Toward Improving the Magnitude of Relationships Between Test Anxiety.

Using the same sentence materials combined with different procedures, reliable test anxiety-memory relationships were generated by not constraining the initial encoding strategy or by placing time pressure at recall. The pattern of correlations was consistent with the model's predictions and indicated the mediating role of worry in comprehension. Encoding efficiency was more highly related to memory than test or state anxiety. The magnitude of component relationships was consistent with prior research. (Author)
Toward Improving the Magnitude of Relationships between Test Anxiety and Sentence Memory

John J. Hedl, Jr.
The University of Texas
Health Science Center at Dallas

James Bartlett
The University of Texas at Dallas

Two additional studies in long-term sentence memory were conducted to determine if certain critical relationships predicted by a cognitive model of test anxiety could be strengthened. Using the same sentence materials combined with different procedures, reliable test anxiety - memory relationships were generated by not constraining the initial encoding strategy or by placing time pressure at recall. The pattern of correlations was consistent with the model's predictions and indicated the mediating role of worry in comprehension. Encoding efficiency was more highly related to memory than test or state anxiety. The magnitude of component relationships was consistent with prior research.
Anxiety and Sentence Memory

Toward Improving the Magnitude of Relationships between Test Anxiety and Sentence Memory

The current theoretical model used to explain the effects of test anxiety on complex cognitive tasks is a cognitive one (Mandler, 1975; Sarason, 1972; 1980; Wine, 1971; 1980). In general, the model can be viewed as a chain of events depicted as follows:

- Test anxiety
- Increased Worry
- Evalitative Stress
- Increased Emotionality
- Decreased Performance

This model includes components of a more general trait-state anxiety model (Spielberger, 1966; 1972; 1975) and includes Liebert and Morris' (1967) bi-dimensional distinction (and instrumentation) between worry and emotionality in terms of state (momentary) anxiety. In the domain of memory, it has also been suggested that anxiety effects should be examined according to a three-stage information processing model (Benjamin et al., 1981; Eysenck, 1977; Tobias, 1977, 1980). In this view anxiety can have separate effects on encoding, storage, and retrieval operations.

Unfortunately, there have been few attempts to explore fully the proposed theoretical model. In particular, the mediating role of worry in decreasing cognitive performance has seldom been examined explicitly in conjunction with level of test anxiety. In addition, many studies of test anxiety have used tasks (e.g., paired associates, anagrams) which may not be totally relevant to
Anxiety and Sentence Memory

school learning. Our goal in the present research program is to examine the applicability of the cognitive-attentional model to a "comprehension" task (Jenkins, 1974), explicitly examining the links between (a) test anxiety and worry, (b) worry and lowered comprehension, and (c) lowered comprehension and subsequent memory.

Two initial studies (Hedl & Bartlett, 1981) dealt with long-term memory for sentences such as the following:

The house turned to water because the fire got too hot.

The haystack was important because the cloth ripped.

We assessed students' comprehension of the sentences as they were presented by scoring whether or not they determined the underlying inference or elaboration (e.g., igloo, parachute) of each sentence. This paradigm enabled us to examine the relationships of anxiety and worry to encoding, and of encoding to subsequent long-term memory for the sentences. These studies showed reliable correlations between test anxiety and worry, worry and encoding, and encoding and memory performance. This chain of correlations was insufficient, however, to generate an overall correlation between test anxiety and performance. With this chain of correlations an especially weak link was that between worry and encoding efficiency or comprehension ($r = -.32$).

Possibly the correlation between anxiety and memory might be strengthened in situations where students pursue input encoding strategies of their own choice. In our earlier studies, encoding strategy was controlled by having the students perform a semantic task during the input list of sentences. This may have reduced performance variation and thereby lowered the observed relationships. In Experiment 1 of this research, students were not given a semantic encoding task to employ.
Experiment 1: Intentional Learning During Sentence Presentation

Subjects
A total of 56 students (28 males; 28 females) participated in Experiment 1. All students received course credit for their participation. The Test Anxiety Inventory (TAI: Spielberger et al., 1980) was used to assess the level of test anxiety for these students. The data is reported by the combined groups since preliminary analyses indicated no important recall differences between the men and women.

Measures of Worry and Emotionality
The revised Worry and Emotionality Questionnaire (WEQ: Morris, Davis, & Hutchings, 1981) was used to assess the intensity levels of students' cognitive concern (worry) about their performance and physiological arousal (emotionality) during the sentence presentations and recall.

