The Productive Thinking Program (PTP), consisting of 16 programmed lessons designed to develop productive thinking skills, was used as the basis of a program conducted with 546 fifth graders in 21 classes in four upstate New York school districts. One hundred and twenty students from one of these districts were involved in a modified Solomon Four Groups Design. A number of PTP stories were presented to both control and experimental groups, but the experimental group alone was invited to participate with the story characters in their adventures. The Minnesota Tests of Creative Thinking were used to assess productive thinking, with the construct validity being of major concern. Problem solving was defined in the categories of fluency, flexibility, and appropriateness of response. In this study, three experimental groups and three control groups were set up to study the effectiveness of the PTP. There were two teacher involvement and two pre-test conditions for the experimental and control groups. Generally, the findings mildly support that the PTP was effective in bringing about differences in posttest scores on the variables studied. One explanation of these findings could be that this application of the PTP differed from that recommended by its authors. (Author/BJG)
The Effectiveness of the Productive Thinking Program

George Schuler
Ithaca College

The study to be described was part of a larger one that dealt with the interaction between school anxiety and problem solving. One of the hypotheses stated that if problem solving skills were to be increased, school anxiety would be reduced. To teach problem solving, Covington, Crutchfield, Olton and Davies' Productive Thinking Program (PTP) was chosen. The research edition of the PTP consisted of 16 programmed lessons designed to develop skills in productive thinking. The skills included planfulness in dealing with problems, searching for and generating many hypotheses, checking the hypotheses against the facts, and trying to think of new and different ideas. It was written for use with fourth- to sixth-graders. The Program is presented as a series of 15 adventure stories in a comic book format. The PTP involves active participation on the part of the learner in a sequence of steps leading to specifically stated objectives. It also presents a private learning situation. (It was felt that, if school anxiety were to be reduced, the private learning situation offered by an instructional program might be better for students whose anxiety in school settings was related to something like audience sensitivity.)

Subjects consisted of 545 fifth graders in 21 classes in four upstate New York school districts. There were 120 students from one of the districts who were involved in a modified Solomon Four Group design which had three experimental and three control groups.

The control groups read a set of adventure stories which were extracted from the Productive Thinking Program. They were given essentially the same experience as the experimental groups using the PTP, without the invitation to participate with Jim and Lila (the "heroes" of the PTP) as they go through the adventures solving the mysteries that are presented to them.

The definition of problem solving involved three categories: fluency, flexibility, and evaluation or appropriateness of response. Similar categories are used by Torrance (1966) and others in their tests of creativity. Fluency deals with the number of solutions generated; flexibility deals with the number of different categories of ideas, and the evaluation variable has to do with the appropriateness of the responses.

To assess productive thinking, a modified precursor of the Torrance Tests which had been devised by Yamamoto, (1965), The Minnesota Tests of Creative Thinking, Form KVI, were used. The test consists of five tasks. The first three tasks present a picture of the nursery rhyme character, Tom the Piper's son, stealing a pig and away he's running. In Task 1 the students are invited to ask questions that will help to understand what's going on in the picture.

They are instructed not to ask any questions that can be answered just by looking at the pictures. The second task is "Guess Causes". Students are asked to list their ideas about what led up to the situation pictured. The third is "Guess Consequences". Students are asked to list their ideas about what they think the outcome will be. Tasks 4 and 5 present the respondents with a picture of a toy dog and involve the familiar product improvement and unusual uses tasks. Students are asked to list all the ways they can think of to make the toy dog more fun to play with and what can be done with the dog even if it isn't changed.

Each task was scored separately for all three dimensions: fluency, flexibility, and appropriateness or relevance. The scoring was based on the pre-existing criteria of Yamamoto (1964), modified as we had experience scoring it with the population we were using. Scorer reliability was studied by doing three blind scorings of each task for 76 randomly selected booklets. The mean correlation for intrascorer reliability was .9.

Yamamoto (1965) found a test-retest reliability coefficient over 8 weeks of .79. Our stability coefficients were a little lower than his. Over a ten week period we obtained coefficients around .6.

The validity of the problem solving tests was viewed primarily in terms of construct validity. The tests were scored tightly within the confines of the definitions of problem solving. We found, as would be expected, that flexibility was partly dependent on fluency, those measures are not really independent. There was divergence for the evaluation score so that seemed independent of fluency.

