ABSTRACT

In order to enlarge the sample of educable mentally retarded children used in the evaluation of media at the Computer Based Project, a number of classes were transported by bus to the project site. The investigator questioned whether observed differences in manifest anxiety might affect responses to the media being evaluated. A study was made to ascertain whether or not there were differences between transported and nontransported students with respect to: (1) performance in answering criterion item questions, and (2) time required to administer the items. The same film was shown to both groups; at its conclusion, a posttest consisting of 10 or more items related to the film seen and another not seen was administered. The posttest performance (items correct and time required for administration) of transported students was compared with that of a similar age-grade class which was not transported. The results indicated that there is no significant difference between the posttest performance of the bussed and the nonbussed groups. The time required for administering the test was longer for the bussed groups. (Author)
AN INVESTIGATION OF THE EFFECTS OF TRANSPORTING EMH STUDENTS TO A MEDIA EVALUATION SITE

ABSTRACT

In order to enlarge the sample of educable mentally retarded children used in the evaluation of media at the Computer Based Project (Syracuse, N.Y.), a number of classes were transported by bus to the Project site. The investigator questioned whether observed differences in manifest anxiety and attention between transported and non-transported students might affect responses to the media being evaluated. The purpose of this study was to ascertain whether or not there were differences between transported and non-transported students with respect to: (1) performance in answering criterion item questions, and (2) time required to administer the items. The same film was shown to both groups; at its conclusion, a posttest consisting of 10 or more items related to the film seen and another not seen was administered. The posttest performance (items correct and time required for administration) of transported students was compared with that of a similar age-grade class which was not transported. The results indicated that there is no significant difference between the posttest performance of the bussed and the non-bussed groups. The time required for administering the test for the groups indicated that it may take longer for bussed students to complete a test. Significant differences were found between primary and intermediate groups, primary groups receiving lower correct scores.
SPECIAL REPORT No. 731
COMPUTER-BASED PROJECT for the EVALUATION of MEDIA for the HANDICAPPED

Title: AN INVESTIGATION OF THE EFFECTS OF TRANSPORTING EMH STUDENTS TO A MEDIA EVALUATION SITE

BACKGROUND

The Computer Based Project for the Evaluation of Media for the Handicapped, based on contract #OEC-9-423617-4357 (616) between the Syracuse (N.Y.) City School District and the Media Services and Captioned Films Branch, Bureau of Education for the Handicapped (United States Office of Education) for the five year period July 1, 1969 through June 30, 1974. The major goal is to improve the instruction of handicapped children through the development and use of an evaluation system to measure the instructional effectiveness of films and other materials with educable mentally handicapped (EMH) children, in-service training and media support for special teachers, and studies related to the evaluation process and the populations used.

The Project has concentrated on the 600 films and 200 filmstrips from the Media Services and Captioned Films (BHE - USOE) depository; however, specific packages from Project LIFE, various elementary math curricula, and selected programs from Children's TV Workshop have also been evaluated. The evaluation model used requires that: 1) objectives of materials be specified and written; 2) instruments be constructed to test and measure effectiveness; and, 3) children be the major sources of evaluation information. A number of instruments and methodologies are employed in the gathering of cognitive and affective data from 900 EMH children and 80 special teachers to make the effectiveness decisions. Over half of the EMH population can neither read or write; therefore, a unique Student Response System (SRS) is employed, consisting of a twenty station G.E.-1000 SRS which can be operated in a group or individual recording mode and is connected to a remote computer system. The computer capabilities consist of remote telephone connections to the Rome (N.Y.) Air Development Command, the Honeywell time-shared network, and the Schenectady (N.Y.) G E Research and Development Center; and batch mode capabilities of the Syracuse City Schools, Syracuse University, and various commercial sources.

In-service and media support activities provide on-the-job training for teachers, teacher aides, equipment, and materials to the special teachers in the city schools. The research activities have centered around investigations and special problems related to the development of the evaluation model. The four major areas considered are: 1) testing effects, 2) captioning effects, 3) special student characteristics; and, 4) evaluation procedures validation.

Documentation of the major activities appear in the five annual reports and the 600 evaluations prepared on materials used. Staff members were encouraged to prepare special reports and the attached paper is one of these. The opinions expressed in this publication do not necessarily reflect the position or policy of the Computer Based Project, the United States Office of Education, or the Syracuse City School District; and no official endorsement by any of the agencies should be inferred.
AN INVESTIGATION OF THE EFFECTS OF TRANSPORTING
EHM STUDENTS TO A MEDIA EVALUATION SITE.