Stress Instructions
Stress instructions were adapted from Sarason's (1961; 1972; 1973) ego-involving ones and emphasized that the ability to process and remember sentences was an important indicator of general intellectual ability, and that it was important to do one's best so that one's recall scores would accurately reflect one's ability in relation to others.

Experimental Materials
A list of 31 sentences, including seven filler sentences to control for primacy and recency effects, was selected based on pilot data from 22 students. These sentences, modeled after Till et al. (1977) and Auble,
Anxiety and Sentence Memory

Rollins, and Soraci (1979), were of the form:

The house turned to water because the fire got too hot.

The haystack was important because the cloth ripped.

These sentences were developed such that certain inferences or elaborations would be probable for the student engaged in an "elaborate semantic" processing strategy (Craik & Tulving, 1975). Recall was tested with inference cues: igloo was the inference for the first example; parachute for the second. Sentence subjects (e.g., house, haystack) were also used as cues.

Procedure

The experiment was conducted in small groups (about 10 per session). After being given a general orientation, the students were instructed that their primary task was to learn a list of sentences for an upcoming recall test and that they should study them closely as they are presented. Two example sentences were then presented. In contrast to the previous two studies, no semantic orienting task was prescribed for these students.

The sentences were presented via tape recorder. One-half of the sentences were followed by a four-second blank time; the other half by a 12-second time interval. This intersentence time variable produced no effects, and we will not consider it further in this report.

After presentation of the sentences, students were given approximately five minutes for free recall. Instructions for the inference cue test were then read aloud. Students were given a definition of an inference cue and shown two sentence examples along with the cues. These inference cues were then read aloud by the experimenter and about 20-30 seconds was allowed for
the students to write down the actual sentences as presented. Only the 24 target sentences were tested. The order of the cues for the inference test was randomly determined with the constraint that the 4-12 second sentences were tested equally often in the first half and second half of the test list. A second test was developed - the second half of the first test now became the first half and vice versa. An equal number of males and females received each test.

The subject cue test was then administered in a similar fashion. The development of this test followed the same procedures used for the inference cue testing.

Following cued recall, students completed two versions of the WEQ (Morris et al., 1981). One version asked the students to describe their feelings during the sentence presentations, the other to describe their feelings during the cued recall tests. The order of the scales was counterbalanced against sex and tape. The two state measures will not be examined separately since they were very highly intercorrelated ($r = -.88; p < .001$).

Students were then asked to describe, in writing, any strategy(ies) they employed to study and remember the sentences. They were then debriefed and given the opportunity to ask questions.

Scoring

Cued recall sentence responses were scored in terms of the lenient criteria established by Till et al. (1977). That is, sentences were scored correct if they were recalled verbatim, contained synonymous substitutions or omissions of partially redundant information, or specified the agent, verb, and some of the additional information. Reliability of two raters was .94 for a subset of the data.
Results and Discussion

The reported strategies were classified by type using Weinstein et al.'s (1979) learning strategy taxonomy (imagery, verbal elaboration, physical strategies, rote, none). Preliminary analyses indicated that an imagery strategy led to higher free recall (8%) than the others (4%) combined ($F (4,51) = 3.22, p < .02$; to higher inference cue recall (19%) than the others (6%) combined ($F (4,51) = 5.48, p < .001$; and to higher subject cue recall (16%) than the others (5%) combined ($F (4,51) = 5.60, p < .001$). Scheffe's tests also revealed little differences between the other strategies. In general, imagery use led to similar overall recall levels in comparison to the semantic encoding strategy (e.g., derive the main implication for each sentence) used in the two previous studies (Hedl & Bartlett, 1981).

