An expanded abstract including statements of purpose, rationale, design and procedure, findings, and interpretations is given for 18 selected research articles. Following each, a short critical analysis is presented by a professional in the field. Among the research topics covered are achievement, attitudes, laboratories, error trends, and learning styles. This quarterly publication has added a new feature beginning with this issue. A listing of all mathematics education research reports that have been included in ERIC publications (RIE and CIJE) the preceding quarter is given to help readers keep abreast of continuing research efforts. (LS)
INVESTIGATIONS IN MATHEMATICS EDUCATION

Expanded Abstracts and Critical Analyses of Recent Research

Center for Science and Mathematics Education
The Ohio State University
in cooperation with
the ERIC Science, Mathematics and Environmental Education Clearinghouse
NOTES from the editor


Usiskin, Zalman P. *The Effects of Teaching Euclidean Geometry Via Transformations on Student Achievement and Attitudes in Tenth-Grade Geometry*. *Journal for Research in Mathematics Education* 3: 249-259; November 1972. Abstracted by STEPHEN S. WILLOUGHBY. 63
In this issue, *Investigations in Mathematics Education* begins a new feature. Using the ERIC documents *Research in Education* and *Current Index to Journals in Education*, we will list the mathematics education research reports that have appeared during the preceding quarter. By following these listings, our readers can keep abreast of continuing research efforts in a simple, but systematic way.

This systematic listing of research reports will complement the present research reviews that appear in the journal. Articles and documents that we review are assigned in chronological order to our panel of reviewers. Our reviewers act independently on their assigned articles, and the order in which the reviews are returned to us helps determine the selection of reviews that will appear in any particular issue of I.M.E. If a reviewer declines to review an assigned document, the document must be reassigned to another reviewer, delaying its final appearance significantly. As a result of these processes, the sequence and order of appearance of reviews in this journal becomes somewhat randomized.

This lack of systematic coverage has concerned the advisory board as well as the editor for some time. It is hoped that the quarterly listing of research reports will help solve much of this problem. Nevertheless, some shortcomings remain in the quarterly listings, and the reader should be aware of these in order to use the list intelligently. Among the more important shortcomings are these:

1. The listing does not contain a comprehensive coverage of mathematics education dissertations appearing in *Dissertation Abstracts International*. Only selected dissertations are announced in RIE, and our quarterly listing reflects this selection.

2. The quarterly listing is systematic, but not necessarily current. The indexing of journal articles in CIJE may be delayed by as much as three months. Since each I.M.E. issue will list the reports appearing in the previous quarter, our listing of journal reports may be delayed by as much as six months. The use of CIJE to generate our listings, however, gives us coverage of more than 500 journals in education. We believe that this extremely broad coverage offsets the time delays.

3. The listing is limited to reports of research studies which analyze quantitative data. It does not contain research reviews, summaries, or position papers. Many such peripheral documents are announced each month in RIE and CIJE, and readers who wish to locate such documents should refer directly to these publications.

Finally, although the list will be as comprehensive as we can make it, no information system is infallible. We would appreciate hearing from readers about research reports (other than dissertations) which have been omitted from the list.

Jon L. Higgins
Editor

ED 067 161  The Status of Number and Quantity Conservation Concepts Across the Life-span. 17p. MF and HC available from EDRS.

ED 067 219  The Effects of Information Concerning the Attributes of Concept Instances and Recall of Relevant Subconcepts on the Level of Mastery of Certain Geometric Concepts. Working Paper 45. 66p. MF and HC available from EDRS.

ED 067 247  The Development of Selected Initiating Activities in the Teaching of Mathematics. 173p. MF and HC available from EDRS.

ED 067 248  Games and Teams: A Winning Combination. 40p. MF and HC available from EDRS.

ED 067 265  Conservation of Identity and Equivalence Among Children from Varying Socio-Economic Backgrounds. 19p. MF and HC available from EDRS.

ED 067 397  A Comparison of Interview and Normative Analysis of Mathematics Questions. 49p. MF and HC available from EDRS.


ED 067 802  Academic Achievement Test Results of a National Testing Program for Hearing Impaired Students. United States: Spring 1971. 64p. MF and HC available from EDRS.

ED 067 870  Student Performance in Computer-Assisted Instruction in Programming. 96p. MF and HC available from EDRS.

ED 067 874  Second Year Evaluation, IPI Mathematics Project; Hall School, 1970-1971. 60p. MF and HC available from EDRS.

ED 068 157  Children's Concept of Number: The Spontaneous Production of Number Symbols in Their Drawings. 31p. MF and HC available from EDRS.

ED 068 354  Computerized Instruction in Mathematics Versus Other Methods of Mathematics Instruction Under ESEA Title I Programs in Kentucky. 59p. MF and HC available from EDRS.
ED 068 440  The Influence of Professional Reference Groups on Decisions of Preservice Teachers. 14p. MF and HC available from EDRS.

ED 068 456  Validation of Minicourse Five for Special Education. Final Report. 152p. MF and HC available from EDRS.

ED 068 491  The Development and Application of a Structured Procedure for the In-Context Evaluation of Instructional Materials. 109p. MF and HC available from EDRS.

ED 068 498  The Effects of Special Instruction for Three Kinds of Mathematics Aptitude Items. 86p. MF and HC available from EDRS.

ED 068 907  The Effects of Instruction by Teachers and Teacher Aides Upon the Performance of Pupils in a Direct Instructional Program. 173p. Not available from EDRS. Available from University Microfilms (72-12,278).

ED 069 496  Some Computational Strategies of Seventh Grade Pupils. Final Report. 96p. MF and HC available from EDRS.

ED 069 507  An Investigation of Various Cognitive Styles and the Implications for Mathematics Education. 126p. Not available from EDRS. Available from University Microfilms (71-17,317).

ED 069 508  The Employment of Non-Standard English in the Development of a Mathematics Course for Seventh-Grade Disadvantaged Students. 230p. Not available from EDRS. Available from University Microfilms (71-17).

ED 069 509  The Effect of Activity Oriented Lessons on the Achievement and Attitudes of Seventh Grade Students in Mathematics. 204p. Not available from EDRS. Available from University Microfilms (71-18755).

ED 069 510  An Experimental Briefing-Teacher Aide Program with Professional Laboratory Experiences for Sophomore-Junior Level Pre-Service Mathematics Teacher Trainees. 302p. Not available from EDRS. Available from University Microfilms (71-19,511).


ED 069 513  A Multi-Experience Approach to Conceptualization for the Purpose of Improvement of Verbal Problem Solving in Arithmetic. 135p. Not available from EDRS. Available from University Microfilms (72-960).
ED 069 514 An Investigation of the Effect of Selected Learning Styles on Achievement in Eighth Grade Mathematics. 270p. Not available from EDRS. Available from University Microfilms (72-351).


ED 069 516 The Ability to Estimate in Mathematics. 108p. Not available from EDRS. Available from University Microfilms (72-4178).

ED 069 517 An Experimental Comparison of Two Methods of Teaching the Addition and Subtraction of Common Fractions in Grade Five. 216p. Not available from EDRS. Available from University Microfilms (72-2954).

ED 069 518 Effects of Varying Concrete Activities on the Achievement of Objectives in Metric and Non-Metric Geometry by Students of Grades Five and Six. 242p. Not available from EDRS. Available from University Microfilms (72-3248).

ED 069 519 The Ability of Selected Sixth Grade Pupils to Function at a Variety of Cognitive Levels on Selected Mathematical Tasks. 158p. Not available from EDRS. Available from University Microfilms (72-3250).

ED 069 520 The Use of an Anecdotal Style of Content Presentation as a Motivational and Instructional Device for Seventh Grade Underachievers in Mathematics. 214p. Not available from EDRS. Available from University Microfilms (71-27,530).

ED 069 521 Differential Performance of Third-Grade Children in Solving Open Sentences of Four Types. 182p. Not available from EDRS. Available from University Microfilms (72-1028).


ED 069 723 Correlates of Achievement in an IP: School. 29p. MF and HC available from EDRS.

ED 069 789 Sex Differences in Mathematics Achievement—A Longitudinal Study. 20p. MF and HC available from EDRS.

ED 069 831 A Study of the Effects of Compensatory Instruction in Language Arts and in Arithmetic on Achievement, Study Habits, and Selected Attitudes of Eighth Grade Students in a Depressed Area School. 374p. Not available from EDRS. Available from University Microfilms.


EJ 067 212  Intellectual Development Beyond Elementary School III--
Ratio: A Longitudinal Study. School Science and Mathematics,
v72 n8, pp735-742, Nov 72.


EJ 067 218  Do Students Learn from and like an Audio-Tutorial Course
in Freshman Mathematics? Two-Year College Mathematics Journal,
v3 n2, pp37-41, F 72.
THE EFFECT OF PROGRAMED REVIEW ON 4TH AND 5TH GRADE ARITHMETIC RETENTION. Bausell, R. Barker; Moody, William B., School Science and Mathematics, v72 n2, pp148-150, Feb 72


Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jerry P. Lecker, Rutgers University

1. Purpose

To explore the effectiveness of programed review materials in elementary mathematics instruction.

2. Rationale

The researchers observe that considerable research seems to show programed instruction to be neither more nor less effective than conventional teacher oriented instruction. However, they point out that programed review may hold promise as a valuable supplement to conventional teaching. For example, programed review work could be individualized and thus provide an efficient adjunct to conventional teaching. Moreover, programed review for elementary school children has not really been tried out on a large enough scale as yet. The researchers also observe that this approach to review may prove to be more palatable to both students and teacher because of obvious differences between the programed technique and the original instruction.

3. Research Design and Procedure

Subjects: Six classes (144 students) were randomly assigned to programed review (R) and six classes (136 students) were randomly assigned to a control group (C). The classes were selected from 30 fourth and fifth grade classrooms that were a part of a more extensive study of teacher proficiency in the Newark Special School District, Newark, Delaware.

Teachers: Teachers were assigned for the five days of instruction who were not normally assigned to these classrooms. They were given teacher's manuals containing minimum content to be covered for the 5 day unit, as well as a list of 12 specific objectives on which students would later be tested. Teachers were free to use whatever teaching techniques they desired.
Instructional and Review Materials: The original instruction dealt with modular arithmetic, covering five days from start to finish, 45 minutes per day. The review programs aimed at reviewing the twelve objectives of the original 5 day unit on modular arithmetic. The five programs contained about 14 linear instructional frames and required no more than 15 minutes each to administer. No new material was added, and each program ended with a 2-3 item mini-test, for self evaluation.

Design: The design and procedures can be illustrated as follows:

30 classes of 4th and 5th graders, receive 5-day unit on modular arithmetic: M-F, 45 mins. 1 day

20-item learning test administered to all students 3 days later, on Monday

Selection of treatment and control groups

Review Treatment (R)
6 classes (144 students) selected at random received program #1 10-11 days after test, and programs 2-5 every 7-9 days thereafter, in addition to regular classroom work (over 8-week period).

Control Group (C)
6 classes (136 students) received no further work on modular arithmetic--only regular classroom work (over 8-week period).

Same 20 item original learning test administered to all subjects in R and C groups.

4. Findings

The R and C groups were taken as "equivalent" after administration of the test the first time since the class means were not significantly different. Also, there were no differences between the two groups in terms of IQ and SCAT Mathematics scores.
Difference scores, computed by subtracting students' scores on the second test from their scores on the first, were used as measures of retention after the second administration of the learning test. Results showed that group R subjects retained significantly more ($p < .01$) modular arithmetic learning than did group C subjects. There was no overall retention loss for group R, whereas there was for group C.

5. **Interpretations**

The researchers interpret their results as strong evidence for the effectiveness of programed review. They further observe that the technique appears to be both practical and economical in terms of classroom time. The results argue for more widespread use of programed review in the elementary schools.

