The Effectiveness of Special Education Placement and Services: A Descriptive Study

Michelle D. Lewis
Western Michigan University

Follow this and additional works at: http://scholarworks.wmich.edu/masters_theses
Part of the Educational Psychology Commons, and the Special Education and Teaching Commons

Recommended Citation
Lewis, Michelle D., "The Effectiveness of Special Education Placement and Services: A Descriptive Study" (1992). Master's Theses. 862.
http://scholarworks.wmich.edu/masters_theses/862

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE EFFECTIVENESS OF SPECIAL EDUCATION
PLACEMENT AND SERVICES:
A DESCRIPTIVE STUDY

by

Michelle D. Lewis

A Project Report
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Specialist in Education
Department of Psychology

Western Michigan University
Kalamazoo, Michigan
August 1992

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This descriptive study evaluated the effectiveness of special education placement and services over a three year period of time. The files of seventeen students were chosen for the study. The study took place in two school districts in Southwestern Michigan.

The findings from this study indicate that: (a) there was an increase in test scores in the area of Performance IQ, and (b) there was a significant decrease in test scores in the area of Math Calculation. There was a general trend, although not significant, of a decrease in scores in all areas of ability and achievement except for the area of Basic Reading Skills. It was concluded that no conclusion could be made as to the efficacy of special education from the results of this study.
ACKNOWLEDGMENTS

I wish to express a special thank you and sincere appreciation to my advisor and committee chairperson, Dr. Howard E. Farris, for the assistance, encouragement and advice that he has given me over the last three years. I would also like to extend a special thank you to Dr. Michael Bahr for his assistance and guidance with the research methodology and data analyses of my project. To Dr. Galen Alessi I express my appreciation for encouragement in both my program and research. The direction and support of my Committee has made this final project possible.

Finally, I would like to thank my family and friends for all of their encouragement and love shown throughout my graduate program. I couldn't have done it without you. Thanks to you all.

Michelle D. Lewis
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700  800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The effectiveness of special education placement and services: A descriptive study

Lewis, Michelle Davina, Ed.S.
Western Michigan University, 1992
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>SECTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>METHOD</td>
<td>5</td>
</tr>
<tr>
<td>Subjects</td>
<td>5</td>
</tr>
<tr>
<td>Setting</td>
<td>5</td>
</tr>
<tr>
<td>Materials</td>
<td>6</td>
</tr>
<tr>
<td>Procedure</td>
<td>6</td>
</tr>
<tr>
<td>RESULTS</td>
<td>7</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>14</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>A. Data Collection Sheet</td>
<td>18</td>
</tr>
<tr>
<td>B. Human Subjects Institutional Review Board Approval</td>
<td>20</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>22</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Variable Analysis (Wilcoxon Matched-Pairs Test) ................................................. 9
2. Difference Scores According to Districts ................................................................. 11
3. District Comparison Analysis (Mann-Whitney U Test) ........................................ 13
INTRODUCTION

Modern societies have generally established equality of educational opportunity for all children as an essential goal. Yet, helping all students learn what is essential for their intelligent participation in a modern society has proved to be a difficult task and many educators have given up the effort. Finding effective and practical learning environments that meet student's needs has been a continuing challenge for researchers and practitioners. The underlying premise of this work is that individuals learn in different ways and at different rates, and that a major responsibility of schools is to accommodate these differences in order to maximize each student's education (Wang & Walberg, 1985).

The educator continually devises and applies new instructional treatments hoping for improved results. Since learners differ, the search for the best methods of instruction should be supplemented by a search for ways to fit the instruction to each kind of learner. Special education was designed to do just that; it would give the most appropriate method of instruction to learners who learn in different ways and at different times. Even so, a study completed by Carlberg and Kavale (1980) found that special classes were found to be significantly inferior to regular class placement for students with below average IQ's. The lack of special education effectiveness has led Glass (1983) to conclude there is little evidence substantiating the benefits of special education services for mildly handicapped children.