Yamamoto's (1965) validity considerations involved two studies using external criteria. One was teacher rating. He had a population of 76 students, and found a correlation .47 between scores on KVB and teachers' ratings. He also did a teacher nomination study in which teachers reported kids as being high, medium, and low on problem solving and creativity variables and found, through analysis of variance, that the teacher nominations separated scorers (on the test).

In the present study, three experimental and three control groups were set up to study the effectiveness of the PTP. There were two teacher involvement and two pre-test conditions for the experimental and the control groups. In one situation, the teachers were merely distributing and collecting the booklets, treating the materials much as a workbook would be handled. In another situation, the teachers were apprised of the objectives of each lesson. They were told of the principles that were being attempted to be implemented and were encouraged to use these objectives and these principles during the rest of the classroom day. We weren't able to observe what the teachers did. Of course, it would have made the study much more useful if we had been able to do some of these observations.

There were also two pretest conditions, and here's where the Solomon Design comes in. One of the experimental and one of the control groups experienced no pretest. Of interest were the determination of whether there was a pretest effect on problem solving posttest scores, and whether there was a pretest x treatment interaction effect.
The students used the PT booklets at about the rate of two lessons a week. At about the same rate and time, the control groups were reading a set of adventure stories written to parallel the stories contained in the Program. In some instances, the control teachers were using the booklets as reading lessons because the children were very interested in the adventures of Jim and Lila.

Changes in performance from pre- to post-test problem solving scores (using notoriously unreliable gain scores) are shown in Table 1.

<table>
<thead>
<tr>
<th>FLUENCY</th>
<th>FLEXIBILITY</th>
<th>EVALUATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Gain</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>e-1</td>
<td>3.72</td>
<td>18.32</td>
</tr>
<tr>
<td>e-3</td>
<td>4.84</td>
<td>10.88</td>
</tr>
<tr>
<td>c-1</td>
<td>.70</td>
<td>12.81</td>
</tr>
<tr>
<td>c-3</td>
<td>-.03</td>
<td>11.36</td>
</tr>
</tbody>
</table>

Analyses of variance and subsequent multiple comparisons of the various treatment group combination using Sceffé's method indicated that scores increased significantly ($p < .01$) for evaluation, i.e., students' responses were more relevant to the tasks in the post-test situation, for the control groups but not for the experimental groups! No other significant gains were observed although generally, the experimental groups did score a little bit higher. The control groups did even better when the teachers didn't try to implement the objectives. The experimental groups did better when the teachers did try to implement the objectives.

There were 88 subjects for whom complete data was available participating in the Solomon Design, with 22 in each cell. Their posttest scores are presented in Table 2. There were no pretest x treatment interaction effects for any of the three scoring categories.

On the fluency variables, although posttest scores favored the group that used the Productive Thinking Program, the difference between control and experimental groups was not significant. There was a significant ($p < .01$) pretest effect. The students that took the pretest scored higher than the students who didn't take the pretest on the fluency posttest scores.
Table 2
Problem Solving Posttest Scores for Solomon Design Group

<table>
<thead>
<tr>
<th></th>
<th>Fluency</th>
<th></th>
<th>Flexibility</th>
<th></th>
<th>Evaluation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>58.55</td>
<td>14.30</td>
<td>37.14</td>
<td>6.92</td>
<td>70.16</td>
<td>13.12</td>
</tr>
<tr>
<td>No Pretest</td>
<td>43.55</td>
<td>14.76</td>
<td>33.82</td>
<td>10.11</td>
<td>79.46</td>
<td>11.36</td>
</tr>
<tr>
<td>Control Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>46.68</td>
<td>16.04</td>
<td>31.50</td>
<td>7.71</td>
<td>73.82</td>
<td>15.16</td>
</tr>
<tr>
<td>No Pretest</td>
<td>42.73</td>
<td>14.22</td>
<td>29.55</td>
<td>9.45</td>
<td>73.40</td>
<td>15.07</td>
</tr>
</tbody>
</table>

Flexibility posttest scores favored the group that used the Program at the .05 level. (Recall that fluency and flexibility scores were not wholly independent).

As far as the evaluation variable is concerned, there were no significant differences.

Generally, the findings mildly support the Productive Thinking Program as being effective in bringing about differences in posttest scores on the variables that were studied.

