Macmillan/McGraw-Hill
Science: A Closer Look 2008

The Macmillan/ McGraw-Hill Company is a leader in the educational community known for its development of successful classroom materials based upon a solid foundation of research. The development of Macmillan/McGraw-Hill’s *Science: A Closer Look* continues in that strong tradition. For any science program to be effective, it must provide a consistent instructional format and design that creates a firm foundation upon which student success is built. Macmillan/McGraw-Hill’s *Science: A Closer Look* program was developed following that model based upon current educational research.

Macmillan/McGraw-Hill’s *Science: A Closer Look* program is an elementary science program written to meet the demands and expectations of the National Science Education Standards and address the Benchmarks for Science Literacy. If students are to succeed on their local and state science assessments, they must be prepared in both the content being assessed and in the format of the assessment. By identifying the key benchmarks and developing specific lessons to meet those expectations, Macmillan/McGraw-Hill’s *Science: A Closer Look* can insure content coverage and student success.

This document summarizes our research efforts and is divided into four parts:

1. A bibliography of the research that was reviewed and studied and which provided the foundation for our instructional pedagogy and lesson structure.

2. A summary of our field study data conducted on the prior Macmillan Science series which provided a basis in the development of *Science: A Closer Look*. This verification data supports the instructional methodology and program design.

3. A review of results of the long term implementation studies on how the Macmillan/McGraw-Hill and its support of staff development initiatives translated into teacher and student success as measured by public release of state test data in the St. Louis Public Schools.

4. A review of State Test Score Results from Districts using Macmillan/ McGraw-Hill Science. This program’s effectiveness provided the foundation for the design of *Science: A Closer Look*
Part I

Bibliography of Research Sources

The following citations were reviewed by our author and editors as part of the development of Macmillan/McGraw-Hill’s *Science: A Closer Look*. The academic studies, educational research reports, and professional articles listed below helped to inform and guide the development process. These studies form the underpinnings of our instructional methodology and learning pedagogy.


Appleton, Ken and Ian Kindt. *How do beginning elementary teachers cope with science: Development of pedagogical content knowledge in science*. 


Blosser, Patricia E. “Science Misconceptions Research and Some Implications for the Teaching of Science to Elementary School Students.” *ERIC/SMEAC Science Education Digest* No. 1, 1987.


Martin, Michael O., Ina V.S. Mullis, Eugenio Gonzalez, Kelvin Gregory, Teresa Smith, Steven Chrostowski, Robert Garden, and Kathleen


Miller, G.A. (1956). The magical number seven, plus or minus two. Some limits on our capacity to process information. Psychological Review, 63, 81-87.


Part II.

Executive Summary

Research and Field Study Results

The following section provides a summary of the field test site results that were gathered as part of the validation on the effectiveness of the instructional model utilized in Science: A Closer Look. These studies were conducted by an independent educational research firm and the results were provided to Macmillan/McGraw-Hill publishing.

Field Studies:

Background

The field studies measured the effectiveness of the instructional model used in an elementary science program published by the McGraw-Hill School Division.

Teachers who participated in this study administered grade appropriate chapter pre tests prior to teaching chapters from Macmillan/McGraw-Hill Science. Post-tests were administered after using the materials in an effort to assess student progress. The results of these tests are included in this report.

Throughout this report, the primary measure of student performance is the “Gap Reduction Percentages” or “GRPs”. GRP reflects the degree to which students have succeeded in closing the gap between the average pre-test score and a perfect score, as reflected by the post-test. Specifically, GRPs are calculated using the following formula:

\[
\text{GRP} = \frac{\text{Average post-test score} - \text{Average pre-test score}}{\text{Average pre-test score}}
\]
A GRP of 0% means that student performance did not improve from pre-test to post-test. A GRP of 50% means that students have closed half the gap between the average pre-test score and a perfect score. For example, an average pre-test score of 50% that is followed by an average post-test score of 75% yields a GRP of 50%; in other words, the gap between the average pre-test score and a perfect score closed by half. Of course, a GRP of 100% means that the gap between the average pre-test score and a perfect score has been eliminated.

The GRP was formulated to measure performance because percentages change, a more typical measure, is unduly influenced by a pre-test score. For example, a post-test score of 90% yields a percentage change of only 12 percent if the pre-test score was 80. By contrast, a post-test score of 40% yields a percentage change of 33% if the pre-test score was 30%. In these examples, a ten percentage point increase yields a very different and potentially misleading percentage change figures. The GRP attempts to eliminate this variability which occurs because of high pre-test scores minimize the possibility of significant percentage changes, while low pre-test scores all but ensure them.

