This study was designed to answer two questions: (1) Is the systematic desensitization of test anxiety effective with secondary school students? (2) Is relaxation per se as useful a technique as systematic desensitization? High test anxious secondary-school students were assigned to one of two experimental conditions, desensitization or relaxation, and met for 20 minutes daily for a period of 6 weeks. Their results were compared to those of a no treatment control group. It appeared that the experimental subjects (Ss) underwent a significant reduction in anxiety. However, only among grade 13 Ss was there a trend toward improved academic performance. (Note: In the Ontario School System, grade 13 is for students planning to attend university.) The relaxation per se treatment was more effective in reducing general anxiety than systematic desensitization. The results and their implications for future research were discussed. (Author)
Abstract

This study was designed to answer two questions: 1) Is the systematic desensitization of test anxiety effective with secondary school students; 2) Is relaxation per se as useful a technique as systematic desensitization.

High test anxious secondary school students were assigned to one of two experimental conditions, desensitization or relaxation, and met for 20 minutes daily for a period of 6 weeks. Their results were compared to those of a no treatment control group. It appeared that the experimental Ss underwent a significant reduction in anxiety, however, only among grade 13 Ss was there a trend toward improved academic performance. The relaxation per se treatment was more effective in reducing general anxiety than systematic desensitization.

The results and their implications for future research were discussed.
Over the past decade evidence has accumulated to show that there is a negative correlation between scores on test anxiety scales and academic performance (I. G. Sarason, 1961; Mandler & Sarason, 1952; Alpert & Haber, 1960). It has been suggested that, in situations involving implied personal threat or fear of failure, high test anxious Ss evoke task irrelevant and defensive responses such as blocking or trembling that impair performance (Mandler & Sarason, 1952; Child, 1954). There is evidence, too, that the retarded performance of high test anxious students in threatening situations does not necessarily indicate inferior intelligence because high anxious Ss perform comparably with low anxious Ss in non-threatening situations (Mandler & Sarason, 1952; Sarason, 1956; Sarason, 1961; Sarason, 1963).

From these findings, it has been postulated that in a threatening situation, lowering the arousal level of high test anxious Ss should reduce the number of task irrelevant responses, increase the proportion of task relevant responses and consequently cause a concomitant increment in performance.*

* In their study of test anxiety, Alpert & Haber (1960) proposed a bidimensional theory consisting of the debilitating factor (task-irrelevant responses) suggested by Mandler & Sarason (1952) and a facilitating factor (task-relevant responses). These two anxiety dimensions have a low negative correlation (about -.35) and are therefore relatively independent of each other. Persons displaying a large number of task irrelevant responses do not necessarily have very many task relevant responses and vice versa.
This hypothesis has been tested in several studies using desensitization procedures (Katahn, Strenger & Cherry, 1966; Paul & Shannon 1966; Emery & Kronholtz, 1967; Donner, 1968). They found that when high test anxious university students were exposed to desensitization, there was a reduction in anxiety and an improvement in grade point averages. The desensitization procedures employed in these studies were based on the reciprocal inhibition principle proposed by Wolpe (1958, 1961). He suggests that if a response which inhibits anxiety such as relaxation can be made to occur in the presence of anxiety-evoking stimuli it will weaken the bond between these stimuli and anxiety. Implicit in this theory is the assumption that some proportion of the stimuli that normally elicit anxiety can be simulated in a clinic or laboratory. Thus a new association formed between simulated testing situations and relaxation responses should inhibit to some extent the old connections between anxiety elicited in the actual testing situation. One may also expect that the newly learned relaxation response could generalize to similar situations. Since the generalization gradient for avoidance responses is steep, a major change in general behavior is not to be expected. Wolpe (1958) also proposes that the reciprocal inhibition procedure should be most successful where specific stimuli elicit anxiety responses. When anxiety responses have generalized to a large number of different stimuli (or what Wolpe calls pervasive anxiety), S's progress through the anxiety hierarchy is retarded and the change in behavior may be transitory (Lang, 1964).

Previous research with test anxious high school students has shown that they have a high level of general anxiety (Laxer & Quarter, 1967). This raises the question of the appropriateness of systematic desensitization procedures for generally anxious Ss as opposed to Ss with specific anxiety.
It might be expected that relaxation exercises per se (without the anxiety hierarchy) would be more effective than systematic desensitization to specific stimuli for Ss with pervasive anxiety. A general relaxation procedure might still establish associations with the large number of stimuli that formerly elicited anxiety.