Unlike the previous two studies, the correlations between TAI and memory performance (cued recall) are now significant ($p < .05$), although not large. The correlations involving encoding efficiency also form the same pattern as before. That is, worry is significantly related to encoding, emotionality is not. Encoding efficiency was also strongly related to the three recall measures (mean $r = .50$). These encoding-memory correlations are smaller (significantly so in one instance) than the previous study which used the same sentence materials. This might be due to the use of an indirect encoding measure, to the lower overall performance in this study, or to a combination of both factors.
Overall, these correlational findings suggest that significant test anxiety/memory relationships can be found when a specific, semantic encoding task is not provided to the student. However, the test anxiety/memory relationships appear to be independent of differential strategy selection in use by high and low-test anxious students. Using the same TAI cut-offs as the previous two studies, the frequency of reported strategies was examined and the resultant data was:

<table>
<thead>
<tr>
<th>Verbal Elaboration</th>
<th>Physical Strategies</th>
<th>Rote Strategy</th>
<th>No Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High TAI Students</strong></td>
<td>5</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td><strong>Low TAI Students</strong></td>
<td>5</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

The chi square comparing usage was not significant ($x^2 = 2.93, df = 4$). It may be, however, that high test-anxious students did not implement the strategies as effectively as did the low test-anxious students.

**Experiment 2: Time Pressure at Test (Recognition and Recall)**

In Experiment 2, a different strategy was used to examine the possibility of generating stronger test anxiety and worry effects in this sentence comprehension task. Rather than focusing on strategies during sentence presentations, we examined the effects of placing time pressure during testing. Recall procedures are not ideal for manipulating time pressure at test due to differences in writing time for sentences of varying lengths. For this reason, we used a recognition procedure and the test contained inference
cues. The purpose of the test was to assess the students' ability to recognize probable inference cues from the sentences which were presented.

To raise performance levels we adopted the same presentation strategy as in the second experiment of the prior research (Medi & Bartlett, 1981). That is, each sentence was read aloud and repeated with a 10-second intersentence interval and a semantic orienting task was used. We also included an inference cue recognition and recall test under nonspeeded conditions. Free recall was assessed as before.

It was reasoned that time pressure during testing could enhance test anxiety effects for either of two reasons. First, worry could now plausibly reduce retrieval efficiency leading to a larger deficit when time pressures are present. Second, with severe time pressure there could be limited opportunity for flexible retrieval operations, ones that could make up for the poor encoding initially. The suggestion would be that encoding efficiency could now be a stronger factor in the model.

Method and Procedure

Essentially the same procedures were followed in Experiment 2. The 31 sentences were presented to forty-four college students (20 males; 24 females). Each sentence was read and repeated to minimize errors of speech perception. The students' task at encoding was to determine the main implication of each sentence, a semantic encoding strategy. Comprehension (encoding efficiency) scores were derived from the students' written responses to sentences of the input list. These responses were scored as correct if the students listed the main inference (toast) or one that represented some understanding of the sentences (e.g., party, honor, etc.).
As in Experiment 1, free recall data was collected immediately after the sentence presentations. The next test, however, was a timed recognition test using the 24 inference cues for the target sentences and 24 lure cues. Lure cues contained the same number of syllables and had the same first letter as their corresponding inference cues. Lures were unrelated to probable inferences from the target sentences. The cues were presented one at a time on a slide projector. Students were given five seconds to indicate whether or not the cues reminded them of the sentences presented. They indicated their recognition response by checking Yes, No, or Don't Know on their test sheet. Students were then given the opportunity to recognize the cues and recall the actual sentences under more relaxed conditions (20-30 seconds as in Experiment 1 and the previous experiments).

Free and cued recall sentence responses were again scored by the lenient criteria established by Till et al. (1977). However, we will only consider the recognition data from Experiment 2. TAI and state worry/emotionality measures were also collected similar to Experiment 1.

Results and Discussion.

To simplify the comparisons with Experiment 1 and the prior work, only the correlational analyses from Experiment 2 will be presented. Table 2 summarizes the most important correlations among the anxiety and memory measures as well as their theoretical significance from a cognitive processing perspective. The test anxiety relationships with recognition performance were significantly higher in the speeded test condition for both recognition accuracy ($A'$) and response bias ($B'$), than those under the more relaxed testing situation. Although not significant, the same pattern was observed with the hits - false alarms measure of recognition accuracy ($p < .10$).
The cognitive chain is also supported by these data. Test anxiety was significantly related to worry and emotionality at encoding ($p's < .01$). As with the three previous studies worry at encoding was negatively related to encoding efficiency. The relationship between emotionality and encoding efficiency also was negative, but not significant.

As predicted, the encoding efficiency - recognition memory link was quite robust ($r = .69, p < .01$) in the speeded test condition. However, this link was weaker under more relaxed testing in comparison to the speeded test condition ($t = 3.91, df = 41, p < .001$). A similar pattern was found for the response bias measure and hits - false alarms.

