It is assumed that inference is at the heart of thinking and little is known about possible relationships between learners' decision-making styles and teaching strategies. Therefore, it was hypothesized that instructed children score higher on an inference test than those not instructed, and interactions occur between different decision-making styles and the reflective teaching strategy used in the instruction. Subjects were 42 white middle class fifth graders in two classes and two schools. A pre-, posttest 2 x 3 factorial design was used: experimental and control groups were divided into three categories of decision-making style -- "overgeneralize, inference, and cautious" as determined by Hilda Taba's Social Studies Inference Test. The "reflective teaching strategy" consisted of a four week self-instructional programed package designed to develop reflective thinking. Analysis of test results revealed no significant differences between groups on scores associated with inference. Although an interaction pattern did emerge, it was not statistically significant. The interaction pattern suggests that children who do not have an inference decision-making style will need non-reflective teaching strategies to develop their inference abilities. (DJP)
THE INTERACTION OF DECISION-MAKING STYLE, TEACHING STRATEGY, AND DECISION-MAKING CONTENT MATERIAL IN SOCIAL STUDIES

Lyman R. School
University of Washington

Ambrose A. Clegg, Jr.
University of Washington

This paper was presented as part of a symposium, Research in Social Studies Education—II, at the annual convention of the American Educational Research Association, Minneapolis, Minnesota March 5, 1970
Although many studies have been conducted about the development of productive thinking (including decision making), few have considered such individual differences among learners.

Some research has considered individual differences by investigating effects of socio-economic level, I.Q., and sex, such as Olton and others (1967), and Klausmeier (1964).

Kagan's impulsive-reflective dimension, or non analytical-analytical (1963), appears to have relevance for individualized instruction, but more research is necessary to explore teaching strategies which are in fact learning strategies for children who have different styles of conceptualization.

One psycholinguistic study did focus on the reflective-impulsive dimension and was conducted with eight and nine year old boys. Impulsive boys did not perform as well as reflective boys on tasks where the required response had to be generated from past learnings when the channel of communication was an auditory-vocal one. An implication here is that other channels should be explored and enhanced. (Gentile, 1968)

In addition to investigating individual differences, more research is necessary related to inference and to developing this ability in children. Consensus has been reached that inference is "the heart of thinking," (Burton, Kimball, and Wing, 1960), and the basis of reflective activity (Dewey, 1933):

In every case of reflective activity, a person finds himself confronted with a given present situation from which he has to arrive at, or conclude to, something else that is not present. This process of arriving at an idea of what is absent on the basis of what is at hand is inference.
Stitt (1968) and Kagan (1966) have shown that training increased inference ability, and reflective behavior respectively with children. But most research concerning inference and other thinking processes has been with older subjects—primarily secondary schools, since people seem to feel that is the age when more sophisticated thought processes begin to develop. Piaget's work tends to support this contention.

Since inference is perceived as the vital component of reflective thought, and since little is known about possible relationships between learner's decision making styles and teaching strategies, particularly with elementary school children, the present study was conducted to investigate these factors and should be considered exploratory. Certainly more investigations in this area are necessary.

In this study a model was tested dealing with the interaction of the learner's decision-making style, the teaching strategy and the decision making content. The degree of match between the learner's decision-making style and the teaching strategy was hypothesized as an independent variable in the training of elementary school children in decision-making.

The following model illustrates the hypothesis that (1) the more similar the learner's cognitive style is to the teaching strategy, the less the amount of change in the learner's style. Conversely, the more the learner's decision-making style differs from the style used in the teaching strategy, the more the amount of change in the learner's style. In addition this change is hypothesized to be in the direction of the style used in the teaching strategy.
The learner's decision-making style was defined in relation to one specific dimension of productive thinking—inference. To determine which decision-making style a subject tended to use, Taba's Social Studies Inference Test was administered. Children's decision-making styles were categorized as overgeneralizing, reflective or overcautious on the basis of subscores on this test. Combining these categories with Kagan's impulsive-reflective dimension resulted in the following theoretical composite: impulsive-reflective-overcautious. Children were assigned to three treatment groups on the basis of their tendency to use an overgeneralizing, an inferring, or an overcautious decision-making style in making inferences. The three categories functioned as independent organismic variables.