Consider another example, average pre- and post-test scores of 20% and 80%, respectively, yield a percentage change of 300 percent and a GRP of 75 percent. Average pre- and post-test scores of 60% and 90% respectively, yield a much lower percentage change of 50 percent, but the same GRP of 75 percent. The latter is true because in both cases, the gap between average pre-test score and a perfect score closed by three-quarters.

Results:
In total, 2,913 students completed pre- and post-test: 232 students in seven Grade 1 classrooms; 868 students in ten Grade 2 classrooms; 598 students in eight Grade 3 classrooms; 648 students in Grade 4 classrooms; 455 students in seven Grade 5 classrooms; and 112 students in two Grade 6 classrooms.

### Field Test Parameters

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number of Classrooms</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>232</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>868</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>598</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>648</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>455</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>112</td>
</tr>
</tbody>
</table>

Seven public schools participated in this project: three in urban communities, two in suburban communities; and two in rural communities. The schools were located in six states: California, Georgia, Illinois, Missouri, New York, and North Carolina.

Of the 2,913 students who participated, 48% were boys, 52% were girls; 26% are minorities; and 14% receive free or reduced-cost lunch.
Top Line Results

- Research results indicate that test scores increased among students using MMH Science
- Eight out of ten students earned higher scores after using MMH Science
- Scores improved among both boys and girls
- Scores improved among minority and non-minority students
- Scores improved among students who reviewed free/reduced-cost lunch
- Overall the gap between average pre-test score and perfect score closed by 46%, or stated differently, on average, scores increased by 59% after using MMH Science.

Test Comparisons: Grade Level Pre-/Post- Test

Grade Level Pre-/Post-Test Score Comparisons

Individual Grade Results

Grade 1 Results

Two hundred thirty-two students in seven Grade 1 classrooms completed pre-post tests. Students were tested on Chapters 1 through 9

An analysis was performed on the pre-and post-tests scores to determine the magnitude of the gap that was closed from the pre-test to the post-test; that is, how far from a perfect score
students were before instructions compared to how far they were after instruction. The results from each classroom are listed in individual tables.

Grade 1 Pre-/Post- Test Scores

<table>
<thead>
<tr>
<th>Classrooms</th>
<th>Average Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>MV1-1</td>
<td>80</td>
</tr>
<tr>
<td>PK1-1</td>
<td>70</td>
</tr>
<tr>
<td>JW1-1</td>
<td>60</td>
</tr>
<tr>
<td>JW2-1</td>
<td>50</td>
</tr>
<tr>
<td>JW3-1</td>
<td>70</td>
</tr>
<tr>
<td>FH1-1</td>
<td>80</td>
</tr>
<tr>
<td>BH1-1</td>
<td>70</td>
</tr>
</tbody>
</table>

Grade 2 Results

Eight hundred sixty-eight students in ten Grade 2 classrooms completed pre-post tests. Students were tested on Chapters 1 through 12.

An analysis was performed on the pre-and post- tests scores to determine the magnitude of the gap that was closed from the pre-test to the post-test; that is, how far from a perfect score students were before instructions compared to how far they were after instruction. The results from each classroom are listed below. In one class, the gap between the average pre-test score and a perfect score closed by 68 percent.

Grade 2 Pre- Post- Test Scores

<table>
<thead>
<tr>
<th>Schools</th>
<th>Average Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>JW1-2</td>
<td>80</td>
</tr>
<tr>
<td>JW2-2</td>
<td>70</td>
</tr>
<tr>
<td>JW3-2</td>
<td>60</td>
</tr>
<tr>
<td>MV1-2</td>
<td>50</td>
</tr>
<tr>
<td>PK1-2</td>
<td>70</td>
</tr>
<tr>
<td>BH1-2</td>
<td>80</td>
</tr>
<tr>
<td>JE1-2</td>
<td>70</td>
</tr>
<tr>
<td>TJ1-2</td>
<td>60</td>
</tr>
<tr>
<td>TJ2-2</td>
<td>50</td>
</tr>
<tr>
<td>TJ3-2</td>
<td>40</td>
</tr>
</tbody>
</table>
Grade 3 Results

Five hundred ninety-eight students in eight Grade 3 classrooms completed pre-post tests. Students were tested on Chapters 1, 2, 7 through 12 and 14.

An analysis was performed on the pre-and post-tests scores to determine the magnitude of the gap that was closed from the pre-test to the post-test; that is, how far from a perfect score students were before instructions compared to how far they were after instruction. The results from each classroom are listed below. In one class, the gap between the average pre-test score and a perfect score closed by 65 percent; in another class, the gap closed by 63 percent; and in a third class, the gap closed by 62 percent.

