Implementation of a Reformed Mathematics Curriculum: Oregon Achievement Results

by Camille L. Wainwright, PhD, Pacific University and Jeffrey Frykholm, University of Colorado

In research on some of the first large-scale implementations of reform curricula, data indicate that reform students do as well on skills as students who study the traditional curricula, and that they do better on an understanding of concepts and problem solving. Moreover, traditional performance gaps between majority students and poor or under-represented minorities are diminished, though not eliminated.[1]

AS A RESULT OF the impact of the *NCTM Curriculum and Evaluation Standards*,[2] *Standards 2000*,[3] and numerous other calls for reform (e.g., NRC *Everybody Counts*[4]), numerous curriculum projects have flowered across the country, largely through funding from the National Science Foundation. While the content and strategies among these "reform" curricula vary, they tend to have several characteristics in common: a constructivist approach to learning, emphasis on conceptual understanding and process over content; mathematical reasoning, problem solving and communication; heightened student engagement; and issues of equitable access to mathematical literacy for all students.

Data on the effectiveness of these curricula has only recently begun to become available, since development and refinement of these projects has taken nearly a decade and students have been in the pipeline of these multi-year programs. Summarizing the growing body of research on the effectiveness of these curricula, Alan Schoenfeld states:

1. On tests of basic skills, there are no significant performance differences between students who learn from traditional or reform curricula.

2. On tests of conceptual understanding and problem solving, students who learn from reform curricula consistently outperform by a wide margin students who learn from traditional curricula.

3. There is some encouraging evidence that reform curricula can narrow the performance gap between whites and under-represented minorities.[5]

As compelling as these initial findings may be, however, these studies were conducted in various schools across the country — Pittsburgh, Massachusetts, Michigan, etc. Closer to home, what has been the effect of reform curricula in Oregon, as measured by the Oregon Statewide Mathematics Mssessment? Are the results similarly positive?

Implementation of a Reform Curriculum in Oregon: IMP

Across Oregon, perhaps the most highly implemented reform curriculum has been the Interactive Mathematics Project (IMP), which has been adopted by dozens of school districts. IMP's four-year high school program of problem-based mathematics replaces the traditional Algebra I-Geometry-Algebra II/Trigonometry-Precalculus sequence. Through a 'spiral' approach, IMP integrates the traditional areas of mathematics with newer topics such as probability, statistics, discrete mathematics, and matrix algebra – including those that use graphing calculators and computers. IMP encourages cooperative learning and is designed to be accessible to all students regardless of ability or career goals. The IMP curriculum challenges students to actively explore open-ended situations in an inquiry manner that resembles the methods used by mathematicians and scientists in their work. Paying tribute to these characteristics of IMP, an article in the *Mathematics Teacher* provides a detailed overview of the Year 1 IMP program.[6] (For additional descriptions of the IMP curriculum, see the informational website at http://www.mathimp.org)

Given the differences in both content and structure of IMP when compared to more traditional curricula, it necessarily follows that the teaching strategies and pedagogical decisions necessary to teach IMP well tend to differ dramatically from the traditional preparation of most mathematics teachers. Even with extensive support from colleagues and staff, the shift to a different style of teaching can be daunting. As one new Oregon teacher in this study admitted, "The idea of letting a classroom full of students run loose and discover their way into conjectures and rules seemed truly frightening."[7] Consequently, a well-designed program of professional development is a critical component of the adoption and implementation processes for any curricular innovation, and IMP in particular. In the midst of a year-long professional development program designed to bolster teachers' confidence and know-how with IMP, she would later affirm the structure and strategies of the program as well as her contagious enthusiasm for it. As she reflected, "Since having taught the IMP curriculum for six months, I have found my level of job satisfaction has greatly increased. I look forward to the next day and can hardly wait to share the new, fresh material with my students. They constantly tell me how much they enjoy the class. The students' positive attitudes toward the class truly make it a joy to teach. I have also learned to allow for independent learning. I have found that I do not have to tell them how to perform certain algorithms or what approach to use. These students are finding excitement in realizing a new concept all by themselves for the first time. It makes them feel smart to find something out or discover a property all on their own (or with the help of a group)."[8]

IMP and the Oregon Statewide Assessment

As noted earlier, results from numerous national studies show IMP is associated with higher student achievement levels regardless of the initial ability level of the student, as compared to their traditionally taught peers. That is, high-achieving IMP students do better than their high-achieving, traditionally taught peers, while students with weak math skills do better than their traditionally taught counterparts.[9] But can we document similar results in Oregon? In this brief article, we share highlights of a large data set that begin to suggest that these trends also hold in Oregon.