**Abstractor's Notes**

The study appears to be a nice, neat one, with little to critically analyze. However, a few comments seem to be in order. It would be valuable to know more about the content (and organization of the content) contained in the original 5-day unit on modular arithmetic. It might also be valuable to know what the specific objectives were that teachers taught for. Further, it would be valuable to know more about the content and its related organization in the programed review (i.e., some sample frames). Finally, the reader should keep in mind that the programed review was not compared with an alternate approach to review. Suppose for example, that teachers in another treatment group(s) had informally or formally reviewed the material in a conventional manner; would subjects in those groups have retained as much as or more than subjects in the R groups?

Jerry P. Becker
Rutgers University
1. Purpose

To investigate the effects of presentation mode (vertical, horizontal) and placeholder position (left, middle, right) upon first and second grade children's ability to solve number sentences as measured by rate and type of error. To investigate the effects of those variables over the operations of addition and subtraction.

2. Rationale

Square "frames" are used as placeholders in open addition and subtraction sentences in most elementary mathematics programs. Since these number sentences are presented with variation of presentation mode (horizontal and vertical) and position of placeholder (left, middle, right), it is hypothesized that these variables will produce differing types and rates of error by students attempting to solve those open sentences. An analysis of the effect of these variables upon type and rate of error should (a) prove valuable to developers of mathematics curricula, and (b) suggest hypotheses to be tested experimentally.

3. Research Design and Procedure

Workbooks of 132 first and second grade children in two southern Illinois school systems were analyzed. These workbooks contained the entire year's written work of those children in mathematics. All analysis was done by two researchers. The errors were classified as being one of three types:

Computation (C) - not knowing "facts"
Process (P) - using addition instead of subtraction or vice versa
Random (O) - errors that could not be classified as either C or P
The data were organized and tabulated, but no statistical analyses were made.

The children selected for study were not randomly chosen; rather the workbooks were analyzed if they were available. However, a wide range of ability was represented.

4. Findings

a. Occurrence of Sentence Types

The bulk of first grade work (80%) is presented horizontally while second grade is fairly evenly distributed between H and V. The L type of sentence rarely appears in either grade; in fact, the VL format does not appear at all in second grade. When the V format is used in both grades, the great percentage of the sentences are classified as R sentences, that is, the placeholder represents the missing sum for addition and the missing difference for subtraction.

b. Error Rate

<table>
<thead>
<tr>
<th></th>
<th>Addition</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>10.2</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>10.8</td>
<td>5.64</td>
</tr>
<tr>
<td>H</td>
<td>13.2</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td>12.3</td>
<td>7.3</td>
</tr>
<tr>
<td>Grade</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Addition</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>26.2</td>
<td>8.0</td>
</tr>
<tr>
<td></td>
<td>25.1</td>
<td>22.6</td>
</tr>
<tr>
<td>M</td>
<td>20.0</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>13.1</td>
<td>6.2</td>
</tr>
<tr>
<td>R</td>
<td>7.6</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td>12.0</td>
<td>6.2</td>
</tr>
<tr>
<td>Grade</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Comparison of Error Rate for Horizontal (H) and Vertical (V) Presentation Formats
Comparison of Error Rate by Grade Level, Operation, Placeholder Placement and Presentation Mode

<table>
<thead>
<tr>
<th></th>
<th>HL</th>
<th>HM</th>
<th>HR</th>
<th>VL</th>
<th>VM</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Addition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr 1</td>
<td>31.0</td>
<td>19.55</td>
<td>8.30</td>
<td>18.55</td>
<td>23.83</td>
<td>4.84</td>
</tr>
<tr>
<td>Gr 2</td>
<td>8.04</td>
<td>5.34</td>
<td>7.66</td>
<td>none</td>
<td>3.59</td>
<td>6.01</td>
</tr>
<tr>
<td><strong>Subtraction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr 1</td>
<td>14.73</td>
<td>15.52</td>
<td>12.39</td>
<td>28.96</td>
<td>8.09</td>
<td>10.22</td>
</tr>
<tr>
<td>Gr 2</td>
<td>22.66</td>
<td>6.68</td>
<td>6.99</td>
<td>none</td>
<td>3.70</td>
<td>5.32</td>
</tr>
</tbody>
</table>

c. Error Type

Distribution of Error Types for Horizontal and Vertical Presentation Formats

<table>
<thead>
<tr>
<th></th>
<th>Addition</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>58.8</td>
<td>71.8</td>
</tr>
<tr>
<td>P</td>
<td>37.4</td>
<td>19.8</td>
</tr>
<tr>
<td>O</td>
<td>3.8</td>
<td>8.4</td>
</tr>
<tr>
<td>H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>73.5</td>
<td>79.9</td>
</tr>
<tr>
<td>P</td>
<td>19.4</td>
<td>12.7</td>
</tr>
<tr>
<td>O</td>
<td>7.1</td>
<td>7.4</td>
</tr>
<tr>
<td>Grade level</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Distribution of Error Types by Placeholder Placement

<table>
<thead>
<tr>
<th></th>
<th>Addition</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>69.7</td>
<td>69.7</td>
</tr>
<tr>
<td>P</td>
<td>23.8</td>
<td>23.6</td>
</tr>
<tr>
<td>O</td>
<td>6.5</td>
<td>6.7</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>64.5</td>
<td>81.8</td>
</tr>
<tr>
<td>P</td>
<td>31.3</td>
<td>8.6</td>
</tr>
<tr>
<td>O</td>
<td>4.2</td>
<td>9.6</td>
</tr>
<tr>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>82.8</td>
<td>77.5</td>
</tr>
<tr>
<td>P</td>
<td>7.3</td>
<td>15.8</td>
</tr>
<tr>
<td>O</td>
<td>9.9</td>
<td>6.7</td>
</tr>
<tr>
<td>Grade level</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Comparison of Error Type by Grade Level, Operation, Placeholder Placement and Presentation Mode

<table>
<thead>
<tr>
<th></th>
<th>HL</th>
<th>HM</th>
<th>HR</th>
<th>VL</th>
<th>VM</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Addition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr 1</td>
<td>C 84.2</td>
<td>67.3</td>
<td>81.6</td>
<td>31.3</td>
<td>46.6</td>
<td>88.5</td>
</tr>
<tr>
<td>P 10.5</td>
<td>27.7</td>
<td>7.6</td>
<td>59.4</td>
<td>51.6</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>O 5.3</td>
<td>10.8</td>
<td>9.3</td>
<td>1.8</td>
<td>5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr 2</td>
<td>C 61.8</td>
<td>62.0</td>
<td>39.0</td>
<td>100</td>
<td>67.3</td>
<td></td>
</tr>
<tr>
<td>P 36.4</td>
<td>17.3</td>
<td>51.2</td>
<td>0</td>
<td>0</td>
<td>29.2</td>
<td></td>
</tr>
<tr>
<td>O 1.8</td>
<td>20.7</td>
<td>9.8</td>
<td>0</td>
<td>0</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td><strong>Subtraction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gr 1</td>
<td>C 100</td>
<td>85.1</td>
<td>78.1</td>
<td>64.0</td>
<td>67.5</td>
<td>74.7</td>
</tr>
<tr>
<td>P 0</td>
<td>8.3</td>
<td>15.1</td>
<td>28.0</td>
<td>12.5</td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>O 0</td>
<td>8.6</td>
<td>6.8</td>
<td>8.0</td>
<td>20.0</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Gr 2</td>
<td>C 79.3</td>
<td>68.5</td>
<td>54.8</td>
<td>0</td>
<td>40</td>
<td>82.4</td>
</tr>
<tr>
<td>P 10.3</td>
<td>8.7</td>
<td>27.8</td>
<td>0</td>
<td>40</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>O 10.4</td>
<td>22.8</td>
<td>17.4</td>
<td>0</td>
<td>20</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>
5. Interpretations

The types of number sentences are disproportionately distributed. Thus, much of the error would seem to be attributable to a lack of opportunity to learn or to practice. Even with roughly equivalent amounts of practice exercises, certain number sentence formats appear to present more difficulty than others. The reasons for this disparity are not immediately evident, suggesting the need for more controlled empirical research.

Abstractor's Notes

The writers have performed a service in accomplishing what they set out to do. Their findings seem to be in agreement with recent findings of Weaver and Grooms. It is disappointing that they did not extend their investigation somewhat to determine the amount of instructional time devoted to each of these sentence types. Workbook pages often provide nothing more than practice over ideas that were taught elsewhere. While it is likely that instructional time is highly correlated with the number of workbook exercises devoted to a skill, it would have been helpful to have some approximation of the measure of the variable of instructional time.

One statistic that proved rather surprising was the large proportion of Process errors in R type sentences involving both addition and subtraction with second graders. It would be interesting to know if the numbers involved were the "facts" less than 10 or the so-called "teen" facts.

James M. Moser
The University of Wisconsin
1. Purpose

To investigate the relationship between teachers' understanding of modern algebra and their students' achievement in ninth grade algebra.

2. Rationale

"The search for characteristics which distinguish effective teachers has gone on for at least two thirds of a century. It has not been fruitful. We still do not have any way of distinguishing, in advance, the more effective from the less effective teachers.

"Curiously, one variable, which at first glance seems very relevant, has not received much attention. The variable is the degree to which the teacher understands the material being taught."

Seven other studies, all relating to elementary or junior high school, are cited in the bibliography. Algebra was selected because of the larger pool of teachers available and because of the general uniformity of beginning algebra courses across the nation.

3. Research Design and Procedure

Two algebra tests, one on the real number system and the other on groups, rings and fields, were administered to 375 algebra teachers in NSF summer institutes in 1970.

A mathematics inventory drawn from NLSMA scales and a Reference Test for Cognitive Factors were administered to the ninth grade algebra students of these teachers in the fall of 1970. Two algebra tests were administered to the students in the spring of 1971.
One was devoted to algebraic computation and the other to understanding concepts.

Of the original 375 teachers, 308 provided the data of the study. The other 67 were withdrawn for a wide variety of reasons.

Each student was assigned the following scores:

- M - fall mathematics test
- R - fall arithmetic reasoning test
- I - fall induction test
- V - fall verbal test
- C - spring algebraic computation test
- N - spring non-computation test.

Each teacher was then assigned score \( m_C \) which was the mean score for the males in that class and \( f_C \) as the mean score for the females in that class. In a similar manner, each teacher was assigned \( m_M, m_R, m_I, m_V, m_N \).

To account for variation in the initial status of classes, the regression of \( m_C \) on \( m_M, m_R, m_I \) and \( m_V \) was computed. The resulting regression equation was used to compute a "predicted male computation" score (\( P_{mC} \)) for each teacher. Likewise, \( P_{fC}, P_{mN} \) and \( P_{fN} \) were computed for each teacher. The "effectiveness for male students in computation" (\( E_{mC} \)) was then determined by the equation

\[
E_{mC} = m_C - P_{mC}.
\]

Similarly, \( E_{fC}, E_{mN} \) and \( E_{fN} \) were computed.

Correlations between the effectiveness scores and the teachers' scores in algebra were computed. Stepwise regression of each of the four effectiveness scores on each of the two teacher algebra scores was carried out.

4. Findings

(a) The pretests given the students turned out to be good predictors of success, so the effectiveness scores are quite meaningful.

(b) There was substantial variation in the effectiveness of teachers.

(c) Teacher effectiveness with male students was not substantially different from effectiveness with female students.
(d) Teacher understanding of modern algebra (groups, rings and fields) has no significant correlation to student achievement in ninth grade algebra.

(e) Teacher understanding of the algebra of the real number system has no significant correlation with ninth grade students' algebraic computation skills.

