Changes Over an Eight- and a Nine-Year Period in the Readiness Level of Entering First-Grade Pupils.

Data from national entering first-grade populations of 1964 (N=12,225) and of 1972/1973 (N=4275/6653) show an 11-point increase in median raw score on the 102-item Metropolitan Readiness Tests. The increase is shown to be greater at the lower than at the higher ranges of the score scale, with the first quartile increasing by 13 points, the third quartile by 9. Gain in the median total score is spread unevenly over the six subtest areas, being greatest in Alphabet (recognizing lower-case letters by name) and least in Copying. In 1964 the first quartile on the Alphabet test was 5 letters, in 1972/1973 it was 12 letters. These changes in readiness development would seem to have significant implications for first-grade instruction. (Author)
CHANGES OVER AN EIGHT- AND A NINE-YEAR PERIOD IN THE
READINESS LEVEL OF ENTERING FIRST-GRADE PUPILS

Blythe C. Mitchell

Test Department, Harcourt Brace Jovanovich, Inc.

Data from national entering first-grade populations of 1964 (N=12,225) and of 1972/1973 (N=4275/6653) show an 11-point increase in median raw score on the 102-item Metropolitan Readiness Tests. The increase is shown to be greater at the lower than at the higher ranges of the score scale, with the first quartile increasing by 13 points, Q3 by 9. Gain in median total score is spread unevenly over the six subtest areas, being greatest in Alphabet (recognizing lower-case letters by name) and least in Copying. In 1964 Q1 on the Alphabet test was 5 letters, in 1972/1973 it was 12 letters. These changes in readiness development would seem to have significant implications for first-grade instruction.
CHANGES OVER AN EIGHT- AND A NINE-YEAR PERIOD IN THE READINESS LEVEL OF ENTERING FIRST-GRADE PUPILS

Blythe C. Mitchell
Test Department, Harcourt Brace Jovanovich

There is some doubt today about the combined effects of the numerous developments in early childhood education that have occurred over the past few years. There have been Headstart, Sesame Street, and an increase in the number of nursery schools, of day care centers, and of school systems providing both four- and five-year old kindergartens.

Objectives of the Inquiry

Academic development has received different amounts of emphasis in the above programs; and the various results of certain of them have been separately investigated. For an estimate of what might be their overall academic effect this study compares recent measures of readiness at school entrance with those of several years ago. Does the combination of the many efforts to improve the "readiness to learn" level of the 3- to 5-year old population appear to have been successful — particularly with those children at the lower levels of readiness? To date, the author knows of no other large scale chronological study on possible improvement.

Method

This paper presents relevant data in the form of a comparison of Metropolitan Readiness Tests scores made by the standardization group in early October of 1964 with those made by two current nationally representative populations tested at the same grade placement — one in 1972, another in 1973.

Source of Data

The recent data were derived during two early phases of item-analysis experimentation in the development of a 1975 revision of these readiness tests. The current (1964) edition was included as an anchor test for all pupils taking any one of 27 experimental tests given in October of 1972 or of 1973 to a total of 10,928 entering first-grade pupils (4275 in 1972, 6653 in 1973). The school systems taking part in these testing programs numbered 31 from 26 states in 1972, and 50 different systems from 26 states in 1973. They were selected to be a representative sample of today's first grade pupils, with the two SES indices of adult years of education and family income of the school district areas agreeing closely with those for the total U. S. population (1970 Census).

The standardization of the current tests in 1964 involved 12,225 pupils (63 school systems, 299 schools, 12 states). These pupils were a selected part of the experimental groups involved in the comprehensive USOE Reading Research of 1964-65. Pupils in certain special studies, such as those involving bilingual or otherwise atypical children, were eliminated from the ERT norming population, as were others that would prevent the total group from being a representative national sample on a socioeconomic, geographic, and size- and type-of-system basis.
The 1964 and the 1972/73 pupil groups are comparable as to chronological age. For all three groups the median is 6 years 4 months.

In the 1972 and the 1973 testing programs the current edition of the MRT, that on which this study is based, was taken before any of the newly developed experimental tests, so no practice effect upon its scores is involved.

It should be noted that Form A of the tests was used in the standardization program in 1964, the alternate Form B in the 1972 and 1973 programs. The two forms, however, are equivalent as to total score, although certain of the subtests have differences of up to one point of score between the two forms. Adjustments for these differences have been made in the 1964 versus 1972/73 comparisons of subtest averages.

Results

The MRT includes six subtests: Word Meaning, Listening, Matching, Alphabet, Numbers and Copying. The total raw score from 102 items is considered as a measure of overall readiness for formal instruction. Results on this total score for the 1964 and the 1972/1973 pupil groups are shown in Table 1, where a marked similarity between the 1972 and the 1973 results is shown, with the means differing by only 0.26 score points.