The data reported in this study are drawn from students in a large suburban Oregon school district. They include results from three tests — a multiple-choice (MC) test administered to both the IMP and control (TRAD) groups once in 8th grade, and again at the 10th grade level, and a problem solving (PS) test administered to both groups at the 10th grade level. All three tests were components of the Oregon Statewide Mathematics Assessment process. Differentiating the two groups by the 10th grade level were differences in the course and curricular programs in which the two groups engaged. Students in the IMP group had at least one year of coursework in one of the IMP texts between 8th and 10th grade, while the TRAD groups followed a non-IMP, more traditional path through pre-algebra, algebra, and geometry. It should also be noted that, in terms of cumulative class-time (hours spend in mathematics class), the control group received 50% more classroom instructional time than did the IMP group due to various block scheduling arrangements. This part of the context of the study is particularly relevant when one views the comparison of test scores that follows below – the IMP students appear to be scoring equally or better than their counterparts with significantly less instructional time in the classroom.

The Multiple Choice Test

Given these two groups, and the fact that each group was tested at both the 8th and 10th grade levels, the first task was to determine that they were equally distributed (in terms of mathematical ability) at the beginning of 8th grade. Hence, data from the two 8th grade multiple choice tests were used to compare the IMP group to the control (TRAD) group. A *t-Test* for independent samples (since the size of the groups was not equivalent) was used to compare the means of both groups on the PS and MC tests. As indicated in Table 1 below, there was no significant difference between the mean scores of each group on both the problem solving and multiple choice tests. In summary, the IMP group scored slightly below the control group at the 8th grade level on the MC test.

Once it was established that the two groups were in fact equal in 8th grade, then it was possible (using the 10th grade data for both tests) to determine if the IMP group had improved at a greater rate than the control (TRAD) group. Two statistical measures were applied to test this hypothesis. First, the same statistical measure described in the previous paragraph (t-Test for independent samples) was also applied to the 10th grade data (for both the MC and PS tests) in similar fashion. The results of this statistical test are found in Table 2 (next page).

As Table 2 suggests, there appears to be little statistical difference in the mean scores of IMP and TRAD students on both the MC and PS achievement tests at the 10th grade level. In both cases (PS and MC tests), the means for the IMP students were greater than those of the TRAD students. However, the t-Test results did not suggest significant differences at the $p < .05$ level.

8th MC Test	
IMP Average	232.93
TRAD Average	233.87
IMP Standard Deviation	7.643
TRAD Standard Deviation	7.537
t-Test Value	−0.549
Significant at p < .05	No

Table 1: 8th Grade Comparisons of Mean Scores on PS and MC Tests

CONTINUED→

Despite the lack of a statistically significant difference when comparing mean scores, however, another test of significance was applied to the data to evaluate the classification of student proficiencies at the grade levels. A Chi Square test (used for evaluating frequency distributions) was used to determine if equivalent numbers of students in both groups were passing the tests at similar levels.

10th PS Test		10th MC Test	
IMP Average	28.26	IMP Average	240.42
TRAD Average	27.05	TRAD Average	239.73
IMP Standard Deviation	9.19	IMP Standard Deviation	9.10
TRAD Standard Deviation	9.21	TRAD Standard Deviation	6.09
t-Test Value	.582	t-Test Value	.431
Significant at $p < .05$	No	Significant at $p < .05$	No

Table 2: 10th Grade Comparisons of Mean Scores on PS and MC Tests

Categories include: 1) Exceeds Expectations, 2) Meets Expectations 3) Conditionally Meets Expectations, and 4) Does Not Meet Expectations.

Although the mean scores tell us something about the central tendencies of scores across the samples, they do not tell us much about whether or not the distribution of scores was the same in both groups. Hence, a Chi-Square test was administered to test for significantly different distributions of student proficiencies. Contrary to the previous results, the Chi Square test did reveal significant differences between the frequencies at which students in both groups surpassed expectations on the MC test.

Specifically, a Chi Square test measures the observed frequency distributions of test results (in this, case proficiency levels), compared to what might have been expected if there were no difference between both the control and test groups. Hence, the frequency distributions of the IMP students (10th grade) across the proficiency levels were compared to what would have been expected for those categories based on the percentages of TRAD students in those categories. That is, the TRAD results were used to determine expected frequencies for the IMP group. For example, if 15% of the TRAD students were determined to "exceed expectations" for proficiency, one could argue that roughly 15% (accounting for small differences of chance) of the IMP students would also "exceed expectations" if in fact the groups were similar in ability. As noted above, for both the MC and PS tests, the IMP students were categorized into proficiency levels at significantly higher proportions than the TRAD students. These findings are illustrated in Table 3.

The findings summarized in Table 3 indicate that, for both the MC and PS tests, the IMP strand of courses tended to promote more students (in terms of percentages) into acceptable proficiency categories. This was particularly true for the MC test. The test results suggest that less than ten percent of the time, this difference in score distribution could be attributed to chance. For the PS test, although the results were not significant at $p < .10$, they were significant at $p < .20$. That is, less than twenty percent of the time, this difference in score distribution could be attributed to chance.

