The study deals with five research questions: (1) is the Taxonomy of Education Objectives: Affective and Cognitive Domains a reasonable tool for use in stating and measuring attainment of objectives? (2) is the cognitive domain cumulative in social science subject matter? (3) does ability to operate at the various cognitive levels differ with intelligence? (4) is degree of affect cumulative? and (5) what is the relationship between affective and cognitive domains? The subjects, high school junior and senior girls enrolled in home economics classes, were selected from big city, small city, central rural, suburban and area vocational schools. The first question was answered affirmatively by personal teacher experiences of using the taxonomy. The data for the second question yielded very strong support of the cumulative nature of the cognitive domain. The findings for the third question show that students in each ability group were able to perform equally well at each of the seven cognitive levels. The fourth question data indicates no tendency of support for the inputed hierarchical structure of the affective domain. Analysis of the final question reveals no pattern of relationships. Also listed are implications for further study. (Author/MC)
MEASURING RELATIONSHIPS AMONG AND BETWEEN
COGNITIVE AND AFFECTIVE BEHAVIORS
IN A CONTROLLED LEARNING SITUATION

Ruth P. Hughes, West Virginia University
Helen Y. Nelson, Cornell University

Paper Presented at the Annual Meeting of the
American Educational Research Association
and the National Council on Measurement in
Education, Minneapolis, March 5, 1970.
The Taxonomy of Educational Objectives, Cognitive and Affective Domains (Bloom, et al 1956; Krathwohl, et al, 1964) has presented various research possibilities, most commonly its suitability as a model for statement and evaluation of objectives for classroom instruction and validation of its imputed hierarchical structure. The reported research addressed the question of useability and the question of hierarchical structure plus two others. Was IQ related to performance at the various cognitive levels? and what was the relationship between the two domains?

This research extended an ongoing study of the cumulative nature of cognitive behaviors of adolescents, directed by Blackwell and her associates at Cornell. Although the earlier studies (Byrd, 1963; Jacklyn, 1964) did not use the taxonomy as the model, subsequent studies did (Thomas, 1965; Mehaffey, 1967; Zavacki, 1968). The latter studies were primarily investigations of the cumulative nature of the cognitive domain for which Guttman simplex analysis was used (1954). Thomas, working only with the first three main levels, and in a controlled learning situation, found support for a hierarchy. Others at Cornell did not find support for the existence of a hierarchy in the social science materials used.

One of the first to use simplex analysis in testing for a hierarchy was McGuire (1963), who in a study reported to this group found "considerable support...(for)...a hierarchy of intellectual processes" (p. 15), using an adaptation of the cognitive domain.

In 1966 Stoker and Kropp reported a series of studies designed to explore the construct validity of the cognitive domain. Their simplex analysis yielded evidence described as "more confirming than disconfirming" of the cumulative hypothesis (p. 84). In raising the question
of item difficulty in relation to complexity of mental processes, they noted that one could conceivably bias results by manipulating item difficulty. Crawford attacked that problem in a 1968 analysis from which he concluded that "there is not necessarily a direct relationship between the complexity of intellectual processes and the difficulty of items which purportedly measure them" (1968, p. 107).

Little research dealing exclusively with the affective domain has been reported, and none tested for hierarchical structure. It has been used to specify objectives and to note their attainment (Purcell, 1965; Grady, Hitchens, and Johnson, 1967), and several have discussed its importance in the educational process (for example, Loree, 1965). Only one study reported use of the affective and cognitive domains together. Johnson (1966) assessed the degree of satisfaction in response students got from answering questions at various cognitive levels, and found a low, positive correlation. He also found weak support for the hierarchical structure of the cognitive domain.

Reported procedures and findings indicated that the present research should include a controlled learning situation, an assessment of affect as well as cognition, and control for difficulty of cognitive items. In addition, we required experienced teachers, a topic with a strong affective component, and a large n.

(Although the cognitive domain has six major levels and the affective has five, only the first three levels, plus major sub-levels of the first two were used. That is, the seven levels in the study were: 1.1, 1.2, 1.3, 2.1, 2.2, 2.3, 3.0.)
Specifically the research questions were:

1. Is the Taxonomy of Education Objectives: Affective and Cognitive Domains a reasonable tool for use in stating, and measuring attainment of, objectives?

2. Is the cognitive domain cumulative in social science subject matter?

3. Does ability to operate at the various cognitive levels differ with intelligence?

4. Is degree of affect, as postulated in the affective domain, cumulative?

5. What is the relationship between the affective and cognitive domains?

Sample

Subjects, primarily girls, were 604 high school juniors and seniors enrolled in home economics classes. The unit of instruction was "Preparation for the Dual Role: Homemaker-Wage Earner", a topic with a heavy affective component. The 25 lessons incorporated a variety of instructional methods. The 32 classes were taught by 29 teachers in 26 schools through different regions of New York State. The variety of schools included big city, small city, central rural, suburban, and area vocational.

