This study sought to determine whether student teachers' questioning strategy can be modified to increase their use of higher level cognitive questions and also whether student achievement was higher in classes where the teacher asked more high level questions. Subjects of the study were 20 student teachers assigned to fifth-grade classes in public schools, who were pretested for similar questioning behavior. The 10 students in the experimental group participated in seminars on purposes and use of different cognitive levels of questions. All student teachers taught a 4-day unit on the same substantive material. Each day, their questions were recorded by observers using a specially developed observation schedule. Following the unit, pupil achievement was assessed by a specially prepared test. Analysis of data revealed that the trained student teachers asked a significantly greater number of higher level cognitive questions than those not specially trained. Pupil achievement in the two groups was not significantly different. This could be attributed to the limited time that students were exposed to higher level questions and to the probability that student teachers have little "power" to influence pupils in their responses. Further research is recommended on the relationship between teacher behavior and student achievement and on the place of questioning strategy training in teacher education. (Author/FT)
Questions are recognized as being important in fostering pupils' intellectual development. Nevertheless, analyses of classroom questions utilized by teachers indicate that memory and comprehension, two low-level mental operations, most commonly are emphasized in classroom discourse, often to the exclusion of higher order operations (e.g., Adams, 1964; Floyd, 1960; Stevens, 1912). Most questions in textbooks and other instructional materials are at the lower cognitive levels (e.g., Chew, 1966; Davis and Hunkins, 1966; Windley, 1966).

In view of this discrepancy between the possible usefulness of questions at higher cognitive levels and observed instructional practices, efforts have been undertaken to demonstrate the usefulness of higher order

---

1This paper was delivered at the convention of the American Educational Research Association, Minnesota, March 1970.
questions and to change educational practices. Hunkins (1966) reported that sixth grade pupils' social studies achievement was increased by using text-type materials which incorporated questions emphasizing higher-level cognitive operations. Clegg and others (1967) reported that student teachers taught to classify questions according to their cognitive level themselves asked higher level questions in student teaching. Taba (1966) and Parsons and Shaftel (1967) found that experienced teachers changed their classroom questioning behaviors following special intervention programs. Microteaching procedures, involving several dimensions of modeling, feedback, and practice, have been found useful in stimulating the acquisition of questioning behaviors by secondary teacher candidates (Berliner, 1969; Claus, 1969; Morse and Davis, 1970).

Research efforts to date, however, have not related differential teacher use of higher order questions to pupil learning. Consequently, this study was designed as a first investigation into an area largely unexplored.

Procedure

Subjects

Twenty senior university student teachers, assigned to fifth-grade classrooms in the public schools, and the pupils enrolled in their classes served as subjects for this study. Ten student teachers were assigned randomly to the experimental group; the remaining ten student teachers constituted the contrast group. The economic levels of the families of the children were classified as average to low. Approximately one-third of the children were Mexican-American children; the remainder were Anglo-Americans.
The mean IQ score of the children in the experimental classes was 104.10 and in the contrast classes 103.56.

**Experimental Preparation Program**

Student teachers in the experimental group \((N = 10)\) met with the principal investigator for five two-hour seminars over a period of four weeks at the beginning of the Spring Semester, 1968. These seminars focused on the purposes and use of varying cognitive levels of classroom questions before the student teachers taught an instructional unit. **Classroom Questions: What Kinds?** (Sanders, 1966) served as a basic source of study in these seminars. Student teachers learned to classify and compose questions in each taxonomic category for several teaching fields. Emphasis, nevertheless, was placed on questions in teaching the social studies. Audio and video tapes, microteaching, and role playing of alternative questioning behaviors were utilized. Some tasks were oriented to the system adopted textbooks; other tasks were not dependent on a textbook. The appropriateness of the question to the objectives for the lesson was stressed throughout the seminars. Subjects taped several of their own classroom lessons and these records were analyzed by the group and/or individually by the student teachers.

No special program was provided student teachers in the contrast group. They did meet as regularly with their university supervisors and dealt with common problems of student teaching in those seminars.

**Instructional Unit**

All student teachers in the study were directed to plan and teach a four-day unit with a culminating test, on the West Indies. This unit was
one in *Journey Through the Americas* (Drummond, 1962), the social studies textbook adopted for fifth grade in the cooperating school district. The daily lessons were to be 30 to 45 minutes in length, the usual time allotment for social studies in the system.

