Meta-analysis of 17 studies was conducted to determine the effectiveness of retaining elementary school students in grade. Calculations with data from the 17 studies produced 217 effect size measures, most of which were measures of differences in reading and mathematics achievement. Some measures of differences in social adjustment, behavior, and attitude toward school were included in the sample of studies. Results provided support for the conclusions of an earlier study revealing that research evidence consistently points to negative effects of nonpromotion and arguing that proponents of retention plans are thereby obligated to show how their retentions plans will be successful when so many other plans have failed. Whereas some studies included in the meta-analysis showed positive outcomes of retention on academic achievement, the positive results appear to be found among middle-class, suburban, predominantly white samples and to be due to program characteristics related to the provision of individualized remedial help. It is concluded that such help can be provided through part-time pull-out programs, thereby avoiding the multi-billion dollar yearly expense of grade retention. (RH)
A SYNTHESIS OF RECENT RESEARCH ON NONPROMOTION: A FIVE YEAR FOLLOW-UP

C. Thomas Holmes
University of Georgia

Although reviews in the past have consistently failed to establish significant positive benefits of grade level retention, educators have continued to remain divided in their opinion of the benefits, or lack thereof, of this widespread practice (Gredler, 1984; Hess, 1978; Holmes, 1983; Holmes & Matthews, 1983; Holmes & Matthews, 1984; Jackson, 1975; Johnson, 1984; Reiter, 1973). Approximately one-half of the teachers and one-half of the parents surveyed in Utah agreed with the statement, "Children learn more academically by repeating a grade" (Fait, 1982). Teachers in Louisiana retaining large numbers of children could not be distinguished from teachers retaining relatively few on the basis of experience or amount of formal education. They were, however, different in their response to items eliciting their beliefs concerning the effect of retention practices on children retained as well as others in the classroom (Bennett, 1981).

A survey of the fifty states by *Education Week* (Changing Course, 1985, February) reported that 8 states had promotional gates tests in place and an additional 3 states had them under consideration. Fifteen states had exit exams and 4 were considering them. With the continuing movement toward promotional gates exams and exit exams, the question is becoming ever-increasingly important.

**Method**

The method employed in this investigation is referred to as meta-analysis. Meta-analysis, as defined by Glass (1978), is based on the concept of effect size. In this study, effect size was defined as the difference between the mean of the retained group and the mean of the promoted group, divided by the standard deviation of the promoted group. This procedure results in a measure of the difference between the two groups expressed in quantitative units which are additive across studies.

**Data Collection**

An attempt was made to gather all research reports that have become available since the meta-analysis first reported on at the AERA Annual Meeting in Montreal was completed (Holmes & Matthews, 1983). To this end, a systematic search of the following indexes was made: *Current Index to Journals in Education* (ERIC), *Research in Education* (ERIC), *Dissertation Abstracts International*, and *Master's Thesis In Education*. The original bibliography was expanded to include approximately 800 entries. In addition to
being recent, the reports had to meet the following selection criteria: the report must have: (a) presented the results of original research on the effects on pupils of grade-level retention in some combination of grades k-8, (b) described an investigation with an identifiable control group, and (c) included sufficient data to allow for the calculation or estimation (see Holmes, 1984) of effect sizes. For the purpose of this study transition classes were treated as retention.

Seventeen studies were identified which met all of the above criteria. Citations of these studies are contained in the reference list and are preceded by an asterisk. Three of the studies were paper presentations at AERA Annual Meetings, six were doctoral dissertations, three were master's theses, and five were reported in professional journals.

Results

Calculations with data from the seventeen studies produced 217 effect size measures. The largest portion of these were measures of differences in academic achievement, specifically reading and mathematics achievement. In addition to academic achievement, social adjustment, behavior, and attitude toward school were all addressed.

Academic Achievement

Fourteen of the seventeen studies addressed the effect of retention on the academic achievement of those children that are retained in grade. All of the effect sizes in this category
except four represented contrasts between grade peers. One would expect this type of comparison to favor the retained children because they have spent an additional year in school and are one year older than the comparison counterparts.

Table 1 displays a conventional count of the statistical significance reported in each of these reports. The overall picture remains confusing as four reports indicate statistically significant results in favor of the retainees, five indicate statistically significant differences in favor of the control or promoted students, and five report no statistical differences.

When mean ES's are calculated for each each of the studies with all measures of academic achievement (see Table 2), wide variations in study means are observed. The average of these study means approaches zero (-.06) with seven studies yielding negative mean effect sizes and seven yielding positive effect sizes. When all 149 individual ES's are plotted by frequency, a more meaningful pattern begins to emerge. The frequency distribution in Figure 1 appears to be bimodal. One mode is centered at approximately -.5 and the other at approximately+.4. The mode at -.5 coincides with the average obtained in the previous meta-analysis of 44 other studies (See Figure 2). The second mode, however, is different and is suggesting positive benefits. The implication is that, in fact, two different types of retention plans were being investigated. Although more of the ES's comprise the distribution centered around +.4 in the figure, the majority of these measures were obtained from only five of the studies.
### Table 1
Significance Reported on the Effects of Nonpromotion on the Academic Achievement of Students