The comparative effects of relaxation per se and desensitization were studied by Davison (1966). Even though the results of this study supported Wolpe's association theory and showed a marked superiority for Ss in the desensitization group, Ss were treated for a specific phobia, a fear of non-poisonous snakes. It is still questionable whether desensitization is a more effective treatment than relaxation per se for Ss whose anxiety is more general.

The present study therefore had two aims: 1) To determine whether the reciprocal inhibition procedure would be successful with test anxious students in secondary schools; and 2) To examine the comparative effects of desensitization and relaxation per se on these Ss.

To assess the effect of the treatments, debilitating and facilitating test anxiety scores, manifest anxiety scores, and school marks were observed. It was assumed that treatments would be considered successful for the experimental Ss if, relative to the control Ss, their level of debilitating test anxiety decreased and their level of facilitating test anxiety increased. Manifest anxiety was used to determine whether any changes in test anxiety generalized to other anxiety provoking situations. School grades were examined to see whether Ss who reduced their anxiety level were able to improve their grade point average.
Method

Subject Selection

The following procedure was used to select Ss for the experiment. All students in four secondary schools in the Metro Toronto Area took the Alpert-Haber Achievement Anxiety Scale (AAS) and the Taylor Manifest Anxiety Scale (MAS) in classroom groups. Students were divided into groups by school, grade and sex. Then the raw data for each variable was converted to a Z distribution. A list of students whose scores met the following criteria was compiled.

a) a Z-score ≤ 1.00 on the debilitating test anxiety scale
b) a Z-score ≥ 1.00 on the facilitating test anxiety scale
c) an average % that was less than B, but high enough so that the student had some hope of passing his year.

The record of each student who fit the above mentioned criteria was reviewed by a counsellor. If there was no evidence of unusual exam tension in the student's history, his name was deleted from the list. A counsellor interviewed the remaining students and the list was further reduced to those who felt that their problem was severe enough to merit participation in a remedial program. Of these students only those who received parental consent participated in the experiment.

The experiment was divided into 2 sections*.

Study A - students in grades 9 - 12
Study B - students in grade 13

* In the Ontario School System Grade 13 is for students planning to attend university. There is a special curriculum and province-wide exams at the end of the year, whereas exams for students in grades 9 - 12 are set by individual schools.
Study A: Within each school Ss were assigned randomly to one of the three treatment groups (relaxation, desensitization and control). There was a total of 33 Ss in each of the three treatments. For purposes of analysis the four schools were grouped together. Debilitating test anxiety, facilitating test anxiety, manifest anxiety, and grade point average were measured both prior to and after completion of the treatments. Grade point average was divided into 3 components: 1) languages and social sciences, 2) maths and sciences, and 3) non-academic subjects.

Study B: The second study included only Ss in grade 13. There were 2 groups, desensitization and control. Fifteen Ss were randomly assigned to the desensitization and control groups as in study A. The same anxiety and academic variables were observed before and after completion of the programme.

Counsellor Training

After a careful screening procedure, four counsellors who were university graduates with teaching experience were hired to administer the remedial programme. The counsellors were given a seven-day training period which included a theoretical overview and several daily practice sessions.

Procedure

One counsellor was assigned to each of four schools. Small groups of two to four students were scheduled according to their timetables. Males and females met in separate groups. In three schools Ss met daily for twenty-minute sessions while in the fourth school there were two twenty-minute sessions per day. All students were encouraged to practice regularly at other times of the day. It was impossible to keep an accurate record of whether Ss co-operated in this regard.
The remedial programme was started six weeks prior to the final exams and continued until the exams began. Students were also encouraged to meet with their counsellors for a relaxation session at the start of each exam, although, in many cases, this was not possible. Each S was retested during the exam week with the AAS & the MAS. S's grade-point average for the final examinations was also recorded. Ss who participated in the remedial programme were given an opportunity to assess the project and their counsellor through a confidential questionnaire.
Stage I  Pretraining - Days 1 and 2

Ss in both the relaxation and desensitization conditions learned to relax while lying on floor mats. They began with their feet and continued in sequential order finishing with their face. S was trained to think of the muscle group, be aware of the strain when the muscles were tensed and to concentrate on the feeling of relaxation as tensing was eased. These exercises were designed to build up a sense of kinesthetic feedback.

Stage II  Relaxation - Days 3 to 5

The procedure of stage II was the same as stage I except that the tensing portion of the exercises was excluded. Ss learned to relax their muscles in the same sequential order as stage I without tensing them first. Any S who had difficulty relaxing in stage II was given additional practice on an individual basis. If necessary, the tensing exercises used in stage I were repeated. Ss in the relaxation condition continued with the stage II procedure throughout the remainder of the experiment (with a minor variation that is described in stage IV) whereas Ss in the desensitization condition moved onto stage III and IV.