**Instruments and Data Collection**

**Teacher Oral Question Observation Schedule (TOQOS).** This criterion measure was a modification of one developed by Davis and Tinsley (1967). The format of the TOQOS made possible a sequential record of the oral questions posed by the student teachers. The seven categories from Sanders' (1966) hierarchy of questions constituted the major dimension of the TOQOS. Four other categories, not cognitive in nature, were included: affective, procedural, textbook and pupil-initiated. Textbook questions refered to those questions read by the teacher from the textbook or other printed materials. Pupil-initiated questions were those asked by a child and repeated by the teacher for student response.

Observers were trained in the use of the TOQOS for approximately twelve hours. During this period, they learned the category system and practiced coding audio and video tapes of classroom interaction. An interrater agreement of .86 over all categories was reached by the end of the training sessions, with reliability by categories ranging from .60 to 1.00.

All student teachers were observed using the TOQOS once prior to and each of the four days during the experimental period. They were observed for 30 minutes of the social studies lesson. Observers rotated between student teachers in order that a different observer was present in each classroom each day.
Student-Teacher Constructed Unit Examinations. Student teachers constructed an end-of-unit examination but did not administer it to pupils. Test items were categorized into one of the seven TOQOS cognitive categories by two raters. Interrater reliability was 1.00.

Pupil Achievement Criterion Test. A criterion measure of pupils' social studies achievement was constructed by the principal investigator. This 35 item, multiple choice test was designed to yield subscores corresponding to the seven TOQOS cognitive categories as well as a total achievement score. Reliability of the total test was estimated as .75. The criterion test was validated for the correctness of the question classification level and the appropriateness of the test for fifth graders by three expert judges. Only questions with one-hundred per cent agreement were used. This test was administered to all pupils on the day following the completion of the instructional unit.

Data Analysis

Data were subjected to analysis of variance procedures. Computer programs employed (Jennings, 1967) were based on multiple linear regression models.

Results

Student Teachers' Oral Questions

Prior to the preparation in questioning of student teachers in the experimental group, all student teachers planned and taught a 30-45 minute lesson on the income tax to their pupils. Their questioning behaviors were observed using the TOQOS. Obtained data were subjected to ANOVA procedures.
and results revealed no statistically significant differences between the groups in the percentages of questions asked in each TOQOS category. Consequently, at the beginning of the experiment, the two groups of student teachers may have been assumed similar (not significantly different) in their oral questioning behaviors.

Results of the analysis of variance of the oral questions asked by student teachers during the four-day unit are presented in Table 1. Five of the seven TOQOS cognitive category variables were significantly different beyond the .05 level of significance. This result is itself statistically significant (Sakoda, Cohen, and Beall, 1954). Group C student teachers asked a higher percentage of application questions ($p < .11$). No statistical difference was noted between the groups' uses of memory questions. All subjects in both groups asked questions in the three lower categories of the taxonomy. While no Group C student teachers asked analysis, synthesis, or evaluative questions, and the differences between groups for these categories were significant, that ten Group E student teachers in four class periods of instruction asked so few high level questions seems important to note. This analyses, does reveal that there was a significant difference in the cognitive emphasis of questions asked by student teachers who have special preparation in the purposes and use of varying cognitive levels of questions and those who do not have this preparation.

The mean percentage of affective and procedural questions asked orally by the student teachers during the unit are presented in Table 2. Group E student teachers asked a significantly lower percentage of procedural questions than did student teachers in Group C. Student teachers in the two
TABLE 1
Comparison of Cognitive Questions Asked Orally by Student Teachers During Experimental Period

<table>
<thead>
<tr>
<th>Question Category</th>
<th>N&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Group E (N=10) Mean (%)</th>
<th>N&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Group C (N=10) Mean (%)</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>10</td>
<td>50.7700</td>
<td>10</td>
<td>56.7610</td>
<td>0.879</td>
<td>0.6365</td>
</tr>
<tr>
<td>Translation</td>
<td>10</td>
<td>10.0500</td>
<td>10</td>
<td>5.9310</td>
<td>4.463</td>
<td>0.0465</td>
</tr>
<tr>
<td>Interpretation</td>
<td>10</td>
<td>17.5060</td>
<td>10</td>
<td>5.6810</td>
<td>25.337</td>
<td>0.0002</td>
</tr>
<tr>
<td>Application</td>
<td>6</td>
<td>2.0150</td>
<td>3</td>
<td>4.6767</td>
<td>3.397</td>
<td>0.1058</td>
</tr>
<tr>
<td>Analysis</td>
<td>8</td>
<td>1.9025</td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Synthesis</td>
<td>8</td>
<td>1.3175</td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Evaluation</td>
<td>4</td>
<td>1.5950</td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