Grade 4 Results

Six hundred forty-eight students in thirteen Grade 4 classrooms completed pre-post tests. Students were tested on Chapters 1 through 9 and 14.

An analysis was performed on the pre-and post-tests scores to determine the magnitude of the gap that was closed from the pre-test to the post-test; that is, how far from a perfect score students were before instructions compared to how far they were after instruction. The results from each classroom are listed below. In one class, the gap between the average pre-test score and a perfect score closed by 83 percent; in another class, the gap closed by 73 percent.
Grade 5 Results

Four hundred fifty-five students in seven Grade 5 classrooms completed pre-post tests. Students were tested on Chapters 2 through 4, 6, 8 through 11, and 13.

An analysis was performed on the pre-and post-tests scores to determine the magnitude of the gap that was closed from the pre-test to the post-test; that is, how far from a perfect score students were before instructions compared to how far they were after instruction. The results from each classroom are listed below. In one class, the gap between the average pre-test score and a perfect score closed by 62 percent; in another class, the gap closed by 58 percent.
Demographic Results

Test Scores and Race

Improveing from pre-test to post- test for minority and non-minority students was tested. The results are listed for minority students are below.

![Minority Pre- / Post- Test Results](image)

**Test Scores and Gender**

Improvement from pre-test to post- test for boys and girls were tested. The results are listed for gender in Table 2.

![Gender Pre-/Post- Test Scores](image)
Test Scores and Free or Reduced-cost Lunch

Improvement from pre-test to post-test for students who received free or reduced-cost lunch was tested. The results are listed for free or reduced-cost lunch in the table below.
Part III.

Efficacy Studies

The following section provides a review of an independent study conducted in the St. Louis Public School system. The results show that with the implementation of Macmillan/McGraw-Hill Science student achievement in science improved as measured by the Missouri Assessment Program (MAP). An additional benefit was noted as a result of this initial study. Student reading achievement also improved as a result of increase science instructional time. Additional supporting studies are also included.

St. Louis Study

St. Louis Public Schools
Higher Scores on the Latest Missouri Assessment Program (MAP)
Continue Multi-year Improvement in Elementary Schools;
Science Shows Biggest Improvement of Any Subject

Overview

The results of the latest scores on the Missouri Assessment Program (MAP) tests show that elementary students in the St. Louis Public School district continue to “catch up” to the Missouri State average in science, communication arts, and mathematics. MAP test scores* of students in SLPS elementary schools have improved at such a fast rate over the past five years that the District will actually meet the state average in all three subjects by next year or the following year just by continuing to improve at this rate.

Data

Figures 1 through 3, for Communication Arts, Science and Mathematics, respectively, illustrate just how dramatic this improvement has been. In all three subjects, students in the tested elementary grades of SLPS have improved more than the state average for the past five years, with science achievement being the most improved from 2000 – 2004 followed in order by communication arts and mathematics. The degree of improvement is quantified in Figure 4. Figure 4 shows that elementary students in SLPS improved much more than their
counterparts in the rest of Missouri, and also improved much more than students in the middle and high schools of SLPS.

Conclusions/Implications

An Urban Systemic Initiative from the National Science Foundation (NSF) to improve Mathematics and Science has been in place in SLPS since 1998-9. Although intended to improve student achievement in K through 12, objective improvement has been confined almost exclusively to the elementary grades (the one exception being a small improvement in middle school mathematics). Similar patterns of improvements being limited to the elementary level have also been reported by other NSF Urban Systemic Initiative sites. These findings suggest that systemic reform of curriculum and teaching in middle and high schools may be too late to provide much help for students already lacking in prerequisite fundamental knowledge, skills and abilities that should have been acquired in previous grades. This hypothesis would be further strengthened in the coming years if the higher-achieving students now emerging from those elementary schools improved by systemic reform practices continue to achieve at higher in middle and high schools.

As indicated by Figure 4, the greatest degree of improvement was seen in elementary school science, with an improvement on MAP of 20.1% from 2000 to 2004. And most of this improvement occurred during Test Years 2002, 2003, and 2004, years that correspond to the adoption of new textual materials in science from McMillan/McGraw-Hill.