Nine of the teachers, each with one class and each in a different school, formed the pilot group which helped plan and test the curriculum. One hundred of the students were in this group, and their responses on evaluative instruments were the basis for statistical analyses of those instruments. The pilot teachers were all experienced teachers known to the writer, and they participated at her original request. Schools and teachers in the main study were selected from a group recommended by the Bureau of Home Economics Education, State Education Department. Some of the main study data were combined with pilot data for analysis but most, of necessity, were separated.
Instrumentation

Instruments were designed to measure cognitive and affective behaviors as defined by the cognitive and affective domains of the taxonomy. Included were equivalent forms of an achievement test, a paper and pencil attitude test, and an interview schedule. Responses of students in pilot classes, $n = 100$, were used for test analyses.

Equivalent forms of the achievement test were necessary since pre and post measures were needed to assess student gain but for the purpose of testing the hierarchical structure of the cognitive domain the items from level 2.0 up had to be new. The table of specifications was simple: seven taxonomy levels, each represented in each of six subject matter topics. Tests were equivalent in content, mean, variance, reliability, discrimination indices, and level of difficulty; coefficient of equivalence was .82.

The achievement test was carefully controlled for difficulty at each taxonomy level. Overall difficulty for the posttest which was used in the simplex analysis was 52, with a range among levels from 46 to 56. Reliability was .80 and mean index of discrimination was .36. Construct validity of the taxonomy questions was a problem, and although outside judges were used, the question of their necessity remains unanswered.

The attitude test, with changes in verb tense, was used both pre and post instruction. It was a series of statements at the selected taxonomy levels, to which the students indicated agreement or disagreement. Reliability was estimated at .89 by test-retest. Construct validity was an even greater problem with the attitude tests. Although judges agreed on items, they suggested that verbal distinctions between levels were based on such minute shades of meaning that they may have been lost on the careless or less able reader.
The interview schedule was revised several times during the pilot study, and the interviewers considered it a much better indication of student feelings than the paper and pencil tests. However, no statistical estimation of its reliability was made. Questions dealt with student attitude toward a dual role, with their perception of the desirability of their mothers' role, and with their attitude toward marks in school. Students from each class were selected randomly for the interviews.

The important, practical question asking whether the taxonomy was a reasonable tool for use in stating and measuring attainment of objectives was answered subjectively by the researchers. The decision, based on teacher evaluation of the package and on the personal experience of using the taxonomy, was decidedly affirmative. Although ordinary use by teachers might not be as detailed nor as demanding as use for research, it was considered not only valid but readily usable, particularly within the main categories.

**Findings**

Question two asked whether the postulated cumulative nature of the cognitive domain was supported. With Guttman simplex analysis (1954), the data yielded very strong support of the cumulative nature of the cognitive domain. Effects of type of questions, number of questions per level, and size of sample were all studied. Increasing the number of questions per level improved the simplex structure. Use of sets of questions, each set measuring aspects of the same principle but at different levels of mental complexity yielded an improved simplex structure, but this finding is not conclusive since sample size for the trial was increased from an \( n \) of 100 to one of approximately 250. Effect of sample size was strikingly clear, since an \( n \) of 533 revealed a simplex on the same test item for an \( n \) of 100 revealed no visible tendency toward a simplex.
Table 1

Correlation Matrix for Seven Cognitive Taxonomy Levels,
Six Topics, Pilot and Main Study Students Combined
(n = 553)

<table>
<thead>
<tr>
<th>Taxonomy Level</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>2.1</th>
<th>2.2</th>
<th>2.3</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>.460</td>
<td>.363</td>
<td>.341</td>
<td>.333</td>
<td>.304</td>
<td>.266</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td>.412</td>
<td>.376</td>
<td>.347</td>
<td>.237</td>
<td>.241</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td></td>
<td></td>
<td>.391</td>
<td>.392</td>
<td>.266</td>
<td>.323</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>.393</td>
<td>.270</td>
<td>.246</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.389</td>
<td>.297</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.252</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A "simplex" is Guttman's term for a simple order of complexity, in this case tests representing levels of mental complexity. If such an order obtains, it will be revealed when test scores are intercorrelated, as in Table 1 above. The highest correlations lie along the main diagonal and taper off as one moves across the rows. Coefficients should also decrease in size from bottom to top. Although by no means perfect, inspection of the correlations in Table 1 supports the existence of a hierarchy, and Guttman says "if inspection reveals a hierarchy, then the hierarchy exists" (p. 279).

Possible differences in performance at the selected cognitive levels as a factor of differences in intelligence, the third question, was studied within an analysis of variance format. Although students with more ability, as measured by standard intelligence tests, were able to perform better at each level than those with less ability, students in
each ability group were able to perform equally well at each of the seven cognitive levels.