N<sup>1</sup>: Number of subjects asking questions in that category

* P < .0001
### TABLE 2

Comparison of Affective and Procedural Questions Asked Orally by Student Teachers During Experimental Period

<table>
<thead>
<tr>
<th>Question Category</th>
<th>N(^1)</th>
<th>Group E (N=10) Mean (%)</th>
<th>N(^1)</th>
<th>Group C (N=10) Mean (%)</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affective</td>
<td>10</td>
<td>5.0050</td>
<td>9</td>
<td>4.3333</td>
<td>.103</td>
<td>.7500</td>
</tr>
<tr>
<td>Procedural</td>
<td>10</td>
<td>10.8320</td>
<td>10</td>
<td>23.6210</td>
<td>5.059</td>
<td>.0353</td>
</tr>
</tbody>
</table>

\(N^1\): Number of subjects asking questions in that category.
groups did not differ significantly in the percentage of affective questions asked. Group E individual student teachers' mean percentages for procedural questions ranged from 6.69% to 17.17% and Group C individual student teachers' mean percentages ranged from 7.27% to 63.8%.

**Student Teachers' Test Questions**

Table 3 presents the results of the analysis of variance of the percentages of questions at various cognitive levels composed for the unit examination constructed by the student teachers. Group E student teachers formulated a significantly higher percentage of questions at six of the seven cognitive levels than did the Group C student teachers. This number of significant differences is itself significant (Sakoda, Cohen, and Beall, 1956). Group C student teachers asked a significantly higher percentage of memory questions than did student teachers in Group E. The contrast group composed no questions from categories other than memory and interpretation. Thus, a difference did exist in the cognitive level of test questions developed by student teachers with preparation in questioning and those who have not had such special preparation.

**Pupil Achievement**

Pupils' mean achievement on the achievement test (total score and subscores) is displayed in Table 4. Only one statistically significant difference was observed. Pupils taught by Group C student teachers achieved higher scores on the analysis subscore than did pupils taught by Group E student teachers. This result (one significant of eight contrasts), however, is itself not significant (Sakoda, Cohen, and Beall, 1954). This result
### TABLE 3

Comparison of Questions Composed by Student Teachers for a Unit Examination

<table>
<thead>
<tr>
<th>Question Category</th>
<th>( N^1 ) Group E (N=10) ( \bar{X} ) Mean (%)</th>
<th>( N^1 ) Group C (N=10) ( \bar{X} ) Mean (%)</th>
<th>( F )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>6</td>
<td>9</td>
<td>42.707</td>
<td>.0001</td>
</tr>
<tr>
<td>Translation</td>
<td>1</td>
<td>0</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>5</td>
<td>2</td>
<td>.354</td>
<td>.5815</td>
</tr>
<tr>
<td>Application</td>
<td>2</td>
<td>0</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>1</td>
<td>0</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Synthesis</td>
<td>4</td>
<td>0</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td>4</td>
<td>0</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

* \( P < .001 \)
TABLE 4

Summary of Pupil Achievement on the Criterion Test

<table>
<thead>
<tr>
<th>Test Scores</th>
<th>Group E (N=265)</th>
<th>Group C (N=268)</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Score</td>
<td>17.2528</td>
<td>17.9701</td>
<td>2.7084</td>
<td>.1002</td>
</tr>
<tr>
<td>Memory</td>
<td>3.8030</td>
<td>3.9403</td>
<td>2.0073</td>
<td>.1499</td>
</tr>
<tr>
<td>Translation</td>
<td>3.3208</td>
<td>3.3246</td>
<td>.0014</td>
<td>.9704</td>
</tr>
<tr>
<td>Interpretation</td>
<td>2.6415</td>
<td>2.6754</td>
<td>.0964</td>
<td>.7570</td>
</tr>
<tr>
<td>Application</td>
<td>1.8189</td>
<td>1.8769</td>
<td>.3198</td>
<td>.5725</td>
</tr>
<tr>
<td>Analysis</td>
<td>1.5887</td>
<td>1.9515</td>
<td>14.7708</td>
<td>.0001</td>
</tr>
<tr>
<td>Synthesis</td>
<td>1.9925</td>
<td>2.1157</td>
<td>1.3854</td>
<td>.2396</td>
</tr>
<tr>
<td>Evaluation</td>
<td>2.0868</td>
<td>2.0858</td>
<td>.0001</td>
<td>.9933</td>
</tr>
</tbody>
</table>
indicates that there was no significant difference in pupils' social studies achievement between classes taught by student teachers with preparation in the purposes and use of varying cognitive levels of questions and those who had no special training in questioning.