Figure 1: Communication Arts

![Figure 1: Communication Arts](image-url)
Figure 2: Science

Grade 3 Test Scores on the Missouri Assessment Program (MAP):
St. Louis Public Schools vs. Missouri Average for Last Five Years

<table>
<thead>
<tr>
<th>Test Year</th>
<th>Missouri</th>
<th>SLPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>215.5</td>
<td>182.0</td>
</tr>
<tr>
<td>2001</td>
<td>216.8</td>
<td>186.0</td>
</tr>
<tr>
<td>2002</td>
<td>218.7</td>
<td>197.7</td>
</tr>
<tr>
<td>2003</td>
<td>220.0</td>
<td>208.0</td>
</tr>
<tr>
<td>2004</td>
<td>224.4</td>
<td>218.5</td>
</tr>
</tbody>
</table>

Figure 3: Mathematics

Grade 4 Test Scores on the Missouri Assessment Program:
St. Louis Public Schools vs. Missouri Average for Last Five Years

<table>
<thead>
<tr>
<th>Test Year</th>
<th>Missouri</th>
<th>SLPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>209.7</td>
<td>182.9</td>
</tr>
<tr>
<td>2001</td>
<td>211.4</td>
<td>185.8</td>
</tr>
<tr>
<td>2002</td>
<td>210.7</td>
<td>186.2</td>
</tr>
<tr>
<td>2003</td>
<td>210.4</td>
<td>193.7</td>
</tr>
<tr>
<td>2004</td>
<td>214.4</td>
<td>208.7</td>
</tr>
</tbody>
</table>
Figure 4:
Percent Improvement in MAP Test Index Scores from 2000 to 2004 by Grade and Subject
St. Louis Public Schools (SLPS) Compared to the Missouri Average

<table>
<thead>
<tr>
<th>Grade</th>
<th>Subject</th>
<th>District</th>
<th>Year 2000</th>
<th>Year 2004</th>
<th>% Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three/Four</td>
<td>Comm. Arts</td>
<td>SLPS</td>
<td>168.3</td>
<td>194.1</td>
<td>15.3</td>
</tr>
<tr>
<td>Three/Four</td>
<td>Science</td>
<td>SLPS</td>
<td>182.0</td>
<td>218.5</td>
<td>20.1</td>
</tr>
<tr>
<td>Three/Four</td>
<td>Mathematics</td>
<td>SLPS</td>
<td>182.9</td>
<td>208.7</td>
<td>14.1</td>
</tr>
<tr>
<td>Three/Four</td>
<td>Comm. Arts</td>
<td>Missouri</td>
<td>197.2</td>
<td>201.9</td>
<td>2.4</td>
</tr>
<tr>
<td>Three/Four</td>
<td>Science</td>
<td>Missouri</td>
<td>215.5</td>
<td>224.4</td>
<td>4.1</td>
</tr>
<tr>
<td>Three/Four</td>
<td>Mathematics</td>
<td>Missouri</td>
<td>209.7</td>
<td>214.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Comm. Arts</td>
<td>SLPS</td>
<td>154.8</td>
<td>153.9</td>
<td>-0.6</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Science</td>
<td>SLPS</td>
<td>136.5</td>
<td>134.9</td>
<td>-1.2</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Mathematics</td>
<td>SLPS</td>
<td>134.4</td>
<td>144.3</td>
<td>7.4</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Comm. Arts</td>
<td>Missouri</td>
<td>190.8</td>
<td>191.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Science</td>
<td>Missouri</td>
<td>169.3</td>
<td>168.6</td>
<td>-0.4</td>
</tr>
<tr>
<td>Seven/Eight</td>
<td>Mathematics</td>
<td>Missouri</td>
<td>167.6</td>
<td>173.4</td>
<td>3.5</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Comm. Arts</td>
<td>SLPS</td>
<td>154.2</td>
<td>142.5</td>
<td>-7.6</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Science</td>
<td>SLPS</td>
<td>135.4</td>
<td>127.1</td>
<td>-6.1</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Mathematics</td>
<td>SLPS</td>
<td>135.8</td>
<td>129.3</td>
<td>-4.8</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Comm. Arts</td>
<td>Missouri</td>
<td>182.9</td>
<td>185.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Science</td>
<td>Missouri</td>
<td>166.2</td>
<td>167.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Ten/Eleven</td>
<td>Mathematics</td>
<td>Missouri</td>
<td>162.2</td>
<td>171.1</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*NOTE: MAP scores are presented as Index Scores ranging from 100 to 300. Scores were calculated by the formula developed by the Missouri Department of Elementary and Secondary Education (DESE) for purposes of comparing scores among different demographic groups and among different schools.