1. We would like to express our appreciation to Douglas Quirk of the Clarke Institute of Psychiatry, Toronto, for his assistance and advice. The training procedure used in the schools was adapted from tapes designed by Mr. Quirk. The modified version of the tapes used in this study can be obtained by contacting authors of this paper at The Ontario Institute for Studies in Education, Toronto.

2. The muscle clusters were: arches of the feet; calf muscles; thigh muscles; stomach and abdomen muscles; back muscles; front, back and upper shoulder muscles; upper arm and lower arm muscles; wrist; hand and neck; forehead; eyes; lips and throat.
Stage III Desensitization - Days 6 to 20

On the 6th day of training Ss in the desensitization condition were presented with the anxiety hierarchy items. Everyone used a standard hierarchy.

In the desensitization groups, the counsellor devoted the first ten minutes of each session to the stage II relaxation procedure, and the last ten minutes to the presentation of items from the hierarchy. An item was presented in the following manner. After Ss were completely relaxed the counsellor verbally described the item and asked Ss to imagine the situation it described. To make the image more vivid Ss closed their eyes. The item description was followed by regularly spaced relaxation-inducing statements. Each item was presented several times, often with slightly different wording to prevent monotony and to provide a more accurate description.

If upon presentation of an item, S felt himself tensing, he was asked to lift his forefinger. When this happened the counsellor reverted to stage II of the treatment until Ss totally relaxed. Subsequent items were not commenced until all members of the group had been desensitized to previous items.

The first four items in the hierarchy were presented on days 6 and 7. On days 8 to 20 one item was covered daily.

Stage IV Continued Desensitization - Days 21 to 30

From day 21 to 30 the remaining items in the hierarchy were presented. Since these latter items usually elicited more anxiety, a new one was initiated approximately every other day.
During days 21 to 30, Ss in both the desensitization and relaxation conditions sat in classroom desks. It was felt that this would facilitate the transfer of relaxation to a classroom context.
Results

Ss were excluded from the data analysis for any of the following reasons:

(a) dropping out of school
(b) missing over 50% of the sessions
(c) missing the posttest

Ten Ss were excluded from Study A, leaving 31 in the control group and 29 in each of the two experimental conditions. In addition, 13 Ss did not take math or science courses and the df for that criterion was reduced accordingly. For study B, (grade 13 students) 3 Ss were excluded leaving 13 in the control group and 14 in the experimental group.

The following criteria were used to evaluate the success of the program:

(1) debilitating test anxiety
(2) facilitating test anxiety
(3) manifest anxiety
(4) languages and social science grades
(5) mathematics and science grades
(6) non-academic grades (shop, physical education and drafting).

For criteria 4 to 6, S received a score out of 100 that was computed by averaging all the scores of the school subjects in that category. Not all Ss had the same number of subjects in a category. For Ss in grades 9 to 12, the posttest grade point average was computed by combining the term work and the final exam score. For grade 13 students, the final exam grade was the total posttest score.
A multiple linear regression model was used to analyze the data. For Ss in Study A, the pretests, as well as sex, grade level, and school programme type (four or five year)* were used as predictors for each of the post test criteria. For Ss in grade 13, the pretest scores and sex were the predictors. Thus it was possible to control statistically for sources of variance affecting the relationships between treatments and criteria.

**STUDY A**

---

Insert Tables 1 and 2

---

The criteria means for each of the three conditions are presented in table 1. When the means for each criterion are analyzed (cf table 2) the results show that only in the case of manifest anxiety are the treatment effects significant (p = .04). Orthogonal comparisons reveal that Ss in the relaxation condition had considerably lower manifest anxiety than the control group, whereas the difference between the relaxation and desensitization conditions and between the desensitization and control groups were not significant.

For debilitating test anxiety, the treatment effects are near-significant (p = .07). Orthogonal comparisons show that Ss in the desensitization condition had significantly lower anxiety than the control group (p = .03); Ss in the relaxation condition also tended to have lower anxiety (p = .06); and there was no difference between relaxation and desensitization Ss.

* In the Ontario School System the 5-year program is for students planning to attend university, whereas the 4-year program is generally for students who intend to terminate their education after their fourth year of secondary school.
For facilitating test anxiety, the treatments did not account for a significant portion of the variance ($p = .15$). However, the trend among the means was similar to that of manifest and debilitating test anxiety.