The Computer Based Project for the Evaluation of Media for the Handicapped
(CBP, 1969) is determining the Instructional effectiveness of selected captioned
films for educationally mentally handicapped (EHM) children. In an effort to
broaden the sample of children used in the evaluation process, a number of
classrooms are transported to an established center in an inter-city school of
a large city school system. At this center the children view a given film and
respond to cognitive and affective items intended to measure effectiveness.

PROBLEM

Over a period of time some observations have been made of apparent differ-
ences in anxiety, attention, and to a lesser degree, in attitude, of those
students who have been transported from other schools into the Computer Based
Project (CBP, 1969) site for media evaluation purposes and those not transported.
It would seem that the additional commotion brought upon by the bussing process
may heighten the exhibited levels of activity for many of the bussed students
to a greater extent than those who are residents (not bussed in) at the school
site. This heightened behavior may reduce the effectiveness of the educational
process afforded at CBP as a discrepancy in the reported performance on the
criterion test and the length of time required to administer the various pre
and post test items, consisting of several cognitive and affective questions
administered to the bussed and resident populations.

QUESTION

The specific question of this study will be to determine if there are
differences between the bussed and resident EHM children on posttest items,
between levels, and time to administer the test instrument.
ASSUMPTIONS

Several general assumptions are made in conducting this study, specifically:

1. All the children in EMR classrooms perform as though they are retarded.
2. Classrooms with the same age-grade designation are representative samples of that strata of the EMR population.
3. The conditions for viewing films at the center are standard.
4. The matrix sample of media items used as a posttest is a valid measure of cognitive achievement.
5. The effects of different films are the same.
6. The effects of a film are considered standard for a given level of child.

LIMITATIONS

Because this study was done using data regularly generated for the evaluation process rather than being organized as an independent study the following limitations are recognized.

1. The actual items making up the posttest may vary from presentation to presentation because of the nature of making item sets uses an independent sorting process which results in different items being placed together for a given presentation.
2. Classrooms do not have equal numbers of members.

DESIGN

The standard evaluation process at CBP are followed. The classrooms arrive at their regular scheduled time and view a film scheduled for evaluation without modification except for increasing the number of classrooms viewing a given film.
SAMPLE

The classrooms scheduled into the on-going regular evaluation process were classified as transported (Group I) or resident (Group II). A sample consisting of six classes in each group who had viewed the same film and taken the posttest consisting of at least four of the items presented to all groups. A further subgrouping is made of three classrooms, each at the Primary and Intermediate levels as determined by the ago-grade classification of the School District's Special Education Department.

INSTRUMENTATION

The criterion instrument will be the regularly scheduled posttest of 10 cognitive items. Five are prepared to measure achievement for the film "seen" and five having to do with a film that was not seen ("unseen"). Four items dealing with liking the film are added to the above posttest set and administered immediately following the viewing of the film. The criterion test given each group varies slightly from class to class because of the matrix sampling procedure employed to select items. An instrument of responses will be statistically constructed by computing a proportion correct for each class on the posttest cognitive items and using this as a score for the group.

TREATMENT

Each classroom was shown a film under the regular evaluation process and scheduled for viewing a film in the evaluation center's student response room, and following the film, administered the posttest. A stop watch was started as the first item was asked and stopped after the last item had been responded to by the class. The total elapsed time for the posttest administration was so recorded for each classroom.
DATA ANALYSIS

The sample population was organized into a 2 x 3 design with two independent measures -- cognitive score and time to administer the items. Two categories of academic level are recognized as primary and intermediate with three classrooms representing each level with one classroom matched on one of three films and the criteria of being transported and not being transported as shown in Figure 1.

![Diagram of study design](image)

**FIGURE 1 - SCHEMATIC SOLID DEPICTING THE STUDY DESIGN**

Two separate data collections were done for Primary and for Intermediate to allow for the different films used in each grade level because of the schedule and classification of appropriateness in the evaluation process.

The data was processed through the Project's student response media evaluation system and reduced by a computer program giving criterion test item scores for each class on each item in the various questions administered.
Immediately after the students have viewed the media. Because the number of items answered varied between classrooms a mean item score was computed by dividing the number of items answered into the mean criterion score obtained for the class thus yielding a value comparable across classrooms, as shown in Table 1 below.