Science & Reading Improvement in Grade 3
by St. Louis Public Schools (SLPS):
MAP Test Results Currently (2004) and Four Years Ago (2000)
Compared by Students’ Race and to the Missouri (MO) Average

<table>
<thead>
<tr>
<th>Subject</th>
<th>MO/All</th>
<th>SLPS/Black</th>
<th>SLPS/White</th>
</tr>
</thead>
<tbody>
<tr>
<td>CommArts</td>
<td>31.7</td>
<td>34.6</td>
<td>29.1</td>
</tr>
<tr>
<td>Science</td>
<td>44.9</td>
<td>51.2</td>
<td>43.0</td>
</tr>
</tbody>
</table>

District/Student Race, Subject, and Test Year
Section 2:

Supporting Study: University City, MO Study

University City, MO. University City is an “inner” suburb of St. Louis. Being an “inner” suburb, it has about the same urban demographics as the St. Louis City Schools, at about 80% African American students, but it is only about 10% the size of St. Louis City school district. For instance, it tested just under 350 students in Grade 3 Science on the state MAP test in 2004, compared to about 3500 who took the test in St. Louis City public schools. 2003-4 was the first year of their adoption of *Macmillan/McGraw-Hill Science*, and it was implemented by a new staff person who had previously worked (through Washington University) providing professional development in science to teachers in St. Louis City schools. You can see the results of the new texts and new emphasis on Science in the elementary schools: the District’s Grade 3 students shot up on the 2004 MAP test in Science to just a few score units below the state average. Data from following years of testing will be required in order to determine whether this improvement constitutes a continuing trend or not.
Part IV.

State Test Score Results
(Science 2005 Edition)

The following section provides a summary of the on-going student achievement results being gathered as part of a long-term program validation study. These studies are being conducted by independent researchers whose purpose is to study the effectiveness of the Macmillan/McGraw-Hill Science instructional model in actual classroom use. The studies are designed to measure the effectiveness of the basal science program by comparing the results of published state science test data. Identification of scores prior to the implementation of Macmillan/McGraw-Hill Science are compared to the state science test results after implementation. These results are on-going in nature and are being revised on an annual basis.

State Studies include:
Georgia
Implemented Macmillan/McGraw-Hill Science in 2002. Georgia State Science Test is the CRCT
Georgia-Dekalb- Grade 3 Science

Most Improved Schools in District

Georgia-Dekalb- Grade 4 Science

Most Improved Schools in District
State Studies include: Illinois
State Studies include: Lafayette Parish, Louisiana
Louisiana State Science Test is the LEAP 21

**Lafayette, Louisiana Grade 4 Science Proficiency**

![Bar chart showing Lafayette Parish elementary schools' science proficiency by year.]

State Studies include: Virginia Beach, Virginia
Proficiency Study- Demographic Data
Virginia State Science Test is the SOL

Virginia Beach student scores compared to State Average after one year of use.

**VIRGINIA BEACH SOL 2003-2004**

<table>
<thead>
<tr>
<th>GRADE 5</th>
<th>School</th>
<th>Division</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Students</td>
<td>91</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>Female</td>
<td>88</td>
<td>88</td>
<td>83</td>
</tr>
<tr>
<td>Male</td>
<td>93</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>African American</td>
<td>93</td>
<td>77</td>
<td>70</td>
</tr>
</tbody>
</table>
Virginia Beach, VA demographic comparisons to state average after one year of using Macmillan/McGraw-Hill Science.

<table>
<thead>
<tr>
<th></th>
<th>Division</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Students</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>African American</td>
<td>77</td>
<td>70</td>
</tr>
<tr>
<td>American Indian</td>
<td>100</td>
<td>87</td>
</tr>
<tr>
<td>Hispanic</td>
<td>86</td>
<td>73</td>
</tr>
<tr>
<td>Limited English</td>
<td>84</td>
<td>68</td>
</tr>
</tbody>
</table>
State Studies include: Texas
Comparisons of Comparable Districts: Macmillan/McGraw-Hill
Science to Leading Competitor
Texas State Test: TAKS

TEXAS ASSESSMENT OF KNOWLEDGE AND SKILLS (TAKS)- Science
STATE/DISTRICT PROFICIENCY RESULTS 2004

<table>
<thead>
<tr>
<th>District-Competitor</th>
<th>% Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amarillo-Lubbock</td>
<td>76</td>
</tr>
<tr>
<td>Houston-Dallas</td>
<td>61</td>
</tr>
</tbody>
</table>

Texas District Comparison Chart- 2004
Grade 5 Science

Macmillan/McGraw Science District (blue) versus Leading Competitive District (purple)