Evaluating the effectiveness of special education programs has certainly become a major focus in recent times, both in the public arena, as well as in the professional literature. There has been a general conception in the special education
effectiveness literature that placement of mildly handicapped students into special
education programs is of questionable value (Tindal, 1985). As Johnson (1962)
noted almost 30 years ago, “There is almost universal agreement that the mentally
handicapped children enrolled in special classes achieve, academically, significantly
less than similar children who remain in regular grades” (p. 58).

In the rush to condense the literature and summarize the conclusions, there has
been a general dismissal of the many methodological issues involved in conducting
evaluation research. Very little attention has been given to the manner in which
empirical findings have been produced, with more focus on results than on methods.
As a consequence for this, researchers essentially know less than has been proven.

In an attempt to identify methodological features that may influence study
outcomes, several features of the primary research were analyzed to determine their
potential correlation with effect size (Tindal, 1985). For example, variables that were
considered included IQ differences between regular and special education classes,
length of placement, age of students, interval between treatment and outcome
measurement, attrition from regular or special class placement, sample size, date of
publication, blindness to treatment, assignment of teachers to groups and reactivity of
measurement. These factors represent equally important qualifications and may have
important influences on the validity of the findings in the research. Due to the lack of
valid empirical evidence and sound methodology, the only conclusion that can be
made at this time is that no conclusion is yet available about special education efficacy.

Mainstreaming is a movement bringing profound changes to long established
practices in special education. Its legislative mandate, P.L. 94-142 (Stainback,
Stainback, & Forest, 1989), called for educational placement in the least restrictive
environment. A growing number of parents and educators are beginning to advocate
that all students be integrated into the mainstream of regular education, including those who have traditionally been labeled severely and profoundly handicapped (Stainback et al., 1989). For many exceptional children, this meant placement in the regular class instead of the special education classroom, which has been the favored educational arrangement. The advocates of this movement essentially believe that it is time to stop developing criteria for who does or does not belong in the mainstream, and that the spotlight should be turned instead toward increasing the capabilities of the regular education mainstream to meet the unique needs of all students.

Justification for the mandate was found in a series of efficacy studies beginning in the 1930s which suggested that the special class may be inappropriate for the education of exceptional children (Carlberg & Kavale, 1980). The results of existing research, when integrated statistically, demonstrated that special education class placement is an inferior alternative to regular class placement in benefiting children removed from the educational mainstream. There is a lack of conclusive findings because of the fact that, in the studies that were done, there was not ideal internal validity within the studies. Methodologically, random assignment of students is the preferred method for initially assuring the equivalence of groups. Because random assignment is difficult in practice, due to many ethical issues, various techniques such as, matched pairs, analysis of covariance, residual gain scores, and wishful thinking have substituted for random assignment (Carlberg & Kavale, 1980). No great differences among classes of outcome measures were identified. Thus, regardless of whether achievement/personality/social, or other dependent variables were chosen for investigation in these studies, no differential placement effects emerged across studies.
The purpose of this descriptive study is to examine the gains made by students who are and who have been, receiving special education services. Descriptive data will be obtained between present information and information gathered three years ago. Any significant differences will be reported. The hypothesis for this study is: “Is there a significant difference between ability and achievement when placed in a special education program over a three year period of time?”
METHOD

Subjects

Seventeen students' files were randomly chosen for this study. Files from four females and thirteen males were used. The subjects ranged in age from 9 years 4 months to 18 years 1 month. The grade level that the students were enrolled in ranged from third grade to eleventh grade.

The subjects included students with mild handicaps (i.e., Specific Learning Disabled, Speech and Language Impaired and Educable Mentally Impaired) students who had been referred and are currently receiving special education services. The students selected were those who had their three year re-evaluation between January and December 1991. The students had qualified for various special education placements and services. Four of the subjects were Educable Mentally Impaired students and thirteen of the subjects were Learning Disabled in various areas of certification. The certification and services that the subjects were given was not a factor in this study.