Discussion

Findings of this study are consistent with other research on the impact of special preparation of student teachers about classroom questioning. Those candidates receiving this preparation asked a significantly higher percentage of questions at the higher cognitive levels in both classroom discourse and on an end-of-unit examination. The preparation seminars focused on the reality of classroom teaching and the discrepancy between desired and real opportunities (demands) for pupils to think at higher cognitive levels. Also, the seminars afforded candidates time and format for practice in developing questioning behaviors and, apparently, confidence in asking questions to which one correct response was not expected. During the instructional unit, observers noted, further, that the Group E student teachers seemed to probe and to refuse to accept childrens' responses at lower levels than intended by the questions asked.

These results should not be interpreted to suggest that the student teaching semester is the appropriate time for attention to classroom questioning by teacher candidates. Neither should it be used to restrict modes of instruction to seminars. Laboratory teaching as an initial component in the preparation program has been found useful in altering candidates' questioning behaviors (e.g., Davis and Smoot, 1969; Morse and Davis, 1970). Success has also been
noted using a game format with undergraduates prior to student teaching (Rogers, 1969). Investigation of a variety of instructional modes employed at different times during the teacher education program should be encouraged. Noted during the seminars was that these student teachers had not had previous direct instruction in the importance and use of classroom questions although they did report general attention having been given to development of pupils' intellectual processes. For candidates to learn to behave differently in their classroom performances, major components of their teacher education program likely will need to be rooted in realism, specificity, and activity (Davis, in press).

The effect of special preparation in questioning on test questions merits special comment. Previous research on planning, including planning for tests, has revealed that secondary student teachers formulated questions mainly at the lowest cognitive levels (Tinsley and Davis, in press). Indeed, the results of the Group C student teachers in the present study were quite similar to those revealed earlier. The present findings, consequently, indicate that the generally low cognitive level of teachers' test questions can be changed.

The relationship of teachers' behaviors to pupil outcomes is not at all clear (Rosenshine, 1969). Especially is this evident with regard to classroom questioning practices. Yet, a paucity of studies report attempts to investigate this relationship. The present study, encompassing only four days of instruction, hardly could be expected to change pupils' psychological set for lower-level cognitive expectations, in classroom discourse, instructional materials, and tests. Probably, pupils had
experienced mainly low-level, mostly memory and interpretation, questions
to the extent that the mass of research on questioning is appropriately relevant.
An extended period of time and a variety of learning opportunities seem necessary
to change pupils' responses. A period of eight days during which daily tests
composed of higher order questions has been shown to be insufficient to change
secondary pupils' set. On the other hand, Hunkins (1966) has reported that an
emphasis on higher order questions in text-type materials over a period of three
weeks can affect significantly pupils' achievement.

Other plausible explanations may be advanced. Student teachers
probably had little "power" to influence pupils in their responses to question-
ing. Had pupils' regular teachers modified their questioning practices, at least
over time, achievement might have been affected. Further, the general "cognitive
climate" of the classroom -- not just in social studies, but all day, every
day -- surely influences pupils' perceptions and behaviors.

Additional research on teacher candidates' use of questioning should
be directed at several vexing problems. Development and test of specific
laboratory and clinical elements in teacher education programs should be
matters of priority. Also, sequencing of program components and their possible
cumulative effects merits early attention. Efforts to understand effects of
candidates' teaching activities and behaviors on pupil learning, while
important, probably should be conducted along with, rather than precede,
special attention to these relationships between experienced teachers and
their regular pupils.
References


Cooper, James M.  Two Types of Social Studies Examinations and Their Effect on Student Learning. Doctor’s thesis. Stanford, California; Stanford University, 1968.