The teaching strategy utilized could be considered reflective. Programmed material developed by Covington, Crutchfield and Davies was used. The material consisted of 16 self-instructional programmed lessons, one of the stated goals of which was to develop reflective thinking. Repeated practice with various "thinking guides" (which might be considered as reflective thinking skills) was provided in the material.

The content of the material consisted of a series of mysterious situations. Decisions were made during the analysis of each mystery. These decisions dealt with problems such as: what plan will we use, what evidence is relevant, what are the possible solutions, and which of the possible solutions check with the facts.

The present investigation, then, focused on one specific aspect of each factor in the three-way interaction. For learner's cognitive style the factor studied was the tendency to use an impulsive, reflective or overcautious decision-making style. The teaching strategy studied was reflective,
and finally the decision-making content consisted of mysterious situations for which decisions were made to resolve discrepancies. This was an exploratory study. The primary purpose of the study was to investigate:

1) if instructed children improve scores on an inference test significantly more than those not instructed,

2) if any interactions occur between different decision-making styles and the reflective teaching strategy used in the instruction, and

3) if any occur, which factors interact and in which ways.

EXPERIMENTAL DESIGN AND TREATMENTS

In a pre-, posttest design, a 2x3 factorial analysis was made with two treatments (experimental and control) and three categories of decision-making styles (impulsive, reflective, and overcautious). (See Figure 1)

The pre- and post-test design was based on two instruments which were both used in the pre-test and the post-test which were individually administered, and a four week instructional program which was administered to each class as a whole. All children were instructed. Three treatment groups were labeled "Impulsive," "Reflective," and "Overcautious" according to the tendency to utilize a particular style in making inferential decisions. Children were assigned to groups on the basis of their scores on the pre-test. Scores for each style were ranked, and students who scored in the top third were assigned to treatment groups for that style in which they ranked high. In cases where students ranked in the top third in more than one category, the students were assigned to that category in which their rank was highest.

During the instructional period, one fifteen-minute lesson was presented each day for three weeks. Makeups were given at another time in the day. After sixteen lessons, the post-test was given.
FIGURE 1

Factorial design with two treatments and three categories of decision-making styles*

<table>
<thead>
<tr>
<th>OVER-GENERALIZE</th>
<th>EXPERIMENTAL (Instruction)</th>
<th>CONTROL (No Instruction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INFRINGEMENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAUTIOUS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 42

*Repeated analyses of variance for this factorial design were made on each of the six sub-scores of the social studies inference test. The following are the six sub-scores which were used individually as the dependent variable in each analysis:

C = Overcaution  D = Discrimination  I = Inference
OG = Overgeneralization  A = Accurate  E = Error
HYPOTHESES

1) Children who are exposed to instruction in reflective thinking will score higher on an inference test than those who have not had such instruction.

2) Children who initially overgeneralize and also those who initially are overcautious will increase their inference score more than those instructed children who initially tend to make correct inferences.

3) Children who overgeneralize will score higher than those instructed children who are overcautious.

METHOD

Subjects: Subjects for this study included two fifth grade classes from two public schools located one mile apart in a Caucasian middle class socio-economic suburban area in Seattle, Washington.

Program Description: Instructional materials were The Productive Thinking Program, Series One: General Problem Solving. This program has as one of its stated goals, the development of reflective thinking (p. 18, Teachers' Guide). It is a sixteen lesson, self instructional programmed package which was developed by Covington, Crutchfield and Davies at the University of California at Berkeley, 1966. Klausmeier found the program significantly increased fifth graders' performance on classic problem solving tasks compared to control subjects who did not have the instruction.

Criterion Measure: Testing materials were Taba's Social Studies Inference Test. This test contains three or four sentence descriptions of various situations and three statements about each situation. The student is required to indicate for each statement whether it is probably true,
probably false, or if he can't tell. The scoring procedure indicates whether the student generally was overgeneralizing, making inferences, or was overcautious. This test (time: approximately 25 minutes) was administered individually so response latency could also be recorded.