This might suggest that under relaxed testing conditions students have the opportunity to practice flexible retrieval strategies which can compensate for poor encoding initially.

**General Discussion**

The experiments suggest two important conclusions about test/state anxiety and cognitive processing. First, the results are consistent with the cognitive/attentional model described earlier. Worry does appear to have a small mediating role within a multi-stage cognitive model (encoding, storage, retrieval processes). Trait test anxiety will predict worry at encoding, worry is related to encoding efficiency, and emotionality is unrelated to encoding. Sentence encoding (comprehension) is a strong, but less than perfect, predictor of free recall, recognition, and cued recall.

Therefore, the suggestions to consider the effects of anxiety in combination with stages of cognitive processing appear to be reasonable (Mueller, 1980; Tobias, 1977, 1980). We are currently examining the
Anxiety and Sentence Memory

structural relationships among the anxiety, encoding, and memory measures via path analytic techniques to further explore the cognitive/attentional model predictions.

A second conclusion suggested by this research is that reliable test anxiety - sentence memory relationship can be obtained when initial encoding strategies are not constrained (Experiment 1) or when time pressure is applied during testing (Experiment 2). This relationship was not significant in our earlier studies (Hedl & Barltett, 1981), which neither constrained initial encoding nor applied time pressure at test. Further, the relationship was not significant in the relaxed-testing condition of Experiment 2.

These findings suggest several implications for the teaching/learning process. First, educational strategies that focus on comprehension of material may attenuate the negative effects of test anxiety and worry on recognition and recall test performance. Recent evidence argues that the test anxiety deficit can be partially explained by ineffective learning (encoding) strategies initially (Benjamin et al., 1981). Wine (1980) also makes the point that task-relevant strategies at encoding can be beneficial to high test-anxious students.

Second, additional research is needed on test anxiety effects as related to time pressure during testing. Experiment 2 produced the interesting finding that relaxed testing conditions reduce correlations between encoding efficiency and recognition (as well as those between test anxiety and recognition). This finding has potentially quite general implications for the use of timed versus untimed (power) tests in educational situations, as well as for test anxiety effects. The finding might also have theoretical
implications for the nature of retrieval processes in recognition (Mandler, 1980). However, any interpretation of our results must be tentative, since the order of the testing conditions in Experiment 2 (speeded, relaxed) was not counterbalanced—the speeded test always was given first. We are planning additional research which will remove this confounding. Regardless of our results, the dependence which holds between encoding activities and retrieval under different test conditions is a crucial problem for educationally relevant research.

In the last two experiments in which the test anxiety-performance correlations were significant, the magnitude of the relationships was quite modest. The similarity of these results with previous research is striking, however. Our test anxiety and performance correlations of about -.25 are quite similar to other studies using a variety of dependent measures (Deffenbacher, 1977; mean r = -.27). The pattern of worry/performance relationships was consistent across the four studies and also similar with prior research. Our worry and performance correlations (r's about -.30) are similar to that reported for anagrams, Miller Analogy Test performance, or classroom performance (See Deffenbacher (1980) for a recent review). Morris et al. (1981) also noted that their WEQ worry accounts for approximately 10 percent of the variance in performance across a number of experiments. They also noted that many factors contribute to performance differences among students in addition to worry.

Lastly, the correlations between test anxiety and worry, and worry and encoding efficiency are considerably less than those between encoding efficiency (semantic or imagery strategies) and memory performance. These
data suggest that within relatively normal school populations, trait and state personality variables can improve our predictive efficiency with regard to basic cognitive processing tasks. But the impact of effective encoding strategies is considerably greater. It may also be that more deviant populations will need to be studied to examine the upper and lower limits of worry and cognition. A similar argument has been advanced with regard to the use of mildly depressed college students to study the effects of depression on a variety of cognitive processing tasks (Depue & Monroe, 1978).
References


Footnotes

1. Paper presented at the meeting of the American Educational Research Association, New York, March 1982. The authors wish to thank Frank Burns for his assistance in collecting and reducing data for analysis. The authors would also like to thank Patsy Moore and Terri Flowers for typing, editing, and clerical efforts associated with the studies. Requests for copies should be directed to John J. Hedl, Jr., School of Allied Health Sciences, The University of Texas Health Science Center at Dallas, Dallas, Texas 75235.

