RAPAPORT, D. *et al. Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945.
17. ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistent and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64-74.
18. ROSENZWEIG, S., & SARASON, S. An experimental study of the triadic hypothesis: reaction to frustration, ego defense, and hypnotizability. I. Correlational approach. *Charact. & Pers.*, 1942, **11**, 1-20.
19. SCODEL, A. Passivity in a class of peptic ulcer patients. *Psychol. Monogr.*, 1953, **67**, 1-15.
20. UNDERWOOD, B. J. Learning and retention. In C. P. Stone (Ed.), *Annual review of psychology*, Vol. III. Stanford: Annual Reviews, 1953.
21. WECHSLER, D. *The measurement of adult intelligence*. (3rd ed.) Baltimore: Williams & Wilkins, 1944.
22. ZEIGARNIK, BLUMA. Ueber das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, **9**, 1-85.
23. ZELLER, A. F. An experimental analogue of repression. I. Historical summary. *Psychol. Bull.*, 1950, **47**, 39-51.

Received June 17, 1957. Prior Publication.