### TABLE ONE: SUMMARY DATA FOR PRIMARY AND INTERMEDIATE GROUPS

<table>
<thead>
<tr>
<th>Media No.</th>
<th>Status Level</th>
<th>Media Type</th>
<th>Mean Score</th>
<th>Number of Items</th>
<th>Mean Item Score</th>
<th>Total Ques.</th>
<th>Time Req.to Adm. Q.</th>
<th>Mean Item Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>4299</td>
<td>Bus. Prim. #1</td>
<td>Prim.</td>
<td>1.8</td>
<td>4</td>
<td>0.45</td>
<td>13</td>
<td>15.5</td>
<td>1.19</td>
</tr>
<tr>
<td>2150</td>
<td>Bus. Prim. #2</td>
<td>Prim.</td>
<td>3.12</td>
<td>5</td>
<td>0.62</td>
<td>14</td>
<td>16.13</td>
<td>1.23</td>
</tr>
<tr>
<td>1010</td>
<td>Bus. Prim. #3</td>
<td>Prim.</td>
<td>1.8</td>
<td>5</td>
<td>0.36</td>
<td>14</td>
<td>15.42</td>
<td>1.1</td>
</tr>
<tr>
<td>1299</td>
<td>Res. Prim. #4</td>
<td>Prim.</td>
<td>3.33</td>
<td>7</td>
<td>0.48</td>
<td>11</td>
<td>11.32</td>
<td>1.02</td>
</tr>
<tr>
<td>2150</td>
<td>Res. Prim. #5</td>
<td>Prim.</td>
<td>2.0</td>
<td>4</td>
<td>0.50</td>
<td>13</td>
<td>9.23</td>
<td>.71</td>
</tr>
<tr>
<td>1010</td>
<td>Res. Prim. #6</td>
<td>Prim.</td>
<td>0.75</td>
<td>2</td>
<td>0.375</td>
<td>10</td>
<td>5.5</td>
<td>.55</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Prim.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Media No.</th>
<th>Status Level</th>
<th>Media Type</th>
<th>Mean Score</th>
<th>Number of Items</th>
<th>Mean Item Score</th>
<th>Total Ques.</th>
<th>Time Req.to Adm. Q.</th>
<th>Mean Item Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>2034</td>
<td>Bus. Inter. #1</td>
<td>Inter.</td>
<td>3.77</td>
<td>5</td>
<td>0.76</td>
<td>13</td>
<td>13.08</td>
<td>1.0</td>
</tr>
<tr>
<td>2034</td>
<td>Bus. Inter. #2</td>
<td>Inter.</td>
<td>2.3</td>
<td>5</td>
<td>.43</td>
<td>14</td>
<td>13.42</td>
<td>.95</td>
</tr>
<tr>
<td>2034</td>
<td>Res. Inter. #4</td>
<td>Inter.</td>
<td>2.67</td>
<td>3</td>
<td>0.89</td>
<td>8</td>
<td>7.52</td>
<td>.94</td>
</tr>
<tr>
<td>2034</td>
<td>Res. Inter. #5</td>
<td>Inter.</td>
<td>3.0</td>
<td>4</td>
<td>0.75</td>
<td>8</td>
<td>4.16</td>
<td>.52</td>
</tr>
<tr>
<td>2034</td>
<td>Res. Inter. #6</td>
<td>Inter.</td>
<td>2.0</td>
<td>4</td>
<td>0.50</td>
<td>8</td>
<td>14.05</td>
<td>.75</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Inter.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The "Number of Items" column indicates the items related to the cognitive information from the film seen. The "Total Questions" column includes all items; consisting of cognitive items on the film seen (generally 5 to a set), cognitive items on a film not seen (generally 5 to a set), and affective items dealing with the class's attitude toward the film seen (generally 4 items).
for a total maximum of 14 items. The variance from this maximum is caused by differences in the makeup of the item set; i.e., number of items available paired or not paired with another film and administering them, i.e., the length of time available for seeing the film and asking questions, presence or absence of intervening activities (fire drills, bus schedules, special school programs, etc.).

Summary Table One indicates the reported data for the primary and intermediate levels, bussed and resident populations. The time required for the system operator to administer the test items is shown in terms of the number of questions divided by total time.

DISCUSSION

The first null hypothesis that "There is no significant \( p = .05 \) difference between bussed and resident groups on number correct and time to administer," was tested using the Mann-Whitney U test. This test is considered to be one of the more powerful of the nonparametric measures for determining whether or not two independent groups have been drawn from the same population (Siegel, 1956; Roscoe, 1971). This test assumes that the ordering of the scores will occur in a systematic way such as

\[
\begin{array}{cccccccc}
B & R & B & R & B & R & B & R \\
\end{array}
\]

or

\[
\begin{array}{cccc}
R & B & R & B \\
\end{array}
\]

and detects a significant change from this arrangement. The actual computations are summarized below in Table II for Primary level and in Table III for Intermediate level.
### TABLE TWO
CRITERION SCORE

**MANN-WHITNEY U TEST FOR PRIMARY LEVEL, BUSSED AND RESIDENT STUDENTS**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SCORES</td>
<td>.36</td>
<td>.37</td>
<td>.45</td>
<td>.48</td>
<td>.50</td>
<td>.62</td>
</tr>
</tbody>
</table>

Computation of $U$: 
$n_1 = 3; n_2 = 3; p = .5$ (Table J)

\[
1 + 2 + 2 = U_r = 5 \\
1 + 3 = U_b = 4
\]

A tabled value of $U = 4$ with $N_1 = 3$ and $N_2 = 3$ (Table J, p. 271, Siegel, 1956) results in a $p = .5$ which is greater than our hypothesized $p = .05$. Therefore, we find that no significant difference exists between primary groups on the response to the media test items (mean criterion test score).