For the three remaining criteria, the academic variables, there was neither an indicated relationship with the treatment nor a consistent trend among the means.

**STUDY B**

Grade 13 Ss did not receive marks for non-academic subjects, therefore, criterion six was excluded from the analysis. For the remaining five criteria, the results were consistently in the predicted direction; experimental Ss showed reduced anxiety and improved academic grades relative to the control group.

---

Insert Tables 3 and 4

---

For the debilitating and manifest anxiety, the treatment effects were significant ($p < .05$), whereas for facilitating test anxiety they were in the expected direction ($p = .17$).

In the analysis of the academic variables, the treatment effects for maths and science were near-significant ($p = .07$); languages and social sciences also were in the expected direction ($p = .19$).
Discussion

This study was designed to answer two questions: 1) Is the systematic desensitization of test anxiety effective with secondary school students; 2) Is relaxation per se as useful a technique as systematic desensitization.

With respect to the former question the results show that the treatments were reasonably effective in reducing Ss anxiety level. However, only among grade 13 Ss was there any suggestion that the change in anxiety generalized to Ss academic work and even there the improvement in academic work was not significant. For Ss in grades 9 to 12, the change in the treatment conditions for all three anxiety measures, debilitating, facilitating, and manifest were in the expected direction, although not always reaching the required level of statistical significance. Academic work was not altered significantly by either of the treatments. For Ss in grade 13, there was a consistently positive change for all 5 criteria, the three anxiety variables and two academic variables. It is quite possible that with a larger number of Ss all 5 criteria might have reached the required level of statistical significance.

The finding that reduced anxiety did not lead to an improvement in academic work should not be used to criticize the treatments. Although lower anxiety appears to be a prerequisite for more efficient performance on exams, the latter depends to a large extent on Ss' desire to study. Many Ss, especially those in grades 9 to 12, seemed to think that the experimental treatment was a substitute for hard work. Some were in vocational programs where academic success was not overly important. Consequently they were not very concerned about their final exams (the post-treatment criterion).
Others were in four-year academic programs because of their inability to cope with courses leading to university (5-year program). Since the exam pressure in the 4-year course was reduced considerably, Ss' high test anxiety probably reflected their negative experiences of the past rather than their present state.

Only among grade 13 Ss was there some suggestion of improved academic work. Since their final exams served as a basis for college admission, they were highly motivated to study and succeed. Furthermore, the final grade 13 marks were the result of standardized province-wide marking, independent of the classroom teacher, whereas classroom teachers marked Ss in grades 9 to 12. By the final exams the teacher stereotype of the student might have lessened the possibility of change.

Also, it is interesting that the reduction in anxiety appears to be more marked among grade 13 Ss than those in grades 9 to 12. Since the former group have a lower level of general anxiety than the latter (Laxer & Quarter, 1967) they might have been more responsive to systematic desensitization. Other theorists (Wolpe, 1958; & Lange, 1964) also predict that systematic desensitization will be more effective in treating phobias and anxieties associated with specific stimuli, rather than generalized anxiety states.

The second question that this experiment explored was whether relaxation per se or desensitization was more successful with Ss in grades 9 to 12. Relaxation per se seemed to be more effective in reducing manifest anxiety, and equally effective in lowering debilitating test anxiety. One possible explanation for these results is that Ss in a relaxed state might inadvertently imagine anxiety provoking stimuli and form new associations.
between these stimuli and relaxation. For Ss in the relaxation per se condition the new associations might have been more general than those for Ss who were systematically desensitized.

Another program currently in progress is designed to rectify some of the present experiment's shortcomings as well as answer some of the theoretical questions raised above. The subject sample has been refined, standardized achievement tests are being employed, and the anxiety hierarchy has been extended to include the simulation of exam situations.
References


Table 1
Adjusted means for six criteria in grade 9-12 study

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Control</th>
<th>Desensit.</th>
<th>Relaxa.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. debilitating test anxiety</td>
<td>73</td>
<td>65</td>
<td>66</td>
</tr>
<tr>
<td>(range is 10-100)</td>
<td>a</td>
<td>b</td>
<td>ab</td>
</tr>
<tr>
<td>2. facilitating test anxiety</td>
<td>31</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>(range is 10 to 90)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>3. manifest anxiety</td>
<td>29</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td>(range is 0 to 50)</td>
<td>a</td>
<td>ab</td>
<td>b</td>
</tr>
<tr>
<td>4. languages &amp; social sciences</td>
<td>53</td>
<td>53</td>
<td>56</td>
</tr>
<tr>
<td>(range is 0 to 100)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>5. maths &amp; sciences</td>
<td>63</td>
<td>60</td>
<td>63</td>
</tr>
<tr>
<td>(range is 0 to 100)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>6. non academic</td>
<td>48</td>
<td>51</td>
<td>50</td>
</tr>
<tr>
<td>(range is 0 to 100)</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

* For each criterion cells having different subscripts are significantly different at .05 level.