RESULTS

The mean squares and F ratios on all subscores including response latency are presented in Table 1.* These analyses of variance show statistical significance (p < .05) on the subscores for discrimination and for error. Two significant differences are for treatment main effects and one is for category main effects. In addition, patterns emerged on the subscores for overcautious, overgeneralize, and inference subscores. (See Table 2.) These scores (Oc, Og, and I) dealt more with actual inference than the subscores above (i.e., discrimination and error).

The significant treatment differences were the following: control groups scored higher on the discrimination subscore, which indicated correct responses; and the control groups also scored lower on the error subscore. (See Table 3.)

The significant category difference was accounted for by the inference groups which scored higher on the discrimination subscore than the overgeneralizing and overcautious groups in that order.

These three differences on the post-test favored the control groups rather than the experimental groups, and therefore were in the opposite direction from the three hypotheses. However, these differences were on subscores related in factual problems and did not depend on direct inferences.
### Table 1

Analysis of Variance for Factorial Design for Two Treatments (Experimental Versus Control) by Three Categories (Overgeneralize, Inference, Cautious)

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (T)</td>
<td>1</td>
<td>.60</td>
<td>.09</td>
<td>3.43</td>
<td>.17</td>
<td>6.88</td>
<td>.39</td>
<td>24.30</td>
<td>.72*</td>
<td>.06</td>
<td>.67</td>
<td>97.52</td>
<td>.49*</td>
<td>3517.88</td>
<td>.66</td>
</tr>
<tr>
<td>Category (C)</td>
<td>2</td>
<td>1.88</td>
<td>.24</td>
<td>14.00</td>
<td>.69</td>
<td>20.85</td>
<td>1.20</td>
<td>15.50</td>
<td>3.02*</td>
<td>5.88</td>
<td>.58</td>
<td>54.45</td>
<td>2.08</td>
<td>1974.00</td>
<td>.36</td>
</tr>
<tr>
<td>T X C</td>
<td>2</td>
<td>4.31</td>
<td>.62</td>
<td>32.00</td>
<td>1.58</td>
<td>12.67</td>
<td>.73</td>
<td>6.50</td>
<td>1.41</td>
<td>2.79</td>
<td>.24</td>
<td>8.31</td>
<td>.46</td>
<td>5695.14</td>
<td>1.03</td>
</tr>
<tr>
<td>Within Replicates</td>
<td>36</td>
<td>.58</td>
<td></td>
<td>20.30</td>
<td></td>
<td>17.42</td>
<td></td>
<td>4.67</td>
<td></td>
<td>11.81</td>
<td></td>
<td>17.80</td>
<td></td>
<td>5294.82</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P<.05

**Key:**
- **C** = Overcaution
- **D** = Discrimination
- **I** = Inference
- **OG** = Overgeneralization
- **A** = Accurate
- **E** = Error
TABLE 2

Means for the following subscores: OG, OC, I.
Over-all change estimate.

<table>
<thead>
<tr>
<th>ERROR SCORES</th>
<th>CORRECT SCORE</th>
<th>OVER-ALL CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overgeneralize</td>
<td>Overcautious</td>
<td>Inference</td>
</tr>
<tr>
<td>Exper.</td>
<td>Control</td>
<td>Exper.</td>
</tr>
<tr>
<td>OG</td>
<td>6.9</td>
<td>6.3</td>
</tr>
<tr>
<td>I</td>
<td>6.4</td>
<td>5.3</td>
</tr>
<tr>
<td>OC</td>
<td>5.9</td>
<td>6.9</td>
</tr>
</tbody>
</table>

TABLE 3

Means for the following subscores: Distinctions, Errors.