Setting

The study took place at two school districts in Southwestern Michigan. The population in the two districts combined was about 10,000. The districts included C and D rated schools. The ratings of these schools is due to the population of the districts in which the schools reside. The population consists of mostly blue collar workers and lower to middle income families. The data collection itself took place in a small private room in the Middle School of one of the districts.
Materials

The information for this study came directly from the cumulative files of the subjects in the study. These files contain many pieces of pertinent information which include: Multidisciplinary Evaluation Team forms, teacher anecdotes, past services received, test protocols and scores, past medical and academic as well as family history. Information coded from these files included demographic information and standard scores from ability and academic testing. A data collection sheet was also created and used to aid in the acquisition of the information needed for this study (See Appendix A).

Procedure

The files were chosen randomly through the following process: The special education director was given a set of criteria for the files that were to be acquired. The criteria specified that the files were to be of those students who were mildly handicapped and who received a three year re-evaluation in 1991. The special education director provided the files of those students who met these criteria. To insure confidentiality the files were then taken to a private room where the data were coded. The type of data that were coded included subject (or I. D.) number, gender, age, grade, school district, and number of retentions. Other data that were coded and recorded included, standard scores from ability and academic achievement tests. These latter data served as the dependent measures for this study. A comparison was made of the standard scores from 1988 and 1991.
RESULTS

The results of this study were analyzed with two nonparametric statistical tests. First, the Wilcoxon matched-pairs signed-rank test was used as a dependent t test. The Wilcoxon was determined to be the most appropriate since the two populations sampled were related, but the assumptions and requirements of the parametric procedures could not be met. In this instance the standard scores obtained from 1988 were not necessarily obtained with the same tests as the scores in 1991. The hypothesis being tested is that standard scores obtained in 1988 for all subjects are higher (lower) than standard scores obtained in 1991.

The Wilcoxon was used to compare the standard scores obtained in evaluations completed in 1988 and 1991 for all seventeen subjects involved in the study. The following are the steps use to obtain the values with which comparisons could be made. The hypothesis posed was there would be no significant difference between scores obtained in 1988 to those obtained in 1991.

Step 1. Obtain the difference between the scores for each subject. It makes no difference whether you subtract the subject's score for condition 1 from condition 2 or vice versa as long as you are consistent for all pairs. For this study the scores obtained from 1991 were subtracted from the scores obtained from 1988.

Step 2. Rank order the absolute value of the difference scores obtained in Step 1. Eliminate from further consideration any pair that has a zero difference. Ties in difference scores are handled by assigning the average rank to the tied scores.
Step 3. Separate the ranks in terms of the sign of the difference scores. For example: 

9 - ranks

7 + ranks

1 ties

Step 4. Sum the ranks for the positive differences to get a Mean Rank +. Sum the ranks for the negative differences to get a Mean Rank -.

Step 5. A comparison of the ranks is done and you get a score which is a Z score in this study. The obtained Z-score is compared with the critical value from the sampling distribution of the Z statistic. The null hypothesis can be rejected at the p=.05 level if the obtained Z value is equal to or less than the critical Z value.

In the first statistical analysis using the Wilcoxon, the following variables were compared: Full Scale IQ, Verbal IQ, Performance IQ, Basic Reading Skills, Reading Comprehension, Math Calculation, Math Reasoning, and Written Expression standard scores from 1988 and 1991 evaluations. For example, the Full Scale IQ from 1988 would be compared to the Full Scale IQ from 1991. Standard scores from all seventeen subjects are presented in Table 1 with values listed for each variable analyzed.