Table 2
Cognitive Posttest Scores by Taxonomy Level and Student IQ, Main Study Classes

<table>
<thead>
<tr>
<th>Taxonomy Level</th>
<th>IQ Group</th>
<th>1 (Below Average)</th>
<th>2 (Low Average)</th>
<th>3 (High Average)</th>
<th>4 (Above Average)</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td></td>
<td>2.76</td>
<td>3.34</td>
<td>3.30</td>
<td>3.67</td>
<td>3.27</td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td>2.60</td>
<td>3.14</td>
<td>3.37</td>
<td>3.70</td>
<td>3.20</td>
</tr>
<tr>
<td>1.3</td>
<td></td>
<td>2.86</td>
<td>3.12</td>
<td>3.67</td>
<td>3.79</td>
<td>3.36</td>
</tr>
<tr>
<td>2.1</td>
<td></td>
<td>2.61</td>
<td>3.37</td>
<td>3.76</td>
<td>3.84</td>
<td>3.40</td>
</tr>
<tr>
<td>2.2</td>
<td></td>
<td>2.64</td>
<td>3.29</td>
<td>3.74</td>
<td>3.91</td>
<td>3.40</td>
</tr>
<tr>
<td>2.3</td>
<td></td>
<td>2.77</td>
<td>3.11</td>
<td>3.38</td>
<td>3.88</td>
<td>3.29</td>
</tr>
<tr>
<td>3.0</td>
<td></td>
<td>2.49</td>
<td>3.01</td>
<td>2.93</td>
<td>3.39</td>
<td>2.96</td>
</tr>
<tr>
<td>Column Means</td>
<td></td>
<td>2.68</td>
<td>3.20</td>
<td>3.45</td>
<td>3.74</td>
<td></td>
</tr>
</tbody>
</table>

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Required F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rows</td>
<td>.58</td>
<td>6</td>
<td>.10</td>
<td>5.26</td>
<td>.01</td>
</tr>
<tr>
<td>Columns</td>
<td>4.29</td>
<td>3</td>
<td>1.43</td>
<td>75.26</td>
<td>4.01</td>
</tr>
<tr>
<td>Remainder</td>
<td>.33</td>
<td>18</td>
<td>.019</td>
<td></td>
<td>5.09</td>
</tr>
<tr>
<td>Total</td>
<td>5.22</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 shows the analysis for students in the main study group for whom IQs were available. Although both F's were significant, the Newman-Keuls range test revealed no differences among row means. This finding, coupled with the strong support for the hierarchical structure of the cognitive domain, was seen to have important practical implications for education.
Question four was asked to determine if there were support for the imputed hierarchical structure of the affective domain. Using the sublevels, no tendency toward a hierarchy was found, irrespective of sample size. When sublevels for 1 and 2 were combined and only three levels were used, there was some indication of a possible hierarchy.

The final question was designed to study the relationship between the two domains of the taxonomy. Exploratory analyses of pilot data revealed no pattern of relationships which would permit analysis beyond that of simple correlation. As expressed by a correlation coefficient of .24, the relationship between student achievement on test items and expressed attitudes toward the topics studied must be described for the data in this study as low, positive, and statistically significant.

Some indication of the complexity of the relationship was indicated by responses to the interview questions about the importance of marks in school. Since students indicated that they would work for marks irrespective of their liking for a subject, and since some students who liked a subject might do poorly, some means other than that used in the present study must be found to determine the relationship between the two domains.

**Implications for Further Study**

The lack of definitive relationship between the two domains points to one recommendation for further study, which is that in assessing affect in school subjects, extrinsic factors which may be of greater importance than intrinsic interest in or commitment to a given topic of study must be recognized.
Another recommendation is that test specialists work on development of objective test items which will measure analysis and synthesis. This research found support for the hierarchical structure of the cognitive domain through level 3.0 using the sublevels; others have found support using major categories through 3.0 and beyond. It would now seem reasonable to accept the hierarchical structure as existent, and to use research time for development of test items at the higher levels.

In addition to developing test items assessing higher mental processes, there is need for inclusion of such test items in both criterion and diagnostic tests developed for use with curriculum materials. And although the taxonomy is a reasonable tool to use in developing and testing materials for a unit of study, other models such as Gagné's might be equally desirable, at least for a single lesson or a brief series of lessons.

For less able learners, development of materials which include provision for use of the more complex mental processes seems indicated. In home economics, materials in the subject areas of family relationships, child development, and consumer economics might well be developed.

Further, considering the increasing importance of individual instruction, the need for and acceptance of carefully constructed curriculum materials for all learners -- but perhaps especially for the less able, the greater value of diagnostic as opposed to norm reference tests, the following recommendation is made most strongly. Based on the experiences with and results of this project, it would seem that the education researcher interested in practical problems might well use available resources to prepare curriculum materials in an important area for use by a variety of students. Achievement of content objectives should be tested
with measurement instruments which assess ability to operate at various levels of mental complexity for the purpose of diagnosing learners' problems and then moving to correct them.
References


Purcell, John, "An Experiment in Developing Attitude Scales for Classroom Use", Phi Delta Kappan, vol. 46, no. 10 (June 1965), pp. 533-534.