<table>
<thead>
<tr>
<th>Helps</th>
<th>Mixed</th>
<th>N.S.</th>
<th>Mixed</th>
<th>Hurts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Askew, 1983</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Dolan, 1982</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hassen, 1980</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Leinhardt, 1980</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>May &amp; Welch, 1984</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Niklason, 1984</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Oldham, 1982</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Peterson, et al., 1985</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandoval &amp; Fitzgerald, 1985</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schuyler &amp; Matter, 1983</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Shepard &amp; Smith, 1985</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Talmadge, 1981/1982</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Vollrath, 1982</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Wright, 1979</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

TOTAL: 3 1 5 1 4
Table 2
Mean Study Effect Sizes
Academic Achievement

<table>
<thead>
<tr>
<th>STUDY</th>
<th>MEAN ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Askew, 1983</td>
<td>-.73</td>
</tr>
<tr>
<td>Dolan, 1982</td>
<td>+.47</td>
</tr>
<tr>
<td>Hassen, 1980</td>
<td>-.96</td>
</tr>
<tr>
<td>Leinhardt, 1980</td>
<td>-.13</td>
</tr>
<tr>
<td>May &amp; Welch, 1984</td>
<td>-.26</td>
</tr>
<tr>
<td>Niklason, 1984</td>
<td>-.12</td>
</tr>
<tr>
<td>Oldham, 1982</td>
<td>+.27</td>
</tr>
<tr>
<td>Peterson, et al., 1985</td>
<td>+.76</td>
</tr>
<tr>
<td>Sandoval &amp; Fitzgerald, 1985</td>
<td>+.10</td>
</tr>
<tr>
<td>Schuyler &amp; Matter, 1983</td>
<td>-.41</td>
</tr>
<tr>
<td>Shepard &amp; Smith, 1985</td>
<td>+.04</td>
</tr>
<tr>
<td>Talmadge, 1981/1982</td>
<td>-.89</td>
</tr>
<tr>
<td>Vollrath, 1982</td>
<td>+.70</td>
</tr>
<tr>
<td>Wright, 1979</td>
<td>+.29</td>
</tr>
</tbody>
</table>

AVERAGE                          | -.06    |
Figure 1
Frequency Distribution of ES's

Effect Size to Nearest Tenth
Figure 2
Effect Size Distribution of 1st Study

\[ \bar{ES} = -0.43 \]
The Effective Programs

The five studies reporting mean ES's that averaged +.50 were reread in an attempt to discover what these programs had in common that differed from the other studies, both the other nine studies identified in Table 1 and the 33 included in the former meta-analysis (Holmes & Matthews, 1984). Although descriptions of the retention plans being investigated were not detailed, a few tentative generalizations could be made.

Sample Characteristics. It appears that all of these studies were conducted in settings described as suburban and included few black subjects. The populations were described as middle class. One of the studies, in fact, reported that all subjects in both groups were white and from two-parent families.

Program Characteristics. The following is a composite of the descriptions given in the five reports. Potential failures were identified early and were given special help. If the decision was made to retain, an individualized and detailed educational plan was prepared for remediation of specific deficiencies. Parents were then consulted and written permission was obtained from the parents to retain the child. It seems key to these "successful" plans that these children were not recycled through the same educational programs but instead received specialized attention.

One of the plans, described in the greatest detail, places its retainees in special classes of 12 to 15 students with a
In addition, many of the children were "mainstreamed" into the regular program with their age peers for part of the day. Pupil-personnel-services teams consulted on a regular basis and a continuous evaluation was made. The results of this continuing evaluation could have at any time allowed for the child to rejoin his or her age cohort.

**Attitude Toward School, Behavior, and Personal/Social Adjustment**

Three studies measured attitude toward the school. Of these, two were among those measuring effective programs as described above. These two studies yielded 13 ES's with a mean of +.28, while the third study yielded two measures with a mean of -.65. Two studies measured student behavior. One of these was among the effective five and yielded 4 measures with a mean of +1.00, while the second study provided two measures with a mean of -.19.

Five studies provided measures of personal/social adjustment. One of these studies was among the five "effective programs" and yielded 36 ES's with a mean of +.23, while the other four studies yielded 24 ES's with a mean of +.06.

**Conclusions**

The greatest majority of the studies reviewed in this analysis provided additional support for the conclusions reached in the first meta-analysis. The concluding statement made at that time was that, "Because this cumulative research evidence
consistently points to negative effects of nonpromotion, the
burden of proof legitimately falls on proponents of retention
plans to show there is compelling logic indicating success of
their plans when so many other plans have failed." (Holmes &
Matthews, 1984, p. 232)

Although it appears that some success has been indicated in
those plans previously referred to as "successful", two notes of
cautions must be emphasized. First, in the only one of the five
studies that looked at differences past one year following
retention, these positive benefits tended to diminish with time
and in many cases vanished completely after two years. Second,
all comparisons were made with grade peers and not with age
peers. It is not unreasonable to expect older children with an
additional year of schooling to perform better on normed tests.
If, however, advances gained at the cost of an additional year
and at what Niklason (1984) estimated to have cost approximately
$5 billion in 1979-1980, are not maintained, one must still
wonder if these programs would really be "effective."

It appears that what was effective was a large amount of
individualized remedial help. The answer may lie in providing
this same help, complete with individualized education plans,
through part-time pull-out programs, and not through retention.
References


*Wright, J. B. (1979). The measured academic achievement of two groups of first grade students matched along five variables when one group has been retained* (Doctoral dissertation,

*Studies included in the meta-analysis.