Using the Wilcoxon matched-pairs signed-ranks test, the only variable in which there was a significant difference between standard scores obtained in 1988 versus those obtained in 1991, was for the Math Calculation variable. The mean standard score obtained for this variable was in 1988 ($M=79$), which was significantly higher than that obtained in 1991 ($M=74.588$); critical value revealed a $Z= -2.1776$, $p < .05$. This finding indicates that achievement in the area of Math Calculation declined from the 1988 scores to those obtained in 1991. There was no
Table 1
Variable Analysis (Wilcoxon Matched-Pairs Test)

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean Rank</th>
<th>Z-Score</th>
<th>2-Tailed P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ-88</td>
<td>17</td>
<td>85.118</td>
<td>18.265</td>
<td>10.00</td>
<td>-1.1376</td>
<td>.2553</td>
</tr>
<tr>
<td>FSIQ-91</td>
<td>17</td>
<td>83.176</td>
<td>15.903</td>
<td>6.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIQ-88</td>
<td>17</td>
<td>85.882</td>
<td>16.896</td>
<td>8.61</td>
<td>-.4912</td>
<td>.6233</td>
</tr>
<tr>
<td>VIQ-91</td>
<td>17</td>
<td>84.941</td>
<td>13.278</td>
<td>8.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQ-88</td>
<td>17</td>
<td>86.294</td>
<td>19.032</td>
<td>7.90</td>
<td>-.5688</td>
<td>.5695</td>
</tr>
<tr>
<td>PIQ-91</td>
<td>17</td>
<td>84.176</td>
<td>17.735</td>
<td>9.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRS-88</td>
<td>17</td>
<td>74.000</td>
<td>13.734</td>
<td>8.83</td>
<td>-.7756</td>
<td>.4380</td>
</tr>
<tr>
<td>BRS-91</td>
<td>17</td>
<td>77.529</td>
<td>12.284</td>
<td>8.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDGC-88</td>
<td>17</td>
<td>75.529</td>
<td>12.846</td>
<td>10.17</td>
<td>-1.2152</td>
<td>.2243</td>
</tr>
<tr>
<td>RDGC-91</td>
<td>17</td>
<td>72.529</td>
<td>11.727</td>
<td>6.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATHC-88</td>
<td>17</td>
<td>79.000</td>
<td>13.238</td>
<td>10.21</td>
<td>-2.1776</td>
<td>.0294*</td>
</tr>
<tr>
<td>MATHC-91</td>
<td>17</td>
<td>74.588</td>
<td>15.399</td>
<td>6.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATHR-88</td>
<td>17</td>
<td>81.882</td>
<td>16.740</td>
<td>8.55</td>
<td>-1.4483</td>
<td>.1475</td>
</tr>
<tr>
<td>MATHR-91</td>
<td>17</td>
<td>78.059</td>
<td>18.613</td>
<td>6.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRITE-88</td>
<td>17</td>
<td>70.588</td>
<td>10.765</td>
<td>8.83</td>
<td>-.5241</td>
<td>.6002</td>
</tr>
<tr>
<td>WRITE-91</td>
<td>17</td>
<td>69.538</td>
<td>13.170</td>
<td>5.43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

other significant differences found between the standard scores of any of the other variables.

Next, the Mann-Whitney U test was used to compare the standard scores obtained in District 1 to those obtained in District 2. The Mann-Whitney U test is a
nonparametric alternative to the independent $t$ test. It may be used in instances were the groups sampled are independent but all assumptions or requirements for the independent $t$ test cannot be met. In this instance, the shapes of the two populations are assumed to be similar. The hypothesis tested is that the scores obtained from one population are larger (or smaller) than the scores from the second population.

When the Mann-Whitney $U$ test is used to compare independent groups, the scores for each group must be ranked and then a Mean Rank must be obtained. An analysis of the two Mean Rank scores will yield a $U$ score which must be compared to the critical value of $U$ at the $p=.05$ level. If the $U$ score is less than or equal to the critical value for $U$ then the null hypothesis can be rejected.

The same variables were compared for the two groups as those used in the Wilcoxon. First the difference (or discrepancy) scores between test scores for 1988 and 1991 were obtained from both districts. Then the Mean Difference scores between District 1 and District 2 were compared and ranked. The scores were then analyzed statistically using the Mann-Whitney $U$ test to determine any significant difference ($p=.05$) (See Table 2 for variable analysis between districts). Note that negative difference scores indicate higher academic achievement 1991 than 1988.