<table>
<thead>
<tr>
<th>DISTINCTIONS</th>
<th>ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exper.</td>
<td>Control</td>
</tr>
<tr>
<td>OG</td>
<td>6.43</td>
</tr>
<tr>
<td>I</td>
<td>8.00</td>
</tr>
<tr>
<td>OC</td>
<td>6.43</td>
</tr>
<tr>
<td>7.00</td>
<td>8.5</td>
</tr>
</tbody>
</table>
The patterns observed in Table 2 suggest an interaction between categories and treatments on the totals of the 3 subscores, OG, I and OC, which all relate to inference. By summing the positive and negative improvement scores (based upon a comparison between the control and the experimental group means), a total improvement score was estimated. The result was that the inference group increased plus one, while the overcautious group decreased minus one, and the overgeneralizing group decreased a minus three. Although this is a gross estimate, it indicates that grouping scores may be valuable and it suggests a logical rank order of categories according to how much they might benefit from the instruction on this measure.

Conclusions
1) That there were no significant differences between groups on the three subscores associated with inference,
2) That the experimental groups actually scored significantly lower than the control groups on two of the three subscores associated with factual statements,
3) That the three categories scored differently on the discrimination subscore, and
4) That patterns emerged which suggest interactions between categories and treatments.

Four explanations are related to the first aspect. First, although inference is a crucial aspect of productive thinking, it is possible that the criterion measure used was just not able to pick up the effects. Second, no total score was given for overall performance. Therefore, it might be that the overall performance was partitioned into so many subscores that it became extremely difficult for any differences to be significant by themselves.
This suggests that another possible analysis might still yield results consistent with the hypotheses. Third, there might be a latent difference, which suggests a second post-test be carried out. Fourth, inference behavior is such a basic type that much more instruction might be required to effect a change.

The second aspect, that of control's higher scores on two factual subscores, might be explained as follows. First, the control groups had been exposed to a factual mode of social studies education during the time the experiment group was exposed to a productive thinking mode; so it might be understandable that the control group would score higher on the factual subscores.

Second, the control group may have had a better memory of the stories in the inference test. A huge amount of time was required for one experimenter to pre- and post-test individual subjects. So, considering the available time to conduct the study, the control group had to be post-tested four weeks after the pre-test, while the experimental group was post-tested five weeks after the pre-test. This might indicate that the control group had a better set to listen to factual detail, since the experimental group's instruction might have interfered with their memory of the pre-test stories.

The third aspect was students in the three categories scored differently on the discrimination subscore. The six groups were not initially randomized on the basis of this score, so wide variability existed, but balanced itself out and was not strong enough to be significant. Separate analyses were computed for the pre-test which when compared to the post-test analyses indicate the experimental groups remained about the same while the control groups increased their scores on this subscore. Total increase was plus 4.2 for the
control group compared to a minus .9 for the experimental group. Not only
did the control groups increase their scores, but also the pattern in which
the increase occurred seemed logical. The inference group which was on the
right track improved most, the overgeneralizers next, and the overcautious
improved least. When this effect was added to the large variation that
existed between the control and the experimental groups on this subscore,
the category difference became significantly large.

The results shown in Tables 2 and 3 were interesting. These show a
pattern which indicates the inference group improved its scores after
instruction and was the only category to do this on an overall grouping of
the subscores which involved inference. Both the overgeneralizers and the
overcautious groups performed worse if they had instruction than they did
without instruction on these subscores. This, while not significant by the
method utilized in this first analysis (Table 1.), is precisely the kind of
difference this study was investigating: Is there an interaction between
types of decision making styles of the learner and that of the instruction,
with decision making content? If it turns out that only those children who
tend to make correct inferences can benefit from this program, then it is
necessary to have other teaching strategies for those children who are over-
generalizers and overcautious decision makers. The next analysis planned for
this data will involve calculating difference scores in terms of z scores
and statistically grouping those that are related to inference questions, to
see what the overall effect might be. The analyses completed thus far were
computed to investigate whether children actually might significantly change
that one subscore that indicated their major tendency, rather than to see if
there was an overall effect.
If an interaction appears, more testing of the children might be planned. Previous studies by the Berkeley developers of the instructional materials and by Klausemaier have indicated that children who have had this instruction perform significantly better on classic problem solving tasks. These studies, however, did not break down their data on the dimension of learner's decision-making style. The data compiled thus far can be utilized to investigate if there are interactions between learner's decision-making style and that of the instruction when the tasks are classic problem solving tasks, instead of the stories and inference statements which were the criterion measures in this study.