(f) Teacher understanding of the algebra of the real number system is significantly correlated with student achievement in understanding algebraic concepts, but the correlation is so low it is educationally insignificant.

5. Interpretations

If there is some minimum amount of knowledge of algebra that teachers need to teach effectively, apparently the teachers in this study were above that point.

Teachers should be provided with a solid understanding of the courses they are expected to teach but selection and retention should be based more on actual classroom performance.

Abstractor's Notes

The study was obviously carried out with great care and the report is thorough and concise. The size of the sample is impressive. This writer observed the data gathering in one institute setting and was impressed by the serious professional concern displayed by the teacher participants.

Discouraging results such as these seem to compel the mathematics teacher training community to do some very hard thinking—the kind one learns by studying logic.

William M. Fitzgerald
Michigan State University
SOME CONSEQUENCES OF LEARNING THEORY APPLIED TO DIVISION OF FRACTIONS. Bidwell, James K., School Science and Mathematics, v71 n5, pp426-434, May 71


Expanded Abstract and Analysis Prepared Especially for I.M.E. by John G. Harvey, The University of Wisconsin

1. Purpose

The study investigated differences in meaningful learning between the common denominator method, the complex fraction method and the inverse operation method of dividing common fractions.

2. Rationale

The author identifies the teaching of the division of fractions as a long-standing problem in the elementary school mathematics program and states that, while considerable research has been done in this area, no real consensus of opinion exists. In order to find a solution a study based on the following three assumptions is necessary:

1. The only proper learning methods for the modern school are those which emphasize student understanding of all the content learned.

2. Recent developments in the theory of cognitive learning are important in finding a solution to the problem.

3. The teaching of the division of fractions is to be done via meaningful reception learning as opposed to rote reception learning or meaningful discovery learning.

3. Research Design and Procedure

The research design and procedure is incompletely described. For example, the selection of the sample, the grade level of the Ss, the duration of the experiment, the design of the treatments, the tests administered and the data analysis procedures are not included in the paper. Thus portions of this part of the abstract are inferences made by the abstractor.
First, the investigator task analyzed each of the three methods of dividing fractions using techniques advocated and developed by Gagne (1965). Then he applied the theory of Ausubel (1968) and examined the lower levels of each task analysis for effective advance organizers. Even though the investigator concluded that neither the common denominator method nor the complex fraction method have effective advance organizers, based on these logical analyses an instructional treatment was prepared for each of the three methods. The common denominator treatment was taught to 155 Ss, the complex fraction treatment to 141 Ss, and the inverse operation treatment to 152 Ss.

The investigator indicates that each instructional treatment and its accompanying task analysis were evaluated "using methods suggested by Gagne." In addition a 13 item computational skill posttest was administered to all Ss as was a similar retention test three weeks after the posttest. The mean scores of groups were compared on the posttest using ANOVA and using ANCOVA with IQ being the covariate. A similar analysis of the means was used to compare the performance on the retention test. Finally, a comparison was made among students with IQ's below 100 using both the posttest and retention test scores.

4. Findings

The findings given by the investigator are:

1. The inverse operation structure was a more effective learning structure than either of the two other structures.

2. On the posttest and on the retention test the mean and the adjusted mean score of the inverse operation treatment groups is significantly better than that of the complex fraction treatment group.

3. On the posttest and on the retention test the mean of the low IQ inverse operation subgroup is significantly better than that of the low IQ complex fraction subgroup.

5. Interpretations

It can be concluded that:

1. More effective classroom procedures can be discovered by the use of the results of learning theories such as those of Gagne and Ausubel.
2. The division of fractions can best be approached by teaching it as a part of the elementary mathematics structure, emphasizing its relationship to multiplication as an arithmetic operation and other properties of a number field.

3. Instead of requiring computational competency in division of fractions, it may be better to take the opportunity to extend the mathematics structure of inverse operation and at the same time provide minimum competency in a most effective manner.

Abstractor's Notes

It is impossible to raise questions about this experiment when a detailed description of it is not given.

John G. Harvey
The University of Wisconsin
1. Purpose

To investigate the following factors:

(a) Whether conservation failures are a result of young children's dependence on perceptual comparisons.

(b) Whether the relation between quantities being compared affects performance on conservation and measurement problems.

(c) Whether recognition of compensatory relationships between dimensions plays any significant part in conservation judgments.

2. Rationale

The author points up the results of two studies by Berison (1969) and Carpenter (in press) which appear counter to the generally accepted Piagetian notion that non-conservers rely almost exclusively upon perceptual cues when solving conservation problems. Evidence from the two cited studies indicates that young children may attend to certain numerical cues more readily than perceptual cues. The basis for investigation of the second factor was provided by the apparently conflicting data of Carpenter (in press), Beilin (1968), Piaget (1968) and Rothenberg (1969) and that of Zimiles (1966), Carey and Steffe (1968) and Harper and Steffe (1968) regarding the differential difficulty of conservation problems dependent upon equivalence and order relations. Piaget's assertion (1952) that recognition of compensatory relationships is a significant factor in the development of conservation provided the basis for the investigation of the third factor.

3. Research Design and Procedure

Throughout the study the ideas investigated were referred to as factors. Five types of problems were developed to facilitate the investigation of the factors.
(a) Traditional volume of liquid conservation task. No unit of measure involved.

(b) Similar to (a) but two different units of measure employed, the larger of which was visually distinguishable from the smaller.

(c) Similar to (b) but larger unit was not visually distinguishable.

(d) Similar to (a) but single measuring unit was used for transfer of liquid.

(e) Similar to (d) but liquid first presented in differently shaped containers and then transferred to similar containers.

Three different types of items were developed for use with all or some of these problem types.

i. Equivalence—equal quantities were transformed to appear unequal.

ii. Nonequivalence I—unequal quantities were transformed so that the dominant dimension in each quantity was equal.

iii. Nonequivalence II—unequal quantities were transformed so that the difference of the inequality appeared to be reversed.

The subjects for the study were 75 first graders and 54 second graders who formed intact class groups from a school system in Wisconsin. The subjects were randomly assigned to two testing groups. One group received Part A of test (all problem-type a, b and c questions), the other group Part B (problem-types b, c, d and e but only items types i and iii). The design for Part A was a 3 x 3 repeated measures and for Part B a 2 x 4 repeated measures design. The analysis of the data was by a MANOVA procedure.

4. Findings

Part A of test:

1. Significant differences were observed between items of type ii and those of types i and iii for all type a and b problems.

2. No significant differences between type b and type c problems with respect to item types i and iii.
3. No significant differences between problem-types a and b for any of the three relations (item-type i, ii and iii).

Part B of test:

Significant differences were observed between problem-types b, c, d and e.

5. Interpretations

1. Factor A

Evidence was provided that the order of presentation of cues in conservation tasks is the major factor in determining to which cues children attend. It was also suggested by the data that problems in which correct cues are numerical are significantly easier to solve than those in which correct cues are visual for either conservation or measurement problems. These findings provided conflict with those of Piaget and Bruner, Oliver and Greenfield which the investigator suggests may be possible because in experiments of other investigators all distracting cues were visual. The investigator also suggested that measurement is meaningful for the majority of the students in first and second grades. This, too, conflicts with the observations of Piaget; however, the measurement tasks used by Piaget were less structured than those used by the investigator.

2. Factor B

The relation between quantities did not appear to affect performance. Non-equivalence I items were easier, but it was suggested that this may have been because that type of question did not require genuine conservation.

3. Factor C

Ambiguous results regarding the role of compensation in conservation were gained, and the investigator suggests that caution should be exercised in interpreting the results of study pertaining to this factor.

Abstractor’s Notes

Much criticism has been made in the past of Piaget’s work based on lack of experimental manipulation, sample description and procedural description. This study by Carpenter involved considerable variable manipulation with respect to problem types which
facilitated the presentation of considerable evidence, much of which is not supportive of Piaget's general ideas regarding conservation. The generalizability of the data, however, has to be guarded unless in a more complete account of the study there is more detailed information regarding the subjects. It seems possible that the early acquisition of measurement ideas and the frequency of use of numerical cues could be explained on the basis of background experience. The investigator claims that protocol and procedure were kept as constant as possible. It is assumed that a detailed account of these is available in a more complete report of the investigation.

It would seem to be of value to ask some further questions to:

(a) Determine whether similar findings could be observed with respect to measures of area and length.

(b) Determine the role of compensation in conservation—this might help in the resolution of some of the conflicting data.

Heather L. Carter
University of Texas at Austin
Expanded Abstract and Analysis Prepared Especially for I.M.E. by H. Laverne Thomas, State University of New York, Oneonta

1. Purpose

To study the nature of prospective elementary school teachers' beliefs about mathematics and mathematics instruction on a formal-informal dimension. Further, to study the degree of ambivalence with which such beliefs are held.

2. Rationale

The specific dimension of formal-informal was chosen because of the researcher's belief that mathematics education is changing emphasis towards a less formal view of school mathematics. Previous observations that some individuals appear to respond in the same way to sets of statements that are judged contradictory motivated a study of the ambivalence of responses and development of a measure of such ambivalence.

3. Research Design and Procedure

Two Likert-type scales were developed, one each for beliefs about mathematics (BAMS) and for beliefs about mathematics instruction (BAMIS). From an item pool, 80 items were validated by a panel of judges, an equal number in each of four categories: mathematics, formal and informal; and mathematics instruction, formal and informal. Item analysis following a trial administration of the scales to 200 students in elementary education allowed selection of final scales with ten items in each category.

The scales were administered to all students in elementary education at the researcher's university. This population was partitioned on (1) amount of instruction received in college mathematics and in mathematics teaching (four stages), and (2) level of mathematics achievement relative to amount of instruction received.
(two levels for each stage of instruction). Stage I subjects had received no college mathematics instruction, Stage II—one mathematics course, Stage III—two mathematics courses, and Stage IV—two mathematics courses and a methods course as well, in a three course sequence. The number of students available in the cell representing low achievers at the highest level of instruction determined the number of students per cell (33) in a 4 x 2 factorial design.

Responses for each scale item were on a six point interval ranging from strongly disagree to strongly agree. In each scale items representing a formal view were deemed negative and received a score of seven minus the scale value checked. Item scores were totaled separately for each scale (BAMS and BAMIS). For each scale a total score of 70 was considered neutral with scores greater than 70 in the informal direction.

A numerical comparison of the responses of each individual to positive and negative items judged to describe the same belief in different ways provided a measure of ambivalence of response, called the ambivalence quotient (AQ). It was hypothesized that the major component of the AQ would be the individual’s consistency of response rather than individual-judge disagreement on interpretation of items. An AQ was computed for each individual on each of BAMS and BAMIS.

A two-way ANOVA was carried out for each scale for total score and for the AQ. In each case preparation (four stages) and achievement (two levels) were the independent variables.

4. Findings

The ANOVA for total scale score on the BAMS and on the BAMIS showed significant differences (p<.01) on the preparation variable on both scales and on the achievement variable for the BAMS. Individual t tests between stages on the preparation variable showed significant differences (p<.05) between stages III and IV (the highest stage) on both scales. Individual t tests between achievement levels at each preparation stage yielded significant differences (p<.05) at each of preparation stages II, III, and IV on BAMS. Mean total scores generally increased with stage of preparation on both scales and were higher for high achievers than for low achievers.

The ANOVA for AQ yielded significant differences for the BAMS on the preparation variable (p<.01). For the BAMIS, differences were significant on the achievement variable (p<.01) and on the
preparation variable (p<.05). On both scales, lower AQ was identified with the later stages of preparation. On BAMIS, lower AQ was identified with the high achievers.