The difference of 10-11 points in the average performance of typical national groups across an 8-9 year interval is highly significant; the mean gain of 10.05 amounts to more than half (.57) the S.D. of the standardization group. The median of 66 for the 1972 and the 1973 group
corresponds to a percentile rank of 73 in the 1964 distribution of scores, and is only one point below the third quartile in 1964. And Q₁ for the 1972/73 groups is almost equal to the median (55) of 8/9 years ago.

Table 1. Total Scores on 1964 Edition of Metropolitan Readiness Tests for 1964, 1972 and 1973 Pupil Groups

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>Q₁</th>
<th>Median</th>
<th>Q₃</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>12,225</td>
<td>41</td>
<td>55</td>
<td>67</td>
<td>53.21</td>
<td>17.75</td>
</tr>
<tr>
<td>1972</td>
<td>4,275</td>
<td>54</td>
<td>66</td>
<td>75</td>
<td>63.38</td>
<td>15.94</td>
</tr>
<tr>
<td>1973</td>
<td>6,653</td>
<td>53</td>
<td>66</td>
<td>76</td>
<td>63.12</td>
<td>17.12</td>
</tr>
</tbody>
</table>

Increase, 1964 to 1972/73: 13/12 11 8/9 10.05

An attachment to this paper graphs the cumulative frequencies for the 1964, 1972 and 1973 distributions on a normalized percentile scale. From it one may note the difference between the 1964 and the 1972/73 scores at any point along the percentile scale.

Subtest Differences

It is of interest to know how the total score increase of 10-11 points is distributed over the six skill or knowledge area measures. Table 2 shows, for each subtest, the 1964 to 1972/73 increase in the mean, both absolutely and as related to the S.D. for the 1964 norm group.

Table 2. Comparison of Subtest Means for 1964 and 1972/1973 Groups

<table>
<thead>
<tr>
<th></th>
<th>Word Meaning</th>
<th>Listening</th>
<th>Matching</th>
<th>Alphabet</th>
<th>Numbers</th>
<th>Copying</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average 1972/73</td>
<td>9.31</td>
<td>10.87</td>
<td>9.44</td>
<td>12.95</td>
<td>13.54</td>
<td>7.12</td>
</tr>
<tr>
<td>Gain 1964 to 1972/73</td>
<td>.64</td>
<td>.98</td>
<td>1.95</td>
<td>3.56</td>
<td>2.17</td>
<td>.70</td>
</tr>
<tr>
<td>SD in 1964</td>
<td>3.10</td>
<td>2.85</td>
<td>4.04</td>
<td>4.70</td>
<td>4.70</td>
<td>3.88</td>
</tr>
<tr>
<td>Relative Gain</td>
<td>.21 SD</td>
<td>.34 SD</td>
<td>.48 SD</td>
<td>.76 SD</td>
<td>.46 SD</td>
<td>.18 SD</td>
</tr>
</tbody>
</table>
For an indication of the different amount of gain at three points along the distribution for the several subtests, Table 3 presents $Q_1$, $Md$, and $Q_3$ values, each being given as an actual integral score. It is noted that for Word Meaning and for Listening the increases at $Q_1$, $Md$, and $Q_3$ are the same; for Numbers and for Copying almost so, being $3/2$, 3 and 3 for the Numbers; 1, 1 and 0/1 for Copying. But for Matching and for Alphabet there is a distinct difference across the three quartile points, with the increase in score decreasing from $Q_1$ through $Q_3$. On the Alphabet test of 16 letters the lower fourth of the 1964 population recognized 5 or fewer letters; in 1972 and in 1973, the lower one-fourth of the scores extended up through 12. The score setting off the lower half was 10 in 1964; by 1972/73 it had risen to 15.

**Summary**

The results of this study indicate that today's entering first-graders are decidedly ahead of their 1964 counterparts in readiness development, as measured by a widely used readiness test. Greatest amount of improvement has been made (1) by the lower-scoring pupils and (2) in knowledge of letter names. The large overall gain, its variation along the score scale, and the
differences in gain across the several skill areas would appear to be in line with what is generally known about current pre-school programs. Most often they are directed to the "disadvantaged" child and, also, tend to place emphasis on routine drill to develop certain associative skills, such as name labels for letters and numerals.

Evidence of improvement in readiness for first-grade instruction has been presented. The implications of this change for first-grade curriculum and instruction extend beyond the field of measurement. Within this field, however, is the inevitable question about the effect of this improved beginning-of-year readiness upon end-of-year achievement, a matter now under investigation by the writer.
PERCENTILE SCALE