** For criteria 1 and 3 it was expected that if the treatments were successful, Ss in the control group would have the highest means, whereas for criteria 2, 4, 5, and 6 the opposite was predicted.
Table 2
Multiple Linear Regression Analysis
for Ss in grades 9 to 12

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>$r^2_f$</th>
<th>$r^2_r$</th>
<th>$0.ff$</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. debilitating test anxiety</td>
<td>.26</td>
<td>.21</td>
<td>.05</td>
<td>2,77</td>
<td>2.69</td>
<td>.07</td>
</tr>
<tr>
<td>2. facilitating test anxiety</td>
<td>.30</td>
<td>.27</td>
<td>.03</td>
<td>2,77</td>
<td>1.91</td>
<td>.15</td>
</tr>
<tr>
<td>3. manifest anxiety</td>
<td>.52</td>
<td>.48</td>
<td>.04</td>
<td>2,77</td>
<td>3.25</td>
<td>.04</td>
</tr>
<tr>
<td>4. languages and social sciences</td>
<td>.47</td>
<td>.46</td>
<td>.01</td>
<td>2,77</td>
<td>0.69</td>
<td>.50</td>
</tr>
<tr>
<td>5. maths and sciences</td>
<td>.41</td>
<td>.39</td>
<td>.02</td>
<td>2,64</td>
<td>0.91</td>
<td>.40</td>
</tr>
<tr>
<td>6. non-academic subjects</td>
<td>.55</td>
<td>.53</td>
<td>.02</td>
<td>2,77</td>
<td>1.28</td>
<td>.28</td>
</tr>
</tbody>
</table>

$r^2_f$ is the total criterion variance accounted for by all the predictors

$r^2_r$ is the portion of criterion variance accounted for by all prediction minus the treatments

0.00 is the portion of criterion variance due to treatments
Table 3

Adjusted means for 5 criteria in grade 13

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Control</th>
<th>Desens.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. debilitating test anxiety</td>
<td>81</td>
<td>69</td>
</tr>
<tr>
<td>(range 0 to 100)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>2. facilitating test anxiety</td>
<td>34</td>
<td>38</td>
</tr>
<tr>
<td>(range is 0 to 90)</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>3. manifest anxiety</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>(range is 0 to 50)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>4. languages &amp; social sciences</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>(range is 0 to 100)</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>5. maths &amp; sciences</td>
<td>51</td>
<td>57</td>
</tr>
<tr>
<td>(range is 0 to 100)</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

* For each criterion cells having different subscripts are significantly different at .05 level.

** For criteria 1 and 3 it was expected that if the treatments were successful Ss in the control group would have the highest means, whereas for criteria 2, 4, & 5 the opposite was predicted.
Table 4

Multiple Linear Regression Analysis
for Ss in grade 13

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>$r^2_f$</th>
<th>$r^2_r$</th>
<th>$0.\text{ff}$</th>
<th>df</th>
<th>f</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. debilitating test anxiety</td>
<td>.47</td>
<td>.34</td>
<td>.13</td>
<td>1,20</td>
<td>5.09</td>
<td>.03</td>
</tr>
<tr>
<td>2. facilitating test anxiety</td>
<td>.37</td>
<td>.31</td>
<td>.06</td>
<td>1,20</td>
<td>2.03</td>
<td>.17</td>
</tr>
<tr>
<td>3. manifest anxiety</td>
<td>.69</td>
<td>.58</td>
<td>.11</td>
<td>1,20</td>
<td>6.72</td>
<td>.01</td>
</tr>
<tr>
<td>4. languages &amp; social sciences</td>
<td>.70</td>
<td>.67</td>
<td>.03</td>
<td>1,20</td>
<td>1.79</td>
<td>.19</td>
</tr>
<tr>
<td>5. maths &amp; science</td>
<td>.80</td>
<td>.77</td>
<td>.03</td>
<td>1,20</td>
<td>3.55</td>
<td>.07</td>
</tr>
</tbody>
</table>

$r^2_f$ is the total criterion variance accounted for by all the predictors

$r^2_r$ is the portion of criterion variance due to all the predictions-treatments

$0.\text{ff}$ is the portion of the criterion variance due to treatments