### Title and Subtitle
Pretask versus within-task anxiety measures in predicting performance on a concept acquisition task

### Author(s)
Tennyson, K.D.; Boutwell, R.C.

### Abstract
This paper describes a study designed to investigate whether within-task measures of the anxiety of a learner are significantly better predictors of learner performance than pretask trait or state measures of anxiety. A time series design was used to obtain the subjects' typical anxiety level during a task that involved identifying RX2 atom crystals. Regression analyses of the study findings showed that the within-task anxiety measure was a significant predictor of performance. Repeated measurements of trait and state anxiety demonstrated that anxiety fluctuated over time, and that environmental changes affected state anxiety. The implications of the study are that aptitude-treatment interactions using within-task measures may be more useful in designing adaptive instruction than the current notion of using pretask measures of anxiety.
LABORATORY

Computer
Applications
Laboratory

PRE 'ASK VERSUS WITHIN-TASK
ANXIETY MEASURES IN PREDICTING PERFORMANCE
ON A CONCEPT ACQUISITION TASK

Robert D. Tennyson and Richard C. Boutwell

August 1, 1973
Working Paper No. 1

To be Published in the Journal of
Educational Psychology

This document is a prepublication draft designed for limited circulation.
The purpose of the Working Paper series is to disseminate information
resulting from activities conducted in the CAL. These papers range from
scientific reports on basic research to nontechnical reports on applied
issues in education. Please do not cite as a source without consulting
the Laboratory and the authors of this document.

Center for Educational Technology
College of Education
Florida State University
Tallahassee, Florida 32306
PRETASK VERSUS WITHIN-TASK
ANXIETY MEASURES IN PREDICTING PERFORMANCE
ON A CONCEPT ACQUISITION TASK

Robert D. Tennyson and Richard C. Boutwell

August 1, 1973
Working Paper No. 1

To be Published in the Journal of Educational Psychology

This document is a prepublication draft designed for limited circulation. The purpose of the Working Paper series is to disseminate information resulting from activities conducted in the CAL. These papers range from scientific reports on basic research to nontechnical reports on applied issues in education. Please do not cite as a source without consulting the Laboratory and the authors of this document.
The purpose of this experiment was to investigate the hypothesis that within-task state anxiety measures are significantly better predictors of learner performance than pretask trait or state anxiety measures. Using a time series design to obtain subjects' typical anxiety level and a crystal identification task, the regression data analysis showed that the within-task anxiety measure was a significant predictor of performance ($p < .01$). The repeated measurements of trait and state anxiety demonstrated the fluctuation of anxiety over a period of time, and the effect of environmental changes on state anxiety. The implications of the study are that aptitude-treatment interactions using within-task measures may be more useful in designing adaptive instruction than the current notion of using pretask measures.
PRETASK VERSUS WITHIN-TASK
ANXIETY MEASURES IN PREDICTING PERFORMANCE
ON A CONCEPT ACQUISITION TASK

Robert D. Tennyson Richard C. Boutwell
Florida State University Bucknell University

Cronbach and Snow (1969) suggested that if development in a wide range of persons is to be facilitated, a wide range of environments suited to the optimal development of each individual must be offered. Thus, instructional units of the same basic subject matter representing various measured learner trait characteristics would be necessary. For example, an interaction of aptitude and instructional treatment would prescribe one type of sequence and media for a learner of certain characteristics, and another learner of differing trait characteristics would receive an entirely different mode of instruction. Cronbach and Snow (1969) advocated that an extensive research program be conducted to identify the characteristics which interact maximally with instructional treatments.

Research studies (Tallmadge, Schearer, & Greenberg, 1968; Dunham & Bunderson, 1969; Merrill, 1970) were undertaken to determine the specific environment assignments from pretask measures. The studies indicated that aptitude treatment interactions (ATI) have an elusive nature. Bunderson and Dunham's (1970) final report of a three-year research project on cognitive abilities and learning, challenged the ATI concept as a viable predictive procedure in "real world" instructional
contexts. Reasons for their skepticism can be summarized as: (a) useful disordinal interactions are rare; (b) disordinal interactions are not sufficiently robust under minor changes in the task or population; and (c) the benefits from disordinal interactions may be less than those attainable through revision of a single optimal treatment. They suggested that, rather than seeking disordinal interactions to assign individuals to different macro-treatments, ATI findings be used to develop the optimal instructional program to reduce the learning burden of low aptitude individuals. After the single best treatment has been maximized, micro-treatment variables can be applied adaptively within the instructional program.