5. Interpretations

With appropriate cautions about generalizations since the sampling at different preparation stages and achievement levels is very possibly from quite different populations, several inferences are drawn from the data. In sum, for beliefs about mathematics, students in elementary education move from an entering neutral position to a slightly informal position, with high achievers holding a more informal stance than low achievers. For beliefs about mathematics instruction, students also moved from an initial neutral position to a slightly informal position.

The AQ adds to these inferences the further conclusion that after concluding the program, students beliefs about mathematics instruction are less ambivalent than upon entering. Examinations of the differences between stages III and IV (effect of methods course) yields the interesting conjecture, among others, that high achievers attained a more informal view of mathematics but with greater ambivalence! It is further conjectured that high AQ may be an indication of unstable attitudes.

Abstractor's Notes

The research reported is in the mainstream of research on attitudes about mathematics held by various segments of the population. As such, the development of the scales and, in particular, the introduction of a measure of ambivalence of belief provides a new avenue for study of attitudes and attitude change.

The author's hypothesis that high AQ indicated unstable attitudes and potential susceptibility to change in attitudes is an interesting one, although it is indicated that preliminary research does not substantiate it. It may be that AQ in this case is confounded not only by individual-judge disagreement in interpretation but also by the dual nature of mathematics. That is, both formal and informal aspects of mathematics may be equally valued rather than contradictory. The scale items suggest this possibility. Inclusion of the scales in the report is a boon to the reader and to potential user of the research findings.

H. Laverne Thomas
State University of New York
Oneonta
A COMPARISON OF FIRST-GRADE CHILDREN'S ABILITIES ON TWO TYPES OF ARITHMETICAL PRACTICE EXERCISES. Engle, Carol D.; Lerch, Harold H., School Science and Mathematics, v71 n4, pp327-334, Apr 71


Expanded Abstract and Analysis Prepared Especially for I.M.E. by James M. Sherrill, The University of British Columbia

1. Purpose

To ascertain whether first-grade children who had studied the addition facts (and other mathematical ideas) without emphasis on closed number sentences could make correct decisions about basic addition ideas stated as either true or false number sentences. The hypotheses tested were: (1) There will be no difference between the abilities of first-grade children to respond correctly to computational type addition exercises and the abilities of the same first-grade children to respond correctly to closed number sentences describing addition ideas of a comparable level. (2) There will be no difference in boys' abilities to respond correctly to the two types of examples. (3) There will be no difference in girls' abilities to respond correctly to the two types of examples. (4) There will be no difference in the abilities of first-grade boys and first-grade girls in regard to responding to addition examples stated in computational form. (5) There will be no difference in the abilities of first-grade boys and first-grade girls in regard to responding to making decisions about closed number sentences.

2. Rationale

Scott (1966) stated that, "...the expanded use of arithmetic sentences is probably one of the most significant contributions of the arithmetic program renovation." Starr (1969) suggests that children should be using number sentences early in their mathematical studies. Many primary classrooms, however, restrict their use of the number sentences to open number sentences. Dienes (1964), "...infers that closed number sentences, both true and false, should be a part of children's early mathematical experiences...." Although use of the closed number sentence may be a very useful vehicle for the development and practice of basic facts, research findings neither support nor contradict the notion that young children can distinguish between true and false statements of mathematical ideas.
3. Research Design and Procedure

Based on the review of five recently published first-grade textbooks, addition facts were used as the content material through which the hypotheses would be tested.

Considering facts such as $3 + 4 = 7$ and $4 + 3 = 7$ as one combination and $0 + n = n$ and $n + 0 = n$ as one combination, the measuring instrument was developed out of the 46 basic addition combinations. The instrument consisted of two one-page sections which were alternated to avoid any possible influence upon student performance. In one section the facts were presented in the computational form either vertically or as an open sentence, $3 + 4 = \square$. In the other section the facts were presented as closed sentences and the students were to indicate whether the sentence was true or false, e.g., $T \, 3 + 2 = 6$. Each of the 46 combinations, with two exceptions, was randomly assigned to one of the two parts of the instrument. The two exceptions were: (1) a combination involving zero was assigned to each part and (2) one basic combination, randomly selected, was utilized in both parts. Within the computational part, the twelve combinations in the vertical form and the twelve in open sentence form were determined by flip of a coin. The order of the addends was also determined in this manner. The 24 combinations assigned to the closed sentence part of the instrument were each stated twice—once correctly and once where the answer given and the true answer differed by one. A flip of coin determined if the incorrect answers would be given as one more or one less than the correct answer.

The instrument was administered to 130 first-grade students from six classes in two schools. All students received the same instructions. The testing time was 35 minutes and the students were allowed to use any drawings or counting activities they wanted, but objects to count were not provided. When approximately one-half of the testing time had expired, the investigators encouraged individuals to start work on the second part, suggesting that they could return to the first part later. Data from pupils not completing at least one-half of the items on each part of the test (12 subjects out of 142) were not included. Raw scores for each pupil on each part of the test were computed as the ratio of the number of correct responses to the number of items attempted.

Mean scores and variances were then computed for boys and girls on each part of the test. The variances were compared using the F-test, then the t-test for equal or unequal variances was used to compare the means.
4. **Findings**

The five hypotheses tested are stated in the section entitled Purpose. Hypotheses 1 to 3 were not rejected. For hypotheses 4 and 5 a t-score significant at a level between .05 and .10 was attained and the two hypotheses were rejected.

5. **Interpretations**

The first-grade children in this study could respond at about the same level of correctness to closed number sentences as they could to computational type exercises. Performance on both types of exercises appears to be limited only by the children's abilities to recall the basic facts and the experiences they have had constructing or establishing the basic facts through manipulation of objects or through making various kinds of drawings.

Girls are deemed better than boys at responding to arithmetical exercises; special instruction procedures should be initiated which provide for these differences.

The role that closed number sentences can play in the elementary mathematics program is greater than merely serving as a vehicle for practicing basic facts. In their early experiences with closed number sentences, children should learn that number sentences can be either true or false. Activities involving closed sentences should then be an integral part in the development of basic number operations and should be utilized in developing skills in problem solving and in helping pupils learn to make logical decisions.

**Abstractor's Notes**

Engle and Lerch present a careful, well planned description of how the measuring instrument was developed. From the article, however, there is some confusion about the data collected by the instrument. The computational part of the test had 24 combinations to be computed (12 in vertical form, 12 in open sentence form) and the closed sentence part of the test had 48 items to be marked "true" or "false." The "raw scores" for each pupil were computed for each part of the test as a ratio of the number of correct responses to the number of items attempted. The means range from 82.9 to 91.7. These numbers can't be net means of the "raw scores" or the mean of the number of items correct. The numbers may be per cents (though they are not marked as such) or the decimal point may have been misplaced. The confusion is more likely a shortcoming of the article than the study.
The investigators also showed some planning in the statement of their hypotheses and avoided the question of the relative difficulty of items in comparing data from a true and false test with data from a computational test.

One "non-criticism" is that the level of significance was apparently chosen to be .10. The .01 and .05 levels are in such common usage that when one sees something else one expects an explanation.

More of the article could have been spent with the procedures of the study to answer such questions as: How were the subjects selected for the study? and Was the formula for the t computed for hypothesis 1 the formula for correlated means?

More time could have been spent watching the conclusions with the results. The final paragraph of the article should not be read as conclusions from the study, but as suggestions from the investigators. Readers may be in agreement with the statements, but the data collected in the Engle and Lerch study don't justify the statements.

James M. Sherrill
The University of British Columbia
TEACHING METHODS AND INCENTIVES IN RELATION TO JUNIOR HIGH MATHEMATICAL ACHIEVEMENT. Gabor, Georgia M., California Journal of Educational Research, v23 n2, pp56-70, Mar 72


Expanded Abstract and Analysis Prepared Especially for I.M.E. by Larry K. Sowder, Northern Illinois University

1. **Purpose**

   Phase 1: To compare the effects of a "modern" vs. a "traditional" method of teaching on performance on an abstract reasoning test and,

   Phase 2: To investigate the effects of teacher praise, no teacher comment, and teacher reproof on student performance on an abstract reasoning test.

2. **Rationale**

   "Surprisingly little do we know about how to promote optimal learning in real life situations [p. 56]." A research-literature search for information on the disagreement among "discovery" and "memorization" advocates, for example, and a disappointing informal poll of the experimenter's teaching colleagues support this quote. In particular, the generalizability gap from contrived "laboratory" settings to real-life classrooms suggests the desirability of in-classroom research.

3. **Research Design and Procedure**

   The 113 subjects were students in the experimenter's junior-high classes. The students were blacks from low SES environment and had below-average means on California Test of Mental Maturity scores, "Arithmetic Reasoning" stanines, and "Arithmetic Fundamentals" stanines.

   Seven testing instruments of 10 items each were designed by the experimenter. Four forms (L, M, N, X) were used in the treatments of Phase 1; the other three (A, B, D) were used to provide dependent measures and were judged equal in difficulty as the
result of a try-out (on different students). Items on each covered filling blanks in series, items like "--- is to --- as --- is to ---," magic squares, and working with "strange-looking numerals used by some ancient nation." Ten minutes were allowed for each test; maximum score for each was 20.

Phase 1: Three treatments were designed: (1) a "modern" (M) method of teaching ("teacher poses a series of cue-giving questions that lead the Ss to discover and state a concept, rule or principle...." Ss give illustrative examples of positive and negative instances), (2) a "traditional" (T) method (teacher "first presents a rule, formula, or principle and then demonstrates its application by presenting concrete examples"), and (3) a control group (C). M-subjects were the students in a "high" seventh grade section and a "low" eighth grade section; T-subjects, a "low" seventh grade section and a "high" eighth grade section; C-subjects, an "average" seventh grade section.

All groups were given a pre-test (Form A). On four subsequent days (15 minutes/day) M-subjects went over Forms A, L, M, and X in a manner roughly as described for the treatment. On the first treatment day, T-subjects received the answers for Form A and the teacher stated "only the concepts and principles...necessary for correct solutions"; on the three subsequent days T-subjects worked on Forms L, M, and N. Form B was given as a post-test on day 6 to all groups; students were told that this was a regular quiz.

The effects of the three treatments on pre- and post-test scores were tested by ANOVA and post-hoc t tests between group means.

Phase 2: With the same control group, the names of the M- and T-subjects were randomly divided into three groups, Praised (P), Ignored (I), and Reproved (R). Before giving Form D (on day 7), the teacher said, "...I wish to compliment the following Ss for having tried hard and getting good grades on yesterday's test (she then read the P-group names). The following Ss did poorly, having made many careless errors (she then read the R-group names)..." Differences between Form B and Form D scores were tested for each group by t tests.

4. Findings

Phase 1: ANOVA indicated significant differences (p<.01) among the M, T, and C effects on post-test scores; n.s.d. on pre-test scores. Post-test (and pre-test) means were M: 7.9 (3.7), T: 8.6 (3.4), and C: 4.6 (3.2). The t tests between groups indicated
that the M and T post-test means were significantly different (p<.01) from the C mean; M and T means were not statistically different (t84 = -.73).

Phase 2: Post-test means (Form D) and pre-test means (Form B) for the groups were: Praised; 5.9(D), 7.6(B); Ignored: 8.0(D), 8.8(B); Reproved: 9.0(D), 8.4(B); Control: 4.8(D), 4.6(B); all (P+I+R): 7.6(D), 8.3(B). Of the P, I, and R treatments, only the R-group performed better (although not statistically). The only statistically significant difference was the Praised group’s negative post-test – pre-test difference (p<.01).

5. Interpretations

1. The superiority of the M and T groups over the C group "supports Gagne's theory that acquisition of learning sets at successively higher stages of the hierarchy are highly dependent upon prior mastery of subordinate learning sets," the C-group having been given no opportunity to learn how to identify subordinate concepts and principles.