2. The strategies reported by the students after the recall tests were scored in the following categories:

- **Rote Strategies** - strategies that emphasized repetition
- **Physical Strategies** - any use of the physical properties of the sentences (e.g., spelling patterns, word patterns, key words)
- **Imaginal Elaboration** - any use of images and/or formation of mental pictures
- **Verbal Elaboration** - any active work with the sentences such as asking or answering questions, determining implications/inferences, or relating the sentence to already known information
- **None** - no strategy reported
3. For comparison of these data to our previous work it becomes convenient to use imagery use (yes, no) as an estimate (albeit crude) of encoding efficiency.

4. The recognition test was developed with similar constraints as the Inference and subject cue tests used in the prior studies. Half of the sentences presented in the first half of the input list were tested in the first half of test recognition test and half were tested in the second part of test. The second half of the input list was similarly split across the two test halves. Within a given test half, 12 lure cues were also included. The order of the target and lure cues was random within each test half with the constraint than no more than four cues of a particular kind could be presented consecutively. For approximately one-half of the students the second block of the recognition test was presented first. The recall test used the same slide order as the recognition test for a given student.
Table I

Intercorrelations of the Anxiety, Encoding, and Memory Measures from Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Present Study</th>
<th>Hedl &amp; Bartlett (1982)</th>
<th>Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAII with worry at encoding</td>
<td>.32 *</td>
<td>.40 **</td>
<td>.54 (n.s.)</td>
</tr>
<tr>
<td>with emotionality</td>
<td>.36 **</td>
<td>.50 **</td>
<td>1.02 (n.s.)</td>
</tr>
<tr>
<td>Worry with Encoding</td>
<td>(.27) **</td>
<td>(.32) **</td>
<td></td>
</tr>
<tr>
<td>Emotionality</td>
<td>(.15) *</td>
<td>(.16)</td>
<td></td>
</tr>
</tbody>
</table>

**Encoding Efficiency and Memory**

| Free Recall                          | (.43) **      | .65 **                  | 1.87 (n.s.) |
| Inference Cue Recall                 | (.49) **      | .74 **                  | 2.45 (p < .05) |
| Subject Cue Recall                   | (.58) **      | .71 **                  | 1.33 (n.s.) |

**TAII and Memory**

| Free Recall                          | -.22          | .03                     |         |
| Inference Cue Recall                 | -.29          | -.13                    |         |
| Subject Cue Recall                   | -.31          | -.18                    |         |

*p < .05; **p < .01

Using a point-biserial correlation with imagery use (yes, no) as an estimate of encoding efficiency.

All recall measures represent total performance since preliminary analyses indicated no rate differences.
Table 2

Intercorrelations of the Anxiety, Encoding, and Recognition Memory

Memory Measures from Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Recall Test Conditions</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speeded</td>
<td>Relaxed</td>
<td>t</td>
</tr>
<tr>
<td>TAI with Worry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>with emotionality</td>
<td>.51 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.53 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worry with encoding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception with Encoding</td>
<td>- .30 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- .21 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worry and Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition Accuracy (A')</td>
<td>- .38 *</td>
<td>- .25</td>
<td></td>
</tr>
<tr>
<td>Response Bias (B')</td>
<td>- .27</td>
<td>- .17</td>
<td></td>
</tr>
<tr>
<td>Hits - False Alarms</td>
<td>- .44 **</td>
<td>- .23</td>
<td></td>
</tr>
<tr>
<td>TAI and Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition Accuracy (A')</td>
<td>- .32 *</td>
<td>- .20</td>
<td></td>
</tr>
<tr>
<td>Response Bias (B')</td>
<td>- .37 *</td>
<td>- .21</td>
<td></td>
</tr>
<tr>
<td>Hits - False Alarms</td>
<td>- .26</td>
<td>- .19</td>
<td></td>
</tr>
<tr>
<td>Encoding Efficiency with Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition Accuracy (A')</td>
<td>.69 **</td>
<td>.35 **</td>
<td>3.91 **</td>
</tr>
<tr>
<td>Response Bias (B')</td>
<td>.55 **</td>
<td>.32 *</td>
<td>2.05 *</td>
</tr>
<tr>
<td>Hits - False Alarms</td>
<td>.74 **</td>
<td>.48 *</td>
<td>4.02 ***</td>
</tr>
</tbody>
</table>

*p < .05; **p < .01; *** p < .001

See Grier (1971) for the calculations of the recognition measures. A' and B' are the nonparametric versions of d' and B.