The use of anxiety measures as a variable for adaptive assignment of macro-instructional strategies resulted from the generalization of the Taylor (1956) and Spence (1958) anxiety theory of competing responses. Low and high anxiety subjects respond differently on tasks of varying difficulty. Recent extensions of this relationship between anxiety and instruction (O'Neil, Spielberger, & Hansen, 1969; Tennyson & Woolley, 1971; Leherissey, O'Neil, & Hansen, 1971; Boutwell & Tennyson, 1972) indicated that trait or state anxiety measured prior to a learning task were not as effective in predicting student performance as state anxiety measured while learning the task. Cronbach (1967) assumed that obtaining periodic measures of the learner increased the precision of assigning optimal macro-treatments to individuals. In an experimental situation, it would be appropriate to desensitize the subjects to the experimental instrument(s) by using a time series design.
Hypothesis

The study proposes that instructional strategies can be adapted according to a learner's within-task state characteristics. In contrast to assigning individuals to different macro-treatments, which ignore the importance of fluid ability (Cronbach & Snow, 1969) and intratask experience, and to Crowder's (1959) intrinsic programming where only the last response made by the learner is considered, a within-task adaptive strategy would make instructional decisions based on a learner's behavior during a segment of the learning task. The reliability of a pattern of state responses, rather than a single response, should increase the validity of adaptive decisions. The within-task method of adaptation is such that remedial "hole patching" (Cronbach, 1967) is avoided.

The experiment investigated the hypothesis that within-task measures are significantly better predictors of learner performance than pretask measures. Two variables were selected to measure subject pretask characteristics: the Taylor Manifest Anxiety Scale and the Spielberger, Gorsuch, and Lushene (STAI) A-State Scale. To obtain a typical performance on the pretask measures, a time series design was used in which the two instruments were administered four times in five consecutive weeks. The STAI A-State Scale was selected as a within-task measure of the subjects' state characteristics.

Method

Learning Task

The behavioral objective of the learning task required subjects to read a definition of RX₂ crystals that included two prompted examples
and two prompted nonexamples. Then the subjects had to identify examples of \( \text{RX}_2 \) crystals from a list containing both examples and nonexamples. A prompted instance is when an attribute is individually separated, identified, and defined. The definition (critical attributes) of \( \text{RX}_2 \) crystals focused the subject's attention to the basic, repeating, two-to-one ratio in crystal structure of the \( \text{RX}_2 \) atoms. Each page consisted of a crystal picture taken from *Crystal Structure* by Wyckoff (1968). Reproductions of the crystals were made from photo plates which provided shaded pictures. Crystals were shaded so that depth perception would not confound identification.

The learning task consisted of 20 examples and 20 nonexamples which ranged widely in difficulty. In order to obtain empirical validation of the item difficulty level (competitive response strength of hard and easy items) Tennyson and Boutwell (197:) defined an instance probability analysis for rating and categorizing the items of a proposed task by the difficulty of the subjects to recognize the items. Each instance (positive and negative) was tested for competitive response strength by administration to 100 subjects from the same target population used in the experiment. High probability items (strong correct response) were those instances identified correctly by 70 percent or more of the subjects. Low probability items (strong incorrect response) were those instances classified correctly by less than 30 percent of the subjects.

**Anxiety measures.** The Spielberger, Gorsuch, and Lushene (1969) State-Trait Anxiety Inventory (STAI) A-State was used to measure the state anxiety. The A-State (Form X-1) scale required the subjects to indicate how they felt "at that moment." Taylor's Manifest Anxiety Scale
(1953) (MAS), was used to measure trait anxiety, which asked the subjects how they "generally feel."

Experimental design. The time series experimental design (Campbell & Stanley, 1963) established the stability of the subject's anxiety baseline. In the past, studies using anxiety as an independent variable have measured the subjects' anxiety prior to the treatment session. The Taylor MAS and the STAI A-State Scale were administered once a week for three weeks, constituting the first three sessions. After the fourth week, the Taylor MAS and STAI A-State Scale were administered immediately before and after the task. The scales were also given in a final session a week later.