2. The low mean scores (out of 20) even after instruction "implies that the majority might not have as yet reached the cognitive stage of 'formal' operations."

3. The absence of different "modern" vs. "traditional" effects might be explained by the hypothesis that lower ability Ss profit less from meaningful teaching than from drill, or that anxious and compulsive children do better in the "traditional," more structured setting.

4. The negative effect of praise on performance might be explained by noting (a) praised Ss perhaps became overconfident and did not concentrate on the post-test, (b) "perhaps some did not want to be praised in front of their peer [sic] and thus purposely did not try to do well..." and (c) praise may not necessarily be a reward, nor failure a punishment, to everyone.

5. As little as one hour of teaching results in a significant improvement in Ss' performance on abstract reasoning tasks.

Abstractor's Notes

It is encouraging when a classroom teacher becomes interested in research and is willing to design and carry out a study. This positive note, however, should not be taken to imply that such studies should not be soundly based and carefully designed. In the present study, for example:
a. The labels "modern" and "traditional" for the treatments should have been avoided; the "traditional" method—the teacher gave instruction only during one part of the first treatment day; she read answers on the other two days—would likely be endorsed by only a few.

b. Why did the "modern" group work with forms A, L, M, and X while the "traditional" group used forms A, L, M, and N?

c. The lack of random assignment and the doubtful underlying normality of either the scores or the gain-scores make the statistical analyses, and therefore the results of the analyses, questionable. In addition, the unit was not the individual but the class section; the resulting loss of degrees of freedom would likely result in n.s.d., especially for the F-tests.

d. The scoring of an item was 0, 1, or 2. How were part scores assigned? A breakdown of item scores and scores by class section (since they were so different) might be valuable.

e. The Praise-Ignore-Reprove results are interesting, but knowledge of what sort of Praise-Ignore-Reprove treatment the students were accustomed to would be welcome information since the experimenter was the regular teacher.

Larry K. Sowder
Northern Illinois University
1. Purpose

To investigate the existence and the effect on mathematics achievement of individual learning styles categorized as inductive and deductive.

2. Rationale

This was an exploratory study. It attempted to establish the existence of a dichotomization of learning styles based on descriptions of how mathematicians solve problems. According to Hadamard, the two approaches mathematicians use are "intuitive" and "logical." The experimenter translated these into "inductive" and "deductive" when referring to learning styles.

3. Research Design and Procedure

The experimenter constructed all learning materials and evaluative instruments used in the study. The learning materials included:

(a) Learning program on the sum of the measures of the interior angles of a triangle—inductive and deductive versions. (I-T, D-T)

(b) Learning program on the sum of the measures of the interior angles of a quadrilateral—inductive and deductive versions. (I-Q, D-Q)

(c) Learning program on the Pythagorean Theorem and its converse—inductive. (I-P)
(ii) Inductive—Above median on post-test following inductive program and below median post-test following deductive program. (22 Ss)

(iii) High—High—Above median on both post-tests. (62 Ss)

(iv) Low—Low—Below median on both post-tests. (77 Ss)

(v) Not classified—At median on one or two post-tests. (105 Ss)

The 54 Ss classified as deductive and deductive learners completed the I-P and the D-A programs in random order. Each S was given a post-test following the completion of each program. It was expected that the deductive learners would have greater mean scores following the deductive program (D %) than would inductive learners and similarly for inductive learners and the inductive program (I-P).

The data were analyzed for sex differences, for independence of categorization attributes, for dependence of classification upon ability, and for achievement differences on the I-P and D-A programs.

4. Findings

(a) The frequency distributions for both males and females on the classification program post-tests showed a preponderance of high scores. There were more high scores on the triangle post-tests than on the quadrilateral post-tests. Irrespective of inductive or deductive organization, the median triangle post-test score was 11 of 12 while the median quadrilateral post-test score was 10.

(b) A Pearson Chi-square test of association was used to test the hypothesis that the categorical attributes used in the classifications were independent. The test was performed for the classifications resulting from D-I and I-Q post-test scores and from I-T and D-Q post-test scores. For each classification the Chi-square value was statistically significant (p<.001).

(c) A Chi-square test showed that learner ability was not an important factor in the classification scheme. Additionally the distribution of previous mathematics achievement scores for the inductive and deductive learners was not found to be different.
(ii) Inductive—Above median on post-test following inductive program and below median post-test following deductive program. (22 Ss)

(iii) High—High—Above median on both post-tests. (62 Ss)

(iv) Low—Low—Below median on both post-tests. (77 Ss)

(v) Not classified—At median on one or two post-tests. (105 Ss)

The 54 Ss classified as deductive and deductive learners completed the I-P and the D-A programs in random order. Each S was given a post-test following the completion of each program. It was expected that the deductive learners would have greater mean scores following the deductive program (D-A) than would inductive learners and similarly for inductive learners and the inductive program (I-P).

The data were analyzed for sex differences, for independence of categorization attributes, for dependence of classification upon ability, and for achievement differences on the I-P and D-A programs.

4. Findings

(a) The frequency distributions for both males and females on the classification program post-tests showed a preponderance of high scores. There were more high scores on the triangle post-tests than on the quadrilateral post-tests. Irrespective of inductive or deductive organization, the median triangle post-test score was 11 of 12 while the median quadrilateral post-test score was 10.

(b) A Pearson Chi-square test of association was used to test the hypothesis that the categorical attributes used in the classifications were independent. The test was performed for the classifications resulting from D-I and I-Q post-tests scores and from I-T and D-Q post-test scores. For each classification the Chi-square value was statistically significant (p<.001).

(c) A Chi-square test showed that learner ability was not an important factor in the classification scheme. Additionally the distribution of previous mathematics achievement scores for the inductive and deductive learners was not found to be different.
(d) For the 54 Ss classified as inductive or deductive, the test results following the I-P and D-A programs were compared using a two-sample t statistic. There were no statistically significant differences. There was, however, a tendency for the Ss classified as inductive to score higher than those classified as deductive regardless of the nature of the program.

5. Interpretations

(a) Support for the hierarchical structure of the learning sequence used to develop instructional materials was found in the fact that no S had a prerequisite skills test score less than 50 per cent and a pretest score greater than 50 per cent.

(b) The large number of unclassified Ss was due in part to the skewed nature of the distribution and to the five opportunities for a S to be unclassified.

(c) Support for the hypothesized learning styles was found in the Chi-square analyses of the classification data.

(d) The hypothesized learning styles were not supported by the analysis of the achievement data resulting from the I-P and D-A programs. Two factors may have affected this situation: (1) the inductive-deductive categorization might be subject matter specific or (2) the categorization might be imperfect due to the high scores on the classification post-tests.

Abstractor's Notes

The research would have been strengthened if the experimenter had related this work to the psychological literature on "learning styles." Is there an accepted definition of "learning style"? Are there "learning styles" which subsume the hypothesized styles called deductive and inductive? Have "learning styles" been experimentally verified?

As the author noted the classification procedure used may have been dependent upon the learning program and the post-tests. It seems tenuous to classify a person as "inductive" when his test scores differ little from one called "deductive." Modifications should have been made after the pilot work.
Although not noted by the experimenter, the low achievement scores on the I-P and D-A programs were alarming. The means ranged from 26 per cent to 55 per cent correct. If little learning occurs, can one ask the effect of "learning style" on that learning? Technical problems such as these leave one knowing little about "learning styles."

Arthur F. Coxford
University of Michigan
1. Purpose

The purpose of this study was to investigate the relationship between third grade children's performance in solving certain open sentences and three factors (Open Sentence Type, Number Size, and Context) associated with these open sentences.

2. Rationale

Although a distinguishing feature of contemporary mathematics programs is the prominence of open sentences in every grade, very little research has been focused on many questions related to children's ability to arrive at solutions to open sentences. Because of this lack of research evidence many curricular and instructional decisions have been made without the benefit of sufficient research evidence. The aim of this study was to begin to alleviate this situation by answering several questions associated with third graders' performance in solving open sentences.

3. Research Design and Procedure

The design of the study with respect to subjects was a 2 x 2 factorial design with Sex and Order of Presentation making up the dimensions of the design. With respect to sixteen repeated measures of distinct problem situations a 4 x 2 x 2 design—determined by four types of open sentences, two number sizes, and two contexts—was used.

The variable, Open Sentence Type, was limited to four types; namely, I: $N + a = b$, II: $a + N = b$, III: $a - N = b$, and IV: $N - a = b$. The rationale for studying only these four types involved consideration for their mathematical and social importance, their frequency of occurrence in a contemporary program, and
the fact that a method of solution beyond "straight-forward processing of numerals" was necessary.

With respect to another experimental variable of interest, Number Size, the investigator was concerned, for example, about whether $9 + N = 16$ is easier to solve, for third graders, than $39 + N = 86$. And if so, to what extent and to what degree do computational errors play a role in student performance. The two levels of this variable were defined by choosing replacements for both $a$ and $b$, in Sentence Types I-IV, to be elements of: The Basic Facts Domain, $\{(a,b,c)|a, b, and c are counting numbers, a + b = c, a \neq b, 11 \leq c \leq 18, and 1 \leq a, b \leq 9\}$ or The Two Digit Domain, $\{(a,b,c)|a, b, and c are counting numbers, a + b = c, a \neq b, 42 \leq c \leq 99, 21 \leq a, b \leq 78\}$. The two levels of the Context variable (Symbolic and Verbal Symbolic) were defined by problem situations comprised of an open sentence or an open sentence together with a related verbal problem context.

Open sentences were randomly generated by selecting eight triples from the Basic Facts Domain and eight from the Two Digit Domain. Verbal problem contexts were randomly generated by assigning the corresponding number triple to an algorithmic constructed verbal context. Sixteen problem situations were thus randomly generated to reflect the sixteen combinations of the 2, 2, and 4 levels of the Context, Number Size, and Open Sentence Type variables.

The sample for the study consisted of thirty-two third grade children (an equal number of boys and girls) randomly selected from three elementary schools in Madison, Wisconsin. These schools were considered representative of the district.

Each of the subjects were given a sixteen problem situation test, individually administered, and in two parts of equal length. There was a time lapse of sixteen days between the administration of Parts I and II. Each problem situation was scored as 1 or 0 and appropriate statistical procedures were applied to the data.

4. Findings

Statistical tests showed no significant difference (at the .05 level throughout) due to Sex or Order of Presentation. Results due to the variable Open Sentence Type were as follows: performance on Type I open sentences was not statistically different from performance on Types II, III, and IV; performance on Type II open sentences was statistically different (and better) from performance on Types III and IV; performance on Type III open sentences was different (and better) from performance on Type IV. Performance on
open sentences with a and b replaced from the Basic Facts Domain was better than with a and b replaced from the Two Digit Domain. There were no significant main effects due to the Context variable. A significant Open Sentence Type X Number Size interaction was obtained.

5. Interpretations

It was concluded that boys' and girls' performances are very similar in solving certain mathematical sentences. The author suggests that if it is expected that third grade children should be able to solve the four types of open sentences equally well, then more instructional attention must be given to open sentences of the form \( N - a = b \). And, if an instructional objective for the third grade is that students be able to solve open sentences from the two-digit domain, then additional instructional activity must take place to accomplish this. Suggestions for further research include extension to open sentences involving integers and other number domains and to sentences which involve division and multiplication.

Abstractor's Notes

This reviewer agrees with the premise made in the paper that we have in the past made curricular decisions without sufficient research information; moreover, he applauds the care used in this study which seeks to answer a well-defined and significant problem related to a curriculum decision. As this study makes very evident, the number of research studies which will be necessary to give information on which major curriculum decisions can be based will be great, probably exponentially large in comparison to the number of major decisions.