Using the Mann-Whitney $U$ test, Table 3 shows that over time, (1988-1991), subjects in District 2 ($M = -4.6$) improved their Performance IQ scores relative to District 1 ($M = 4.9$), $Z = -1.9032$, $p < .05$. There was no significant difference between any other variables between the two districts.
### Table 2
Difference Scores According to Districts

#### DISTRICT 1 (N=12)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>89.08</td>
<td>20.01</td>
<td>85.58</td>
<td>17.50</td>
<td>3.500</td>
<td>5.792</td>
</tr>
<tr>
<td>VIQ</td>
<td>88.33</td>
<td>18.95</td>
<td>87.00</td>
<td>14.32</td>
<td>1.333</td>
<td>7.303</td>
</tr>
<tr>
<td>PIQ</td>
<td>91.42</td>
<td>20.13</td>
<td>86.50</td>
<td>19.62</td>
<td>4.917</td>
<td>9.956</td>
</tr>
<tr>
<td>BRS</td>
<td>77.67</td>
<td>14.42</td>
<td>80.17</td>
<td>11.46</td>
<td>-2.500</td>
<td>12.724</td>
</tr>
<tr>
<td>RDGC</td>
<td>77.75</td>
<td>13.84</td>
<td>73.00</td>
<td>11.54</td>
<td>4.750</td>
<td>10.678</td>
</tr>
<tr>
<td>MATHC</td>
<td>79.92</td>
<td>14.43</td>
<td>77.33</td>
<td>14.98</td>
<td>2.583</td>
<td>8.512</td>
</tr>
<tr>
<td>MATHR</td>
<td>85.33</td>
<td>18.75</td>
<td>82.17</td>
<td>19.85</td>
<td>3.167</td>
<td>9.034</td>
</tr>
<tr>
<td>WRITE</td>
<td>71.42</td>
<td>12.41</td>
<td>71.11</td>
<td>14.84</td>
<td>2.222</td>
<td>19.240 *</td>
</tr>
</tbody>
</table>

*(There were only 9 students who had Written Expression test scores in District 1 so only these 9 scores were used in the analysis.)*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>75.60</td>
<td>8.53</td>
<td>77.40</td>
<td>10.50</td>
<td>-1.800</td>
<td>8.497</td>
</tr>
<tr>
<td>VIQ</td>
<td>80.00</td>
<td>9.67</td>
<td>80.00</td>
<td>9.90</td>
<td>.000</td>
<td>9.354</td>
</tr>
<tr>
<td>PIQ</td>
<td>74.00</td>
<td>8.15</td>
<td>78.60</td>
<td>12.01</td>
<td>-4.600</td>
<td>7.162</td>
</tr>
<tr>
<td>BRS</td>
<td>65.20</td>
<td>6.72</td>
<td>71.20</td>
<td>13.10</td>
<td>-6.000</td>
<td>13.982</td>
</tr>
<tr>
<td>RDGC</td>
<td>70.20</td>
<td>9.12</td>
<td>71.40</td>
<td>13.46</td>
<td>-1.200</td>
<td>7.530</td>
</tr>
<tr>
<td>MATHC</td>
<td>76.80</td>
<td>10.94</td>
<td>68.00</td>
<td>15.95</td>
<td>8.800</td>
<td>6.907</td>
</tr>
<tr>
<td>MATHR</td>
<td>73.60</td>
<td>5.68</td>
<td>68.20</td>
<td>11.39</td>
<td>5.400</td>
<td>9.072</td>
</tr>
<tr>
<td>WRITE</td>
<td>68.60</td>
<td>5.73</td>
<td>66.00</td>
<td>9.06</td>
<td>1.750</td>
<td>6.238 *</td>
</tr>
</tbody>
</table>