Procedure. The 75 subjects were divided into three groups to facilitate task administration. Subjects were seated in a large room at alternating desks with a self-instructional booklet. Testing occurred the same time and day each week. The subjects' responses were marked on IBM answer sheets and checked by an experimenter when completed. The booklet explained that the subjects were participants in an experiment and asked them not to discuss the proceedings with anyone. The first three weekly sessions consisted of the administration of the Taylor MAS and the STAI A-State Scale and averaged 30 minutes.

After the third session, an interval of a week lapsed to minimize practice effects. The fourth session began with the Taylor MAS and STAI A-State Scale, and was followed by the task of crystal identification. The subjects were presented the illustrated definition showing two examples and two nonexamples. They were asked to identify examples of RX₂ crystals from a list of 40 instances. Upon completion of 25
instances, the subjects took the STAI A-State Scale. The directions asked subjects to respond according to how they "feel during the crystal identification."

The final period began with the subjects going to an adjacent room for a ten-minute rest period. At the end of the rest period, the subjects took the Taylor MAS and the STAI A-State Scale a fifth time. Subjects were asked to report how they felt at that moment. This entire session averaged 80 minutes. The experimenters checked the completed answer sheets to be certain that all questions had been answered.

One week after the administration of the experimental task the subjects again took the Taylor MAS and STAI A-State Scale. The directions and procedures were similar to the first three sessions.

Subjects. The subjects for the instance probability analysis were 105 students randomly chosen from undergraduate educational psychology classes at Florida State University. From this same population, but not the same students, 75 subjects were randomly selected to participate in the five-week study. Several subjects were dropped from the analysis because they failed to attend all four sessions. Two subjects who had unusually extreme variations on the Taylor MAS were also omitted from the analysis.

Results

Pretask versus within-task. The hypothesis of this experiment was that within-task measures of subject state variables are better predictors of subject performance than pretask measures, either
trait or state. Regression analysis was used to test this hypothesis. The correct score on the crystal identification task was the dependent variable. The four pretask Taylor MAS scores were averaged to obtain a baseline measure per subject. The same was done for the four pretask STAI A-State Scale scores. The full regression model had five variables as the predictors: (a) the average score of the pretask Taylor MAS measures; (b) the average score of the pretask STAI A-State Scale measures; (c) the STAI A-State Scale score reporting within-task anxiety; (d) the posttask Taylor MAS; and (e) the posttask STAI A-State Scale. These last two anxiety measures were taken during a rest period following the task. The final two anxiety measures, taken a week later, were not included in the regression analysis. Interactions between the variables were not interpretable, therefore, none were included in the model. The first model dropped the pretask Taylor MAS average score and resulted in a nonrejection of the null hypothesis ($F=3.09$, $df=2,70$, $p > .05$). In the second test, the pretask Taylor MAS and STAI A-State Scale average scores were dropped from the full model. This partial model again resulted in a nonsignificant $F$-test ($p > .05$). To test the effect of the within-task measures, the within-task STAI A-State Scale score was dropped. The test of this partial model resulted in a significant $F$-test ($F=9.49$, $df=2,70$, $p < .01$). A fourth test dropped the posttask STAI A-State Scale score resulting in a significant $F$-test ($F=3.98$, $df=2,70$, $p < .05$). An $R^2$ of .26 was obtained for the two pretask measures; for the within-task measures, the $R^2$ was .73.
Anxiety Measures

The two anxiety scales were individually analyzed to demonstrate how the within-task measures can provide a more accurate prediction of performance over the pretask measures.

**Taylor MAS.** The subjects were divided into four quartiles based upon their average score for the pretask Taylor MAS measures. The first quartile was the lowest trait anxiety and the fourth quartile the highest trait anxiety with mean scores for the three sessions of 6.6 and 23.8 (Table 1). After one week of nontesting (fourth week), the Taylor MAS was administered a fourth time prior to the task, a fifth time during the rest period, and a sixth time a week after the experimental treatment. A one-way analysis of variance of repeated measures of the low anxiety first quartile subjects' scores for the six periods resulted in a difference between sessions ($F=2.68$, $df=6,90$, $p<.05$). Duncan's New Multiple Range Test (NMRT) was used to analyze the six means. Sessions four, five, and six, which did not statistically differ ($p>.05$), did differ from session one ($p<.01$). However, the analysis of variance for the high anxiety subjects showed no difference between the six sessions ($F=1.06$, $df=5,68$, $p>.05$).