Also it is this reviewer's opinion that, because many curriculum developers are not interpreters of educational research, researchers in mathematics education must provide occasional critical reviews, evaluations, summarizations and generalizations to transform "curriculum oriented" research reports into "consumer" expositions.

Merlyn J. Behr
Northern Illinois University
AN INVESTIGATION OF THE EFFECTS OF NON-DECIMAL NUMERATION INSTRUCTION ON MATHEMATICAL UNDERSTANDING. Higgins, James E. School Science and Mathematics, v72 n4, pp293-297, Apr 72


Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jack E. Forbes, Purdue University, Calumet Campus

1. Purpose
To contribute to the knowledge as to whether or not non-decimal numeration instruction facilitates pupil understanding of place value, order, reading numerals and writing numerals.

2. Rationale
The author states: "Number bases, other than base ten, have been thought by many to be a means for increasing pupil understanding of base ten. However, opinions have differed and research findings have been inconclusive."

3. Research Design and Procedure
A sample of fifteen fifth grade classes was randomly selected from seventy such classes in one city. These classes were randomly assigned to one of three groups.

Treatment A: Seven lessons of instruction on base ten numeration.
Treatment B: Seven exactly comparable lessons on base five numeration.
Treatment C: Control group; no numeration instruction.

In A and B, no more than three lessons per week were taught. Teachers of these classes received special training and materials.

The first twenty-one items of A Test of Arithmetic Principles, Elementary Form were used as a posttest to measure understanding of the four aspects of numeration listed under "Purpose" in this abstract. An analysis of variance was used to test the null hypothesis that "no significant difference in understanding of these four aspects of numeration exists as a result of the three treatments."
4. Findings

While class means for treatment B were somewhat higher than those for treatment A and both exceeded those for treatment C, these differences were not statistically significant (p<.05).

5. Interpretations

The author notes the limitations of these results due to sample size, geographical distribution, and the grade level at which the study was conducted. He also notes he was studying the effect of the instruction only on understanding numeration. He suggests that such instruction may have other values even if more extensive research were to remove the limitations of this study and obtain replication of its results.

Abstractor's Notes

The fact that neither treatment A nor treatment B produced statistically significant gains over no instruction in numeration leads one to conclude that some or all of the limitations noted by the author were important ones. Of course, the size of the sample alone makes statistical significance very difficult to achieve. This study is, to say the least, not definitive.

Jack E. Forbes
Purdue University
Calumet Campus
1. Purpose

The study attempted to help answer three questions:

1. To what extent does a mathematics laboratory enable slow learners to gain in achievement?

2. To what extent does a mathematics laboratory enable gifted learners to gain in mathematics achievement?

3. To what extent does a mathematics laboratory enable slow learners and gifted learners to develop a more positive attitude toward mathematics?

2. Rationale

Although there is no detailed review of theoretical and empirical literature as the basis for a rationale for the study, there were three motivating ideas. Non-average students were seen as presenting special instructional problems for an over-worked elementary teacher. The mathematics laboratory was seen as a prime innovation for assisting in the development of computational skills, concepts and attitudes. Because of parallels to assumed successful language laboratories, mathematics laboratories were seen as a possible instructional remedy for the above problem. The investigator perceived a lack of empirical evidence at the time the study was carried out (1969-70), hence a study such as this was needed.

3. Research Design and Procedure

The study was carried out in grades 4, 5, and 6 in two schools in a "middle class" (School A) and a "low socio-economic" (School 44

44
B) neighborhood in suburban Houston, Texas. Students in these two rather large elementary schools were screened on the basis of standardized tests in order to identify those either more than one grade level above or more than one half grade level below each of the three grade levels. On the basis of this screening 236 students in school A and 164 in B were then pre-tested on the California Arithmetic Test (Elementary or Upper Primary, Form W). In school A the 20 highest and 30 lowest at each grade level were randomly assigned to experimental and control groups. Due to population limitations, in school B such an assignment was only carried out for the grade 5 students and the slow learning grade 4 students. A complete experimental compliment was drawn at the grades 4 and 6 level, but on a different screening basis.

The laboratory in each school was housed in a special room. Although the lab program was not detailed, students in the lab had access to special facilities: listening stations, film strips, games, devices, and self-instructional material. One regular teacher or one of two doctoral students was always present in each of the labs.

The students grouped by pre-test diagnosis spent 2 forty-five minute periods per week in the lab. The labs operated four days a week with the fifth day given over to planning, material design and evaluation of lab materials. The visits continued from September to May in place of normal arithmetic periods.

Four null hypotheses directly related to the purpose of the study were stated.

- There is no mean difference between sample students in Schools A and B on
  i) IQ
  ii) Pre-tests.

- There is no difference in gains, in each school, between experimental and control groups with respect to
  i) Achievement
  ii) Attitude.

To test the achievement hypotheses parallel forms of the California Arithmetic Tests were administered in September 1969 and May 1970. At the same times a revised Aiken-Dreger 20 Item Attitude Inventory with suitable reliability was administered. IQ data was gathered from records of given Otis Beta tests. All screened students were told they were being considered for admission to the lab
at the time of pre-testing ostensibly to control for the Hawthorne effect.

The hypotheses were tested statistically using analysis of covariance for the first pair given above, and analysis of variance with repeated measure for the second pair.

4. Findings

There were no significant pre-experiment mean differences on the mathematics achievement or IQ scores between schools at each grade level. This result was also observed when comparing experimental and control classes at each grade level.

With respect to the total laboratory sample and the sub-samples of slow and gifted learners, there was a significant gain from the pre-test in September to the post-test in May in mathematics achievement. This gain was observed in both schools. Though there was some evidence in the data favoring the laboratory treatment, there were no significant differences at any grade level between laboratory and control groups with respect to achievement gain.

Considering the total sample in each school there was a significant gain in attitude in school B, but not in school A. While in school A the experimental group mean attitude rose slightly and the control group mean attitude declined slightly, there was no significant difference between groups at either school or at any grade level with respect to attitudes.

5. Interpretations

Although noting no differences between experimental and control groups, the author found a slight differential effect in achievement favoring the laboratory groups for both slow learners and gifted learners especially at the lower grade levels. The laboratory did allow a significant gain in achievement at both schools and in attitude in the lower socio-economic school. The author concludes that laboratories organized to provide personal and individualized assistance are most helpful to learners that are either academically or culturally disadvantaged.

Abstractor's Notes

1. This study is noteworthy because of its length and because it covered the upper elementary school. The sampling process was careful.
2. The data indicates that the lab apparently benefited slow learners in the "middle class" school actually and differentially more than slow learners in the "low socio-economic" school.

3. The lab defined in this study appeared to be an enriched and more individual learning environment. There are numerous studies now reported which suggest that with slow learners it is the type of physical experience and the mathematical content which have an effect. There was no planned quality of laboratory experience identified in this report with respect to either of these variables.

4. The use of standardized tests may have obscured some of the effects of the laboratory treatment. Since the study was so comprehensive, it is unfortunate that a wider sampling of objectives was not tested in the criterion measures.

Thomas E. Kieren
University of Alberta
1. Purpose

(a) To compare two instruments designed to measure attitudes toward mathematics: a Likert-type attitude scale (Mathematics Attitude Scale) constructed by Aiken and Dreger, and a semantic differential constructed by McCallon and Brown.

(b) To determine whether the evaluative scales of the semantic differential instrument differentiate between high and low scorers on the Aiken-Dreger scale.

2. Rationale

As the abstractor understands the rationale of this study, it is intimated that although the Mathematics Attitude Scale is an older instrument and the Likert procedure an older technique of measuring attitudes, they are not necessarily superior to the semantic differential procedure and an attitude instrument employing this procedure.

3. Research Design and Procedure

The 20-item Mathematics Attitude Scale (MAS) and a semantic differential consisting of the concept "Mathematics" and 15 bipolar adjectival scales (7 points per scale) were administered to 68 male and female students in a graduate course in educational statistics. Half the students took the MAS first and the semantic differential one week later; the remaining students took the two attitude instruments in the reverse order. The responses to the semantic differential were then factor analyzed and the factors rotated by a varimax procedure. Finally, the raw scores and the scores on the first factor of the semantic differential were correlated with each other and with scores on the MAS.
4. Findings

Of the two factors extracted, Factor I was labeled "Evaluation," and Factor II "Potency." The correlation between raw scores on the semantic differential and scores on the MAS was .90, while scores on the "Evaluation" factor of the semantic differential correlated .97 with raw scores on that instrument and .90 with scores on the MAS. Finally, subjects were divided into "favorable" and "unfavorable" attitude groups according to their scores on the MAS and the mean scores of these two groups on the 15 semantic differential adjectival scales compared.

5. Interpretations

The investigators concluded that the semantic differential "proved to be as effective a measure of attitude toward mathematics as the MAS." In addition, due to its ease of construction, it was concluded that the semantic differential is a more "practical" approach to the measurement of attitudes toward mathematics. Finally, it was stated that the fact that the means of the "evaluative" (Factor I) scales of the semantic differential were higher for high scorers than for low scorers on the MAS attests to the "external" validity of the semantic differential.

Abstractor's Notes

Actually, the first stated purpose of this investigation was not realized. Simply because scores on two methods of measuring the same variable, administered one week apart, correlate highly, it does not tell us anything about the relative utility of either method. In fact, a single item on the MAS would undoubtedly correlate substantially with scores on the semantic differential instrument. If the investigators had been serious about comparing the reliabilities of the two methods, they would have reported separate test-retest coefficients over a period of several months, or at least separate alpha coefficients. And if they had been serious about the question of differential validity, they would have used a multiple regression approach to predicting statistics grades, or some other criterion, from scores on the two instruments in combination with ability test scores.

The results of the factor analysis of the semantic differential showed that one factor ("evaluation") absorbed most of the total score variance. Therefore, it is not surprising that scores on this factor should correlate highly with the same variable—the MAS—as raw scores on the unfactored semantic differential. Also,
it is clearly predictable from this finding that high scorers on the MAS will make higher mean scores on this "evaluation" factor than low scorers on the MAS.

Furthermore, considering the nature of the sample employed, it is not surprising that one factor should be so important, but the abstractor is not convinced that "evaluation" is the best name for it. From an inspection of the Factor I loadings, it appears that this dimension might better be labeled "like vs. dislike mathematics." In a more recent effort to analyze attitudes toward mathematics, the abstractor has devised an "evaluation" or "value" scale concerned with the recognized importance of value of mathematics. Scores on this scale have been compared with scores on an abbreviated version of the MAS, and the two scales, although positively correlated, appear to measure somewhat different variables.

Finally, the question of the relative efficiency of the Likert attitude-scaling approach with that used in devising the semantic differential requires little comment. The MAS was already available, but the investigators had to construct the semantic differential instrument. So which instrument was more "practical" or efficient for them to use initially? In terms of ease of administration, I have found that students frequently experience difficulty in rating school subjects such as mathematics as "fast-slow," "strong-weak," "hard-soft," "light-heavy," etc. They have expressed no such problem in responding to the MAS or other Likert-type scales.

Lewis R. Aiken
Guilford College
1. Purpose

To compare the mathematical competencies of the following three groups of children entering school: (1) Negro children enrolled in all Negro kindergartens in an all Negro community; (2) Negro children enrolled in integrated, suburban kindergartens; and (3) non-Negro children enrolled in the same set of integrated kindergartens.

2. Rationale

Studies of the relationships of measures of achievement and intelligence to various racial aspects have created considerable controversy during the past several years. Rea and Reys, the authors of this study, cited several recent investigations on this subject including the much publicized work of Jensen related to IQ and scholastic achievement.