Included as part of the study was an attitude survey, the Children's Attitude Inventory Toward Problem Solving (CAPS), designed by Covington (undated). One scale of CAPS deals with the valuation of problem solving activities. It contains items like, "If you can't solve a problem at first, you should keep trying," and "It's important to find solutions to problems." The second scale of CAPS deals with one's feelings about his own competence in problem solving. It includes such items as, "Other kids in the fifth grade usually do better than me on solving problems like this one."

There was a tendency toward greater valuation of problem solving activities for those subjects who used the Program without the teacher involvement. They thought problem solving activities were more important than any of the other groups. There was a pretest x Productive Thinking Program interaction effect. Those students who had the pretest and used the Program scored higher on the valuation of problem solving activities than students who did not take the pretest.

As far as feelings of self-confidence and competence in problem solving are concerned they did not seem to be effected by the treatment that we delivered.
The attitude change is consistent with the findings of Crutchfield and Covington (1966), Hippie and Treffinger (1968), and Treffinger (1969). Even though the scores on problem solving tasks don't change, the attitude, the valuation of problem solving activities, does change. They found the same relationship as we did with feelings of confidence in one's own competence to solve problems. There was no change. Maybe the kids think that, "If they give it to you to do in school then it must be important," but as far as feelings of competence are concerned, perhaps the private learning situation involved in the instructional program does not permit the learner to get the kind of public feedback or approval that would provide information to him about the adequacy of his performance. Solving problems in the systematic way called for by the PiP presents an alien task for many children. They may need some kind of social or affective feedback - feedback from peers or supportive feedback from significant adults - in addition to the cognitive feedback provided by the Program.

Dacey (1967) explored the relationship between problem solving variables and the Productive Thinking Program. He used slightly different tests than we did with eighth graders - and the program is not designed for eighth graders - and found that the greatest effect on posttest scores were the creativity pretests he used. Hippie and Treffinger (1968) working with fourth- through seventh-graders, gave the program in 16 days rather than spreading it out over eight weeks as I did and as its authors suggested. They found that there were only three significant posttest differences out of 40 analyses and concluded that those were artifacts.

Generally, the Program did not work as well for me as it did for its authors, but its use was not wholly consistent with the procedures recommended by its authors. The dependent variables differed, the control experiences differed, and the format of the presentation differed.

The Program's authors have come up with another edition of the Program published by Charles Merrill. In the new edition, there are fifteen lessons designed to be used at the rate of one a week. There's an extensive Teacher's Guide and a book of problem solving sets. The authors suggest using four sessions each week: one devoted to work on the basic lesson, one for follow-up discussions, one for work on the appropriate problem solving set and one for follow-up discussions of the problem solving set. The problem solving set consists of a set of elaborate materials which present problems for kids to work on, which will, hopefully, encourage transfer. Some of the problems they've come up with are really interesting.

I know of no research evaluating the revised published edition. I do know the revised publish edition is fairly expensive. Extending the time over which the materials are used is acceptable for the new Program, but they don't recommend reducing the time over it's used. They report that significant gains in thinking ability will occur even when teacher involvement is minimal but that greater gains are found with active teacher participation.

I think that one of the important things going on here, and I did no systematic study of it, has to do with what the teachers were doing with this Program. One of the teachers in the study had been a research chemist and she herself was a very creative person. One of the creative teaching techniques she used had the students writing up insurance contracts. There was the Tom Sawyer Insurance Company. For ten cents a week, the initial fee, plus five cents a week thereafter, they would give you sickness insurance. There was a list of things that they would pay for: a skinned knuckle was penny, a skinned ankle was two cents and if you vomited, you got a nickel. Breaking your arm was really big. They also had life insurance. If you died they paid $5 and in parenthesis, it was
stated that the money would go to your family. One of the companies even would lend you the money (at a rate) to pay their premiums. I'm talking about this to illustrate that an unusually creative teacher could do a great deal with the Program. Her class got an awful lot out of the Program and one of my feelings is that the Program can be infinitely more effective if we involve the teachers more actively in its application.
REFERENCES

Covington, M.V. *A childhood attitude inventory for problem solving.* Mimeographed, undated.


Yamamoto, K. *Experimental scoring manuals for Minnesota tests of creative thinking and writing.* Kent, Ohio: Bureau of Educational Research, Kent State University, 1964.