**STAI A-State Scale.** The STAI A-State Scale was administered for three consecutive weeks immediately following the Taylor MAS. After one week of nontesting, the STAI A-State Scale was given before the experimental task, again within the task, during the rest period, and, finally, one week later. Within-task (fifth session) STAI A-State Scale scores were divided into quartiles (Table 2). The first quartile mean score was 36.5 (low anxiety) and the fourth quartile mean score was
TABLE 1

First and Fourth Quartile Mean Scores on the Pretask Taylor Manifest Anxiety Scale

<table>
<thead>
<tr>
<th>Sessions</th>
<th>High Anxious</th>
<th>Low Anxious</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.6</td>
<td>8.6</td>
</tr>
<tr>
<td>2</td>
<td>23.3</td>
<td>5.9</td>
</tr>
<tr>
<td>3</td>
<td>23.9</td>
<td>5.4</td>
</tr>
<tr>
<td>4</td>
<td>22.3</td>
<td>4.7</td>
</tr>
<tr>
<td>5</td>
<td>22.5</td>
<td>4.8</td>
</tr>
<tr>
<td>6</td>
<td>22.6</td>
<td>5.6</td>
</tr>
<tr>
<td>Mean</td>
<td>23.0</td>
<td>5.6</td>
</tr>
</tbody>
</table>

*Note:*--High anxious were the fourth quartile subjects, while the low anxious were the first quartile subjects.
A one-way analysis of variance of repeated measures across the six periods resulted in a statistical difference between sessions for the first quartile low anxiety subjects ($F=2.54$, df=5,76, $p < .05$). The results of Duncan's NMRT showed that session five was significantly different from the other sessions ($p < .05$). There were no other differences ($p > .05$). An analysis of the fourth quartile high anxiety subjects across sessions resulted in a significant difference between means ($F=2.96$, df=5,72, $p < .05$). Duncan's NMRT showed a significant A-State increase on the within-task session (5) measure over the remaining sessions ($p < .05$). The other sessions did not differ on measured A-State ($p > .05$).

Discussion

The purpose of this study was to investigate an alternative approach to the ATI method of adapting instruction to individual differences. ATI methodology has focused on pretask measures as the variable for selecting appropriate individualized instructional strategies. However, portions of a learning program may be adjusted once a student starts performing a task. Results of the regression analyses do imply that Bunderson and Dunham's (1969) model of micro-treatments using adaptive approaches is a vital alternative to the macro method. Thus, Cronbach's (1967) thesis of differing environments according to individual characteristics is based on concurrent state measures.

Trait anxiety as a possible pretask measure for instructional purposes was shown to be confounded by a desensitizing effect of repeated exposures. Low anxiety subject scores suggest a desensitizing
effect associated with the administration of repeated measures of the same instrument week after week. Contrasted with the low anxiety subjects' decreasing anxiety level are the high anxiety subjects' which show a consistent anxiety level. The reliability of the Taylor MAS for the high anxiety subjects is consistent, but for the low anxiety subjects the significantly decreasing level of anxiety suggests other factors might limit the accuracy of pretask measures for selecting appropriate macro-instructional sequences.

The data analysis of the STAI A-State Scale scores show changes resulting from the manipulation of the environment by introducing an instructional situation not previously encountered. The A-State level of both high and low anxiety subjects increased during the task. The empirical data were significantly related to performance on the learning task, but the pretask state anxiety was not. This apparent change in A-State implies that state characteristics do fluctuate during a task. As the instructional environment changes, state variables would change accordingly, requiring adaptive models that update the learner's profile for adequate decision making on micro-treatments. Future investigations should help decide whether a learner maintains a trait characteristic within certain types of tasks or for all tasks.
REFERENCES

Boutwell, R. C., & Tennyson, R. D. Memory monitoring as a function of anxiety level and task difficulty in concept acquisition. AERJ, 1972, in press.