In addition, the need for instruments to provide base line data on entering kindergarteners had been cited in the literature and had motivated the investigators to develop the Comprehensive Mathematical Inventory (CMI). This instrument, which is administered individually to children, consists of two parts. The first contains 200 items organized into seven subscales: (1) money, (2) number, (3) vocabulary, (4) geometry, (5) measure, (6) pattern identification, and (7) recall. Each of the items is developed by pictures and manipulative objects appropriate to young children. The second part consists of two open-ended questions: (1) "Can you count for me?" and (2) "Can you count these chips for me?", named rote and rational counting, respectively. Measures of reliability for the instrument were determined by use of the Kuder-Richardson Formula 20 and varied from .91 to .94 for Part I and from .83 to .87 for Part II.
3. Research Design and Procedure

The CMI was administered to a total of 727 children from 30 kindergartens in the St. Louis, Missouri metropolitan area. The children consisted of three samples: (1) 93 Negro children attending an all Negro school in an all Negro community (AN); (2) 27 Negro children attending integrated, suburban schools (N); and (3) 607 non-Negro children attending the same integrated schools (NN). The data were collected during the second full week of classes in the 1968-69 school year.

The measurements on each of the variables (rote counting, rational counting, money, number, vocabulary, geometry, pattern identification, measurement, recall, and total score) were studied through a one-way analysis of variance followed by a use of Scheffe's Multiple Comparison Test to isolate significant differences among the sets of three sample means for each variable.

4. Findings

Highly consistent results were found for each of the variables. In every case, the mean of the AN sample was significantly less than the means of the NN and the N samples, respectively; but no significant differences were found between the means of the NN and the N samples for any variable (p<.01).

5. Interpretations

The authors concluded that the results of this research revealed that: "Kindergarteners from an all Negro community were significantly below the mathematics achievement of both Negro and non-Negro kindergarteners from surrounding municipalities," and "There was no significant difference in mathematics achievement between Negro and non-Negro kindergarteners drawn from integrated suburbs." However, they concluded from analyses of the same data reported elsewhere (but not in this article) that: "This analysis has produced results in line with considerable other evidence which indicate that race is a nonsignificant variable in investigating mental ability or academic achievement."

Abstractor's Notes

This investigation delves into an interesting research area; but as is the case with many studies, it raises numerous questions.
(1) It is not clear from the report whether the samples which were studies were selected randomly from the defined populations or were selected in some other manner.

(2) Since the acceptance of the equality of means of the NN and the N populations for each variable is a crucial finding of the study, one may be interested in learning the size of the Type II error (which is probably quite small).

(3) Considerable interest centers on the 27 Negro children attending the integrated schools. For example, were any of them bused into the schools from outside areas? What were the educational backgrounds and the socio-economic statuses of the parents of these 27 children as compared to the 607 non-Negro children? If one were to match the 27 Negro children with 27 non-Negro children on the bases of variables related to the children and their parents, what differences, if any, might be found between the two matched samples? Similarly, if a matching was done with the 93 Negro children in the all Negro school and 93 non-Negro children, what differences might be found?

(4) A most significant finding is that "...race is a non-significant variable in investigating mental ability or academic achievement," but this result is not a consequence of this study but rather the finding of an analysis "reported elsewhere."

On the whole, one may conclude that this study is certainly interesting but has not produced results which have contributed substantially to a solution of the nature-nurture controversy.

Donald J. Dessart  
The University of Tennessee  
Knoxville
1. **Purpose**

The purpose of the study was to determine whether the Survey of Study Habits and Attitudes (SSHA) was a valid predictor of academic accomplishments of sixth graders in reading and arithmetic.

2. **Rationale**

Researchers have used the SSHA to measure study habits and attitudes in grades 7-12. It is important to identify early problems in these areas and sixth graders were judged not to be too dissimilar to seventh graders, so it was decided to use the SSHA at the sixth grade level.

3. **Research Design and Procedure**

The subjects were 16 boys and 10 girls in the sixth grade of a private school. They were given the SSHA and scores were available from the Metropolitan Reading Achievement Test and the Iowa Test of Basic Arithmetical Skills.

4. **Findings**

Rank order correlations were obtained for the achievement tests with all parts of the SSHA (habits, attitudes, total score). Only the Attitudes subtest of the SSHA was significantly correlated (p < .05) with Reading. None of the scales of the SSHA were correlated significantly with the Arithmetic test (p < .05). Subsequent correlations done separately for boys and girls produced significant correlations between Attitudes and Reading (p < .01) and between Total SSHA score and Reading (p < .05) for boys. For girls, the Attitudes subtest was significantly correlated (p < .01) with Arithmetic scores. The average correlation between subtests of the SSHA was .75 (p < .01).
5. **Interpretations**

The authors concluded that "the SSHA has some utility, at least as a research instrument, in the elementary school." They also concluded that scores of males and females as well as relationships between SSHA and academic achievement scores should be analyzed and reported separately. Finally, the authors stated that this research "should pave the way for the use of the SSHA in the elementary school for the early diagnosis of individual attitudinal problems hindering effective school performance."

**Abstractor's Notes**

This study does not validate the SSHA as a predictive instrument of academic achievement in the sixth grade. A significant correlation coefficient between variables cannot be interpreted to mean that one variable predicts the other. The small number of subjects, the special population and the use of rank order correlations further detract from the value of the study as a validation of the SSHA.

The reported correlations between attitude and achievement add nothing new to present knowledge. The recommendation that predictive measures be related to specific academic areas is a good one but there are probably better attitude tests for use with arithmetic achievement of sixth graders than the SSHA, since the proposed predictive measures should also be area-related.

The SSHA was originally developed specifically for college-bound students. Although it is advertised for grades 7-12, its use in sixth grade may not be appropriate. The suitability of the reading level, reliability data, and the relevance of item content to the subjects' school experiences are not reported. The last of these may have accounted for the absence of significant correlations between study habits and achievement.

In summary, the study does not add to present knowledge about predicting achievement and provides little evidence that the SSHA is valid at the sixth grade level.

Gerald Kulm
Purdue University
1. **Purpose**

The purpose as stated was to analyze word problems found in six commercially produced grade six arithmetic texts to determine readability levels and compare results with a similar analysis of the problem solving sections of three standardized arithmetic achievement tests.

2. **Rationale**

The author argued that the steps involved in finding solutions to word problems were: reading the statement, analyzing the data, and using computational skills to arrive at a correct solution. Of all these steps, ability to read was considered most important. "If a child cannot read the statement, then his abilities to think, analyze and compute are hampered." Several related studies were cited.

3. **Research Design and Procedure**

Word problems found at the end of each chapter of six commercially produced grade six texts were selected for study. The Dale-Chall formula was used to determine readability levels. Each problem was analyzed separately in order to ascertain the range of reading levels over the entire text and then raw scores were combined to determine a single overall reading level. Problem sections designed for superior or gifted students were excluded from the study. Problems from the three achievement tests were similarly treated.

4. **Findings**

Overall readability scores for all six texts fell between fifth and sixth grade level. Two texts were close to grade five...
level, two were close to grade six level and two fell in the middle of grade five. The range of readability levels for individual problems across all six texts varied from grade four and below to grades seven and eight. Two texts ranged from fourth and below to grades five and six, two ranged from fourth and below to grades seven and eight, and the final two ranged from grades five and six to seven and eight.

A similar analysis of the three achievement tests revealed that one of the tests had a readability level of grade four and below, while the others were at a fifth or sixth grade level.

5. Interpretations

The author noted that, while the overall readability of all six texts appeared to be at an appropriate grade level, the readability levels for single problems varied considerably within a given text and across all six. Clearly, it was noted, that while some problems within a single text could be found easily readable by given students, the same students might well find other problems beyond their own reading competence. The range of readability levels also indicated that two of the texts could well be considered easy to read, and two others could prove difficult. Analysis showed further that frequently problems with lower readability levels did not appear early in the texts. Generally the author concluded that, "There was no consistency among the samples within a text."

The grade level readability scores for the three achievement tests showed them to be either at a level similar to or slightly below the level of the textbooks.

Abstractor's Notes

1. This study appears to have followed the general lines of similar readability studies.

2. The procedure used to expose the actual range of readability levels within a text so often masked by a composite or overall level should be most helpful to teachers. It can not be assumed that a composite readability level which falls within the measured reading competence of a student assures that reading may not still be a significant barrier to the development of problem solving skills within a given arithmetic text.
3. The author presented a reasonable argument for the importance of readability to problem solving. He also cited four related studies, but gave not a hint of their findings nor of how they might have been related to the present study other than the fact that all utilized some readability formula. Thus, he failed, in this writer's judgment, to present an adequate theoretical or conceptual framework from which the need and value of the present study would become clearly evident.

Roland F. Gray
University of British Columbia
A STUDY OF THE LABORATORY APPROACH AND GUIDED DISCOVERY IN THE TEACHING LEARNING OF MATHEMATICS BY CHILDREN AND PROSPECTIVE TEACHERS. Unkel, Esther R., Florida Atlantic University, Boca Raton. Spons Agency--National Center for Educational Research and Development (DHEW/CE), Washington, D. C. Pub Date Sep 71, Note--15 p, EDRS Price MF-$0.65 HC-$3.29


Expanded Abstract and Analysis Prepared Especially for I.M.E. by Jenny R. Armstrong, University of Wisconsin

1. Purpose

   The purpose of this study was to determine whether instruction based on guided discovery, the use of inexpensive manipulative aids and maximum pupil participation would increase achievement in mathematics for pupils in grades 1-6 and to determine whether a program based largely on learning through activity and guided discovery is effective in the training of prospective teachers enrolled in a mathematics methods course for the elementary teacher.

2. Rationale

   Despite the increased amounts of money which have been spent in education during the last ten years, the achievement of many elementary level pupils in mathematics is often below that of their potential. This study assumes that a possible cause for this pupil underachievement in mathematics is related to the training of teachers in mathematics methods courses. This study, therefore, examines a teacher training methodology based largely on student-teacher interaction with pupils in an activity-oriented guided discovery approach. The criterion of success for this method was the learning of the pupils with whom the students worked in mathematics and the increase in knowledge exhibited in mathematics by student teachers.

3. Research Design and Procedure

   The sample of the study included 72 undergraduates in elementary education and 29 elementary school pupils. All undergraduates in elementary education at Florida Atlantic University who were
taking the mathematics methods course during the fall and spring quarter were included in the study. The pupils were chosen from one school in Boca Raton, Florida within walking distance of the mathematics laboratory on the campus of Florida Atlantic University. The pupils were selected by classroom teachers as being underachievers in mathematics on the basis of standardized tests and observations. Three pupils were not recommended by teachers, but came on their own initiative.

The pupils were tutored twice weekly by the undergraduates enrolled in the methods class. Twenty-nine of the 66 pupils in grades 1-6 completed the full 9-month period. There were 4 in grade one, 13 in grade two, 7 in grade three, 4 in grade four, 5 in grade five, and 6 in grade six.¹ Those not included in the test sample missed at least one quarter of instruction because they could not attend at a time when it was possible for an undergraduate to tutor them.

The Glennon Basic Test of Mathematical Knowledge was administered at the beginning of the quarter to all undergraduates. At the end of the quarter an adaptation of the Glennon test covering the same number concepts was administered. During the quarter, a weekly seminar was held and attended by each undergraduate immediately following one of the two tutoring periods to discuss the work, ask questions and plan future work for the pupils.

During the weekly seminar the regular classroom chairs were removed and tables and straight chairs were used. Six students worked at each table with the instructor moving about among the tables to discuss the work with the students. Lectures were limited to less than an hour of the three-hour class period and ideas were crystallized only after the students had an opportunity to work with, discuss and think about the content.