* (Only four Written Expression test scores were available from District 2 and were used in the analysis.)
<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean Rank</th>
<th>Z-Score</th>
<th>2-Tailed P</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQD</td>
<td>17</td>
<td>1.94118</td>
<td>6.87814</td>
<td>9.92 (88)</td>
<td>-1.1688</td>
<td>.2425</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.80 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIQD</td>
<td>17</td>
<td>.94118</td>
<td>7.67684</td>
<td>9.21 (88)</td>
<td>-.2642</td>
<td>.7916</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.50 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIQD</td>
<td>17</td>
<td>2.11765</td>
<td>9.80359</td>
<td>10.50 (88)</td>
<td>-1.9032</td>
<td>.0570</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.40 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRSD</td>
<td>17</td>
<td>3.52941</td>
<td>12.76283</td>
<td>9.50 (88)</td>
<td>-.6332</td>
<td>.5266</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.80 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDGCD</td>
<td>17</td>
<td>3.00000</td>
<td>10.01873</td>
<td>10.00 (88)</td>
<td>-1.2672</td>
<td>.2051</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.60 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATHCD</td>
<td>17</td>
<td>4.41176</td>
<td>8.38197</td>
<td>8.40 (88)</td>
<td>-1.2205</td>
<td>.2223</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.30 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATHRD</td>
<td>17</td>
<td>3.82353</td>
<td>8.81926</td>
<td>8.71 (88)</td>
<td>-.3701</td>
<td>.7113</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.70 (91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRITE-D</td>
<td>13</td>
<td>2.07692</td>
<td>16.01802</td>
<td>6.89 (88)</td>
<td>-.1549</td>
<td>.8769</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.25 (91)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DISCUSSION

When examining all seventeen subjects and analyzing significant gains or losses in ability or achievement between standard scores obtained in 1988 and those obtained in 1991, it was found that the only variable showing significant was the Math Calculation achievement score. The standard scores obtained for this variable in 1988 were found to be significantly higher than those obtained in 1991. While this appears to be an unusual finding, it may be due to some common differences in curriculum and instruction between mainstream and special education programming.

Two possible reasons are: (1) there tends to be a deemphasis in the special education classrooms on calculation and a focus on functional math skills such as the use of manipulatives to add, subtract, multiply and divide; and (2) the students may have performed less well on the 1991 test due to the lack of practice. Both would be necessary to improved performance on the types of calculations required for the standardized tests given in the initial evaluation and three-year reevaluation. Another factor that could contribute to the decrease in achievement in the area of Math Calculation could be the fairly common practice of students working below their ability and skills levels in special education settings. For example, students are frequently placed in materials which are below that of which they are capable for behavioral or other reasons, such as classroom overcrowding or lack of diverse instructional materials.

The data analysis from the Wilcoxon Matched Pairs statistical test also revealed a trend for test scores to generally decrease from 1988 to 1991. In all but one instance (Basic Reading), the mean scores of the ability and/or achievement areas...
all decreased from those obtained three years earlier. Although not significant, this raises some questions about the general efficacy of the educational programming for these youngsters. The findings are also consistent with many reported in the literature (Carlberg & Kavale, 1980; Shinn, 1986; Stainbeck et al., 1989).

Based on these findings, it has been suggested that special education may be an inferior placement for students when compared to the option of remaining in the regular education classroom. A meta-analysis by Carlberg and Kavale (1980) clearly support this contention. They found that “regardless of whether achievement, personality/social, or other dependent variables were chosen for investigation, no differential placement effects emerged among studies” (p. 304). In general they noted that special classes were found to be significantly inferior for students with below average IQ's.

When the two school districts were compared using ability and achievement variables of the students, Performance IQ scores for District 2 were significantly higher than the scores for District 1. No other significant differences were found between the districts. The difference in former scores observed between the two districts may be related to differences in the curricula. District 2 has adopted a curriculum for the special education students which requires a great deal of active participation by the students with direct hands-on experiences for problem-solving in math and science (Personal observation). It would be interesting to examine the relative contribution(s) of the curricular differences using curriculum-based measures in both districts with annual followup testing.