### TABLE III

**MANN-WHITNEY U TEST FOR INTERMEDIATE LEVEL, BUSSED AND RESIDENT STUDENTS**

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Bus</th>
<th>Res.</th>
<th>Bus</th>
<th>Bus</th>
<th>Res.</th>
<th>Bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCORES</td>
<td>.43</td>
<td>.50</td>
<td>.75</td>
<td>.76</td>
<td>.89</td>
<td>.92</td>
</tr>
</tbody>
</table>

Computation of $U$: 
$n_1 = 3; n_2 = 3; p = .5$ (Table J)

\[
1 + 1 + 2 = U_r = 4 \\
2 + 3 = U_b = 5
\]
The tabled probability for \( U = 4 \) when \( N_1 \) and \( N_2 = 3 \) is \( p = .5 \).

Here also, no significant difference exists between secondary groups on the response to the mean criterion test items. \( P = .5 > P .05 \). The analysis of the time to administer the items will follow later.

A second hypothesis was the "There are no differences between the scores between primary and Intermediate level." A Mann-Whitney U test is summarized in Table IV below:

| TABLE IV |
| MANN-WHITNEY U TEST FOR DIFFERENCES BETWEEN LEVELS |

<table>
<thead>
<tr>
<th>SCORES</th>
<th>.36</th>
<th>.375</th>
<th>.43</th>
<th>.45</th>
<th>.48</th>
<th>.50</th>
<th>.50</th>
<th>.62</th>
<th>.75</th>
<th>.76</th>
<th>.89</th>
<th>.92</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP</td>
<td>P</td>
<td>P</td>
<td>I</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>

Computation of U: \( N_1 = 6 \quad N_2 = 6 \quad p = .02 \)

\[ U_1 = 2 + 5 + 6 + 6 + 6 + 6 = 31 \]

\[ U_p = 1 + 1 + 1 + 2 = 5^* \]

The probability of obtaining a \( U = 5 \) is \( p = .02 \) which is less than our \( p = .05 \); therefore, there is a significant difference between the levels on the percent correct on cognitive items. In Table I, it can be noted that both the bussed and resident mean item correct values for Intermediate are roughly 50% more than those for primary.

TIME TO ADMINISTER CRITERION TEST

The same procedures described above for the Mann-Whitney U were used on the mean time per item as summarized in Table V.
No significant differences were found for intermediate level sub-group when classed as resident or bussed; however, there was a significant difference for primary level with bussed subjects (mean time = 1.17) taking longer for the testing situation than resident students (mean time = .96). No significant differences were present between levels when control for transportation was not made; however, a perusal of the data suggests a tendency for primary children to take longer than intermediate as indicated by four intermediate mean times being 1.0 or less whereas only two of the primary times were in this category. The extremely long time (1.75) for the #6 intermediate group suggests something may have interfered with the test such as a behavior problem, teacher interruption, etc. If the U test is recomputed omitting this group, a U = 8 is obtained. The probability of this value is p = .12 for N₁ = 5 and N₂ = 6 which still indicates a non-significant difference in sub-group level (at the p = .05 level).

The rank-order correlation between the obtained scores and the time to administer the item is shown in Table V the resulting rho = -.20 suggesting...
essentially no correlation between the percent correct and the time to take the test. The fact that rho is negative may even suggest that the longer time taken for the test, tends to lower the score.

The results indicate no significant differences on cognitive items between groups bussed to the Project and those not being transported to the Project. The significant differences between levels on percent correct were consistent and as are expected; i.e., intermediate children are expected to have higher correct scores than primary children. There does seem to be a difference in the time to administer the test items to primary children, with those being transported taking the longest time. A small negative correlation between the percent correct and the time to take the item was found.

CONCLUSIONS

These findings tend to suggest that transporting children to the Project site has had no appreciable effect on their response and cognitive items as long as the extra time to administer the items to transported primary children can be allowed. It is further implied that transporting the children seems not to have affected their cognitive responses.

REFERENCES