Both undergraduates and pupils were taught using a "Task Card" approach to guided discovery. Instruction was given via Task Cards which directed the performance of a specific task and usually required working with a peer. For example, the individual might be directed to fill boxes of various sizes with one inch cubes, keeping a record of the length, width, and height of the container and the number of blocks required to fill each container. Before

¹There is an unexplained discrepancy in the numbers here. The sum of this breakdown as shown in Table 1 in the original report is 39. The pupil sample number given elsewhere throughout the report is 29. I have no way of knowing which is correct. (JRA)
filling the box with one inch cubes and counting the number it held, the task card directed the pupil to estimate the number of cubes the various containers would hold. The eventual discovery was that width times height times length equaled volume.

When the undergraduates implemented the "Task Card" approach to guided discovery with the pupils it was at the end of the pupils' school day with a maximum of four pupils in each tutorial session. The pupils were given two tests during the second week in October, the California Mental Maturity Test and the California Achievement Test, Form X, 1966 Norms, and one test during the third week in May, Form W of the California Achievement Test. The tests were administered either individually or in small groups of two to six pupils. The arithmetic test data were recorded separately for the following: arithmetic reasoning, arithmetic fundamentals, and arithmetic totals. Multiple t statistics (seventeen in all) were computed for each possible comparison. Ten of the seventeen were found to be significant at the .01 level, one at the .001 level, and two at the .05 level.

4. Findings

The mean error rate for the undergraduates as assessed by the Glennon Test of Mathematical Knowledge and a study developed parallel form of the Glennon Instrument decreased for those undergraduates who participated in the Task Card Approach to Guided Discovery courses. When analyzed separately by grades and subtests (reasoning, fundamentals, and totals), mean pupil achievement in arithmetic changed in all cases except in reasoning at grade three, reasoning, fundamentals and totals in grade four, and reasoning in grade five.

5. Interpretations

Guided discovery through the use of task cards and manipulative aids shows promise as a method of instruction both for undergraduates in teacher training programs and for pupils in elementary grades 1-6.

Abstractor's Notes

The guided discovery approach using task cards and manipulative aids was well supported from a theoretical and reported research base as a potential method for teaching undergraduates and pupils in the elementary school in general, but no rationale was
provided for the use of this type of methodology to enhance the mathematical learning of underachievers. This study does nothing to verify this method's value on an empirical basis. The experimental design of the investigation was not detailed and one suspects not considered. There were no design controls built into the study to eliminate alternative hypotheses of explanation for the changes in behavior found for both undergraduates and pupils. These changes may have been explained by other course instruction, maturation, outside activities and/or measurement or statistical error. The validity and reliability of the measurement instruments were not reported. Certainly, the content validity of the chosen instrument for this particular study is highly questionable. If the chosen instrumentation was specifically chosen to assess the content relayed by the experimental procedure, it is not verified in the report. A more detailed description or listing of the content included on the task cards would provide at least part of the necessary information to properly evaluate this measurement question. The choice of multiple t tests to statistically analyze the data offers no systematic control for error rate. Whether correlation and, thus, non-independence of the pre- and posttest scores was taken into account in the statistical analysis is not clear from the report. The conclusions made were not strong; they could not be on the basis of the work undertaken. One questions, however, whether these same conclusions could have been made without ever undertaking the study.

Jenny R. Armstrong
University of Wisconsin
1. **Purpose**

To create a geometry course using transformations in the framework of standard courses, including a systematic development of basic theorems of Euclidean geometry from postulates. To determine the effects of such a course upon student attitude and achievement, and determine whether teachers can teach such a course without a large amount of retraining.

2. **Rationale**

Some mathematics educators have suggested that the study of geometric transformation be one of the major goals of tenth-grade Euclidean geometry. Several texts in which transformations are given substantial attention have appeared recently. There has been little, if any, research that sheds light on the possible effects of studying transformations on student achievement or attitudes.

3. **Research Design and Procedure**

Preliminary materials were written by Coxford and Usiskin in 1967-68, tried with three classes at the University of Michigan High School and revised for the study. The research was conducted during the 1968-69 school year.

The classes of eight volunteer teachers in six schools were chosen as the experimental group (E). The experimenter and mathematics supervisors chose the classes of nine teachers in seven schools which they believed matched the E schools in size, location, and type. E and C teachers were similar in number of years experience in high school or college mathematics and in teaching high school geometry. The E teachers averaged 2.9 college geometry
courses and C teachers averaged 1.8. The mean E class size was 22.9 and the mean C class size was 26.8. There were 18 control classes which used five different geometry texts, none of which used transformations.

The ETS Cooperative Tests in Geometry were used to test achievement. Part I was given in the first week of school, Parts I and II were given in June. Each class was randomly divided in halves, with half taking Form A of Part I, and half taking Form B of Part I in September. In June, each student was given the form he had not used in September. The Aiken-Dreger Opinionnaire was administered four times (September, December, March, and June) to measure student attitudes.

4. Findings

Using analysis of covariance to adjust for September pre-test scores, the control group (C) had significantly higher post-test achievement scores (p<.01) than the experimental group (E). When the samples were divided according to sex, and into quartiles according to the September tests, similar but not always significant results obtained.

Mean attitude scores were lower in June than in September for both E and C students, but the difference was significant (p<.01) only for the C students. Scores of E students did not fall as far as those of C students, but the difference in change was not significant.

Teachers of experimental sections had strongly positive attitudes towards the material as indicated by a nine item multiple-choice questionnaire and the fact that each school elected to buy additional texts the following year. In-service help was offered to the experimental and control teachers through special sessions or by telephone. The special sessions were badly attended and discontinued by agreement of the teachers. The telephone assistance was not used.

5. Interpretations

The researcher believes the study shows that there is insufficient time to teach the experimental approach and develop the competencies achieved in the standard program. But, he says the study shows that it is possible to develop and teach a transformation geometry course to average students without special training for the teachers. He raises the (non-experimental) question of
whether it is worth developing such a course and responds that because of the usefulness of transformations in other mathematics courses he believes the answer is yes.

While the experimenter called attention to the possibility that repeating the attitude measure four times in a year may have adversely influenced the attitude scores, he expressed great disappointment at the drop in these scores and found scant satisfaction in the fact that the situation was worse in the control group than in the experimental group.

Abstractor's Notes

The initial conditions of this research appear to favor the experimental group. Experimental teachers were volunteers for a special project and averaged almost one extra year of geometry in college. The extra knowledge gained in such a course is probably less important than the indication of their greater interest in geometry. Furthermore, the control classes were, on the average, 17 per cent larger than the experimental classes.

Apparently the reported N's were for those students who actually took all the attitude questionnaires or both achievement tests. The great difference between the original N's and the reported N's was not explained. For the control group there were 18 classes averaging 26.8 pupils per class, for a total of 482 or 483 pupils but for the attitude measure, N=294 and for achievement N=335. Presumably, there were 16 experimental classes which are reported to have averaged 22.9 pupils per class for a total of 366 or 367 pupils but for the attitude measure, N=246 and for achievement, N=324. If the assumption regarding the 16 classes is correct, that means that 6 per cent more of the experimental students completed the attitude instrument all four times and 19 per cent more completed the achievement tests. These facts seem sufficiently unusual to warrant some comment or explanation from the investigator.

This research shows what many similar studies in the past have shown, namely that there is a great variety of things that can be done in a high school geometry course without greatly affecting achievement of usual concepts. Thus, apparently the question of what ought to be taught remains essentially a question of taste and philosophy.

Stephen S. Willoughby
New York University
**IMPORTANT INSTRUCTIONS**

- Order ERIC Reports only by 6 digit ED No. shown in Research in Education (RIE) or other indices.
- Indicate if you want microfiche film (M/F) or paper copies (PC).
- Enter unit prices from the Price List below. All other prices are out of date.
- Enclose check or money order payable to EDRS for orders totalling less than $10.00.

**PRICE LIST**

<table>
<thead>
<tr>
<th>MICROFICHE (M/F)</th>
<th>PAPER COPIES (PC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each ERIC Report - $0.65</td>
<td>Number of Pages per ERIC Report:</td>
</tr>
<tr>
<td>Microfiche Film (M/F) is a 4&quot; x 6&quot; sheet of microfilm on which up to 70 pages of text are reproduced.</td>
<td>1 - 100 $3.29</td>
</tr>
<tr>
<td></td>
<td>101 - 200 6.58</td>
</tr>
<tr>
<td></td>
<td>201 - 300 9.87</td>
</tr>
<tr>
<td>Each additional 100 pages or portion thereof - $3.29</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:**

1. Fourth Class Book Rate or Library Rate postage is included in above prices.
2. The difference between Book Rate or Library Rate and first class or foreign postage (outside the continental United States) rate will be billed at cost.
3. Paper copies (PC), shown as hard copy (HC) in past RIE issues, will be stapled with heavy paper covers.

**SIGNATURE** ____________________________  **DATE** ____________

**TITLE OR DEPT.** ____________________________

**MAKE ALL DRAFTS PAYABLE TO EDRS**

Orders are filled only from ED accession numbers. Titles are not checked. Please be sure you have supplied the correct numbers.

---

**ERIC DOCUMENT REPRODUCTION SERVICE** is operated by Leasco Information Products, Inc. for the U.S. Office of Education.
TERMS AND CONDITIONS

1. PRICE LIST
   The prices set forth herein may be changed without notice; however, any price change will be subject to the approval of the U.S. Office of Education Contracting Officer.

2. PAYMENT
   The prices set forth herein do not include any sales, use, excise, or similar taxes which may apply to the sale of microfiche or hard copy to the Customer. The cost of such taxes, if any, shall be borne by the Customer.
   Payment shall be made net thirty (30) days from date of invoice. Payment shall be without expense to LIPCO.

3. REPRODUCTION
   Materials supplied hereunder may only be reproduced for not-for-profit educational institutions and organizations; provided however, that express permission to reproduce a copyrighted document provided hereunder must be obtained in writing from the copyright holder noted on the title page of such copyrighted document.

4. CONTINGENCIES
   LIPCO shall not be liable to Customer or any other person for any failure or delay in the performance of any obligation if such failure of delay (a) is due to events beyond the control of LIPCO including, but not limited to, fire, storm, flood, earthquake, explosion, accident, acts of the public enemy, strikes, lockouts, labor disputes, labor shortage, work stoppages, transportation embargoes or delays, failure or shortage of materials, supplies or machinery, acts of God, or acts or regulations or priorities of the federal, state, or local governments; (b) is due to failures of performance of subcontractors beyond LIPCO's control and without negligence on the part of LIPCO; or (c) is due to erroneous or incomplete information furnished by Customer.

5. LIABILITY
   LIPCO's liability, if any, arising hereunder shall not exceed restitution of charges.
   In no event shall LIPCO be liable for special, consequential, or liquidated damages arising from the provision of services hereunder.

6. WARRANTY
   LIPCO MAKES NO WARRANTY, EXPRESS OR IMPLIED, AS TO ANY MATTER WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE.

7. QUALITY
   LIPCO will replace products returned because of reproduction defects or incompleteness. The quality of the input document is not the responsibility of LIPCO. Best available copy will be supplied.

8. CHANGES
   No waiver, alteration, or modification of any of the provisions hereof shall be binding unless in writing and signed by an officer of LIPCO.

9. DEFAULT AND WAIVER
   a. If Customer fails with respect to this or any other agreement with LIPCO to pay any invoice when due or to accept any shipment as ordered, LIPCO may without prejudice to other remedies defer any further shipments until the default is corrected, or cancel this Purchase Order
   b. No course of conduct nor any delay of LIPCO in exercising any right hereunder shall waive any rights of LIPCO or modify this Agreement.

10. GOVERNING LAW
    This Agreement shall be construed to be between merchants. Any question concerning its validity, construction, or performance shall be governed by the laws of the State of New York.