There were a number of limitations to this study. First, the sample size was small. Seventeen subjects were selected from those who met the criteria for participation in the study. A larger sample size could have provided a better indication
of the differences between the standard scores of the two districts. Secondly, the type of data available (standard or norm-referenced ability and achievement scores) tend to limit the level and nature of the analyses. As noted earlier, some form of curriculum-based or criterion-referenced assessment would be highly preferable. Another factor limiting the study is the lack of randomization of subjects in the "treatment" conditions due to ethical and practical reasons. It simply is not possible to "randomly assign" students who qualify for special education services to classrooms or programs within or across districts. The type of sample available for the study was also limiting. Of the seventeen subjects selected for the study, thirteen of them were diagnosed as having a learning disability and four of the subjects were certified as Educable Mentally Impaired. Ideally, a larger more diverse group of subjects representing a broad range of disabilities or handicapping conditions would have been desirable. An additional limitation, while not critical, was the type of data analyses procedures appropriate for the study. The Mann-Whitney U and the Wilcoxon were both chosen over more powerful tests due to some of the other limiting factors listed above. While these tests are appropriate for the types of analyses and data examined, a more accurate picture might have been obtained if single-subject or parametric designs could have been used.

While the study revealed only two significant differences, one positive and one negative, the data do show a trend that establish a basis for concern. It is clear from this limited sample that little or no progress has been made by these students after three years in special education. All of the "lack of achievement" cannot be attributed to "test error." As noted earlier, the data are consistent other studies in the literature and minimally suggest that further and more extensive examinations ought to be made to determine whether or not special education placement is effective for
handicapped students. There are several areas of research that would be helpful. These include research with variables known to be important in educational achievement. One variable would be that of "educationally engaged time." Time on-task has been positively related to achievement and could be studied across levels, handicapping conditions, programs and districts. Another possible area of research is related to effective teaching strategies, such as that used in "Direct Instruction". Data from field studies with these strategies clearly show that intense instructional programming, using effective characteristics of teaching, produced positive results i.e., Project follow-through. These types of strategies, consistently applied and tested could be extremely valuable for assessing the effectiveness of special education programming. Curricular variables, such as performance-based lessons with frequent sampling of student performance, could also be valuable determinants of student achievement and program efficacy. This type of research within the current context of special education is not only desirable but critical for future decision-making.

In summary, the purpose of this study was to determine the nature and extent of achievement of students in special education following three years of placement. Indicators used were ability and achievement norm-referenced test scores obtained from the student's files and from a three-year reevaluation being conducted by the school psychologist in the student's districts. The study, while limited in a number of ways, produced results that indicate gains were few if any. The study corroborates findings of others and adds yet another question mark as to the efficacy of special education.
Appendix A

Data Collection Sheet
Data Collection Sheet

I.D.# ____________ District: 1 2 (Circle One)
Gender: ____________ Date Of Birth: ____________

Name of Ability Test: ________________________________
Verbal IQ: ______________________
Performance IQ: ___________________
Full Scale IQ: ______________________

Name of Reading Test: ________________________________

Name of Math Test: ________________________________

Name of Written Expression Test: ______________________________
Appendix B

Human Subjects Institutional Review Board Approval
Date: April 9, 1992
To: Michelle Lewis
From: Mary Anne Bunda, Chair
Re: HSIRB Project Number: 91-12-07

This letter will serve as confirmation that your research protocol, "The effectiveness of special education placement and services" has been approved under the exempt category of review by the HSIRB. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the approval application.

You must seek reapproval for any changes in this design. You must also seek reapproval if the project extends beyond the termination date.

The Board wishes you success in the pursuit of your research goals.

xc: Farris, Psychology

Approval Termination: April 9, 1993
REFERENCES