10th Grade MC Frequency Distribution

	Exceeds	Conditionally Meets	Does Not Meet	N
IMP	9 25%	11 30.56%	16 44.44%	36
TRAD	13 10.57%	57 46.34%	52 42.28%	122
Chi Square Results	Chi Square Value = 5.817 Degrees of Freedom: 2 Significant Difference at $p < 0.10$			

10th Grade PS Frequency Distribution

	Exceeds	Meets	Conditional	Does Not Meet	N
IMP	1 4.76%	10 47.62%	3 14.29%	8 38.10%	22
TRAD	5 3.31%	67 44.37%	5 3.31%	72 47.68%	149
Chi Square Results	Chi Square Value = 4.995 Degrees of Freedom: 3 Significant Difference at $p < 0.20$				

Table 3: Frequency Distribution for Proficiency Ratings on MC and PS Tests

Discussion

There is much that can be said about the student data gathered in connection with the PD-TEAMS III project. Both tests for mean and categorical differences suggested that, at the least, IMP students performed as well as their peers in a traditional curricular program. A similar result has been reported by numerous school districts across the state. Given the low numbers of students enrolled in IMP classes, illustrating statistical significance through the use of any test is more challenging than if comparable percentage differences occurred with equal sized comparison groups. Yet, what is most notable about this study is that, as mentioned earlier, the IMP students out-performed their peers in the control group even though the control group had 50% more instructional time in class. One wonders how much greater the difference in scores would have been if the IMP students had received an amount of instructional time equal to that which the control group received.

In addition to the quantitative analysis of student achievement data, this study also looked at the impact the professional development opportunities associated with IMP had on the participating teachers. Although there is not enough space in this article to elaborate more fully, survey data (from *The Mathematics Teacher Belief Inventory*)[10] and interview data indicated significant growth among IMP teachers in all areas of reform teaching, most notably in "Construction of Knowledge and Learning," and "Depth of Knowledge and Mathematics." These two categories are mentioned in particular given their importance as mediators in mathematics education reform. The research literature on teachers' beliefs is replete with evidence that, unless teachers undergo changes in these areas (with respect to their beliefs and knowledge bases), meaningful, long-term change is unlikely to occur. The literature also indicates the difficulty in changing such beliefs; hence these impressive gains are positive indicators of the success of the IMP professional development program.

Conclusions:

We have an obligation to provide students with the best possible opportunities to gain mathematical literacy. All of our students, regardless of their economic, cultural, or linguistic backgrounds deserve the opportunity to learn powerful mathematics, since "... making decisions in one's personal life, on the job, and in matters of public interest calls increasingly for quantitatively sophisticated reasoning. More than ever before, today's students need to learn to reason and communicate using mathematical ideas." [11]

In closing, notable about the findings in this study is the way in which teachers are growing over time with respect to their validations of IMP institutes in which they have participated. Moreover, it appears to be the case that gains in student achievement are growing in parallel. That is, although there were no significant differences in student test scores in the previous iteration of this study (2000-01 data), there were statistically significant gains this year. Taken together, the data represented in this report suggests that the IMP curriculum is in fact an increasingly valuable asset in helping learners exceed Oregon Content Standards and Benchmarks. These results suggest the need for wider dissemination of programs like IMP that hold the potential to impact the pedagogical preparation and practices of mathematics teachers, as well as the proficiency and achievement of learners in mathematics classrooms.

Acknowledgements

This study was partially funded by a Dwight D. Eisenhower grant.

The authors appreciate the support and input from the NW IMP leaders: Brent McClain (Tigard-Tualatin School District), Sue Yabuki (Portland Public Schools) and Lori Carter (Hillsboro School District).

[1] Schoenfeld, A. H. (2002) "Making mathematics work for all children: Issues of standards, testing, and equity." Educational Researcher, 31: 1, p. 14.
[2] National Council of Teachers of Mathematics (1989) *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: Author.
[3] National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.
[4] National Research Council. (1989). *Everybody Counts: A report to the nation on the future of mathematics education*. Washington, DC: Author.
[5] Schoenfeld, A. H. (2002) "Making mathematics work for all children: Issues of standards, testing, and equity." Educational Researcher, 31: 1, p. 16.
[6] NCTM (1997). *The Mathematics Teacher.* September, 501-502.
[7] Barker, Nicole. (2002) "Regional IMPressions," IMPressions: A Newsletter about the Interactive Mathematics Program. Key Curriculum Press.
[8] Ibid
[9] http://www.mathimp.org/research/index.html
[10] Brendefur (1999). The Mathematics Teacher Belief Inventory.
[11] Schoenfeld, A. H. (2002) "Making mathematics work for all children: Issues of standards, testing, and equity." Educational Researcher, 31: 1, p. 13.