

RESEARCH RESULTS

Into Math *Quasi-Experimental Study:* *Examination of 2023 Illinois Math Test Scores*

STUDY PROFILE

Sample:

155 schools in Illinois

GRADES:

3–8

STUDY DESIGN:

Quasi-Experimental Design (ESSA Level 2)¹

EVALUATION PERIOD:

2022–2023 school year

STUDY CONDUCTED BY:

JEM & R LLC

OUTCOME MEASURE:

Illinois Assessment of Readiness (IAR)

SCHOOL CHARACTERISTICS

This study included a total of 46 Illinois schools that purchased and adopted HMH *Into Math*® starting in the 2020–21 school year. Researchers performed confirmation calls with schools/districts to determine the extent to which identified schools have used *Into Math*. These confirmation calls allowed researchers to determine that potential treatment schools were established *Into Math* users.

IMPLEMENTATION OVERVIEW

Given the nature of the study, researchers relied on self-report measures from confirmation calls to determine usage of *Into Math* as their core mathematics program at Grades 3–8. Data collected included: (1) verification of use of the *Into Math* program and at which grades, and (2) the proportion of students within schools that used this curriculum.

Into Math is an innovative mathematics program for Grades K–8 that focuses on meeting the needs of the whole child and strives to instill in students a positive attitude toward math. The solution emphasizes the development of a conceptual understanding to allow students to transfer their knowledge to new situations and apply it to new contexts. The lessons are intentionally structured to ensure that procedural practice functions to reinforce conceptual knowledge. At all *Into Math* sites, a majority of students at each grade level were taught with the program as their core math instruction.

PARTICIPANTS

The total study consisted of 155 schools with 46 schools using HMH *Into Math* during the study period, and 109 schools serving as the matched sample comparison group.

Given that performance data is available for each grade level, the unit of analyses is *grade level performance* within schools. Therefore, *Into Math* schools were matched by school-level characteristics and grade-level performance measures to unique schools within the state.

Table 1 below displays the final grade level units for treatment and comparison schools included in the final analytic sample.

Table 1 Analytic Sample Size by Grade and Cohort							
Illinois	Grade Level						Total
Group	3	4	5	6	7	8	
Total	52	37	38	39	36	31	233
Control	25	18	18	20	19	15	115
<i>Into Math</i>	27	19	20	19	17	16	118



MEASURES

ILLINOIS ASSESSMENT OF READINESS (IAR)

The Illinois Assessment of Readiness (IAR) is the state assessment and accountability measure for Illinois students enrolled in a public school district. IAR has been in use since 2019 and assesses the New Illinois Learning Standards Incorporating the Common Core in English Language Arts and Mathematics for all students in Grades 3–8. The IAR requires students to demonstrate their acquired skills and knowledge by answering selected-response items and fill-in-the-blank questions. Each assessment consists of multiple units, and one of the mathematics units is split into two sections: a non-calculator section and a calculator section.

The present study uses math assessment data from the IAR. According to the IL Department of Education, the mathematics assessments contain tasks that measure a combination of conceptual understanding, applications, skills, and procedures. Mathematics constructed-response items consist of tasks designed to assess a student's ability to use mathematics to solve real-life problems. Some of the tasks require students to describe how they solved a problem, while other tasks measure conceptual understanding and ability to apply concepts by means of selected response or technology-enhanced items. In addition, students must demonstrate their skills and knowledge by answering innovative selected-response and short-answer questions that measure concepts and skills.

The tests were designed to allow the classification of student proficiency into five performance levels (Level 1, Level 2, Level 3, Level 4, and Level 5). For the present study, percent proficient (i.e., Levels 4 and 5) is the student achievement measure.

PROPENSITY MATCH PROCEDURES

To evaluate the impact of *Into Math*, a comparison group was identified using Propensity Score Matching (PSM). PSM was conducted separately for each grade level within school. To create these comparison groups, researchers conducted a three-step process:

Step 1. Computing Propensity Scores

For each unit included in a model, the probability of receiving the *Into Math* treatment was estimated using a logistic regression model with the *Into Math* treatment flag as the outcome variable and grade-level characteristics as predictor variables.

Step 2. Matching

PSM was implemented with MatchIt version 4.5.0 (Ho et al., 2011) using the nearest neighbor matching method with generalized linear model specified as the distance measure and logit specified as the link function. In each PSM model, each *Into Math* unit was sequentially matched with one potential control unit.

Step 3. Assessing Balance

Baseline equivalence was assessed on key characteristics using a two-sided Independent Sample t-test with a 0.05 alpha-level.

RESULTS

PROPENSITY MATCHING

Statistical analyses were performed to evaluate the match between *Into Math* students and comparison group students on 10 characteristics (e.g., Percent Proficiency, Percent English language learners), see Table 2.

These analyses revealed that there were significant differences ($p < .05$) on baseline proficiency overall, with comparison schools having greater proficiency than *Into Math* schools in 2019. There was also a statistical trend ($p < .10$) with *Into Math* schools' average class size tending to be larger than comparison schools' class size. Given these differences, these two variables were added as covariates in the final analysis.

The remaining analyses revealed no other significant differences between the groups, $p > .05$.

Group	Category	<i>Into Math</i> (n=118)	Comparison (n=115)
Avg. Baseline Proficiency Overall*	--	27.55% (17.08)	30.20% (19.89)
Average Class Size**	--	20.83 (3.35)	20.49 (4.18)
Baseline Proficiency by Grade	3	34.63% (15.32)	37.62% (22.41)
	4	27.09% (15.52)	30.88% (16.86)
	5	25.68% (16.53)	29.44% (19.66)
	6	21.32% (15.37)	26.65% (22.60)
	7	25.49% (18.64)	25.22% (13.65)
Gender (% in School)	8	28.06% (20.71)	28.93% (21.46)
	Male	50.97% (2.35)	50.56% (2.87)
Race/Ethnicity (% in School)	Female	49.03% (2.35)	49.44% (2.89)
	White	70.48% (20.78)	67.12% (27.97)
	African American	3.32% (4.41)	6.83% (10.83)
	Hispanic	18.21% (19.76)	16.90% (24.34)
	Asian	2.81% (8.19)	3.25% (7.44)
Subpopulations (% in School)	Multi-Racial	4.95% (2.61)	5.72% (4.43)
	Economically Disadvantaged	53.00% (21.18)	51.35% (23.72)
	English Language Learner	7.86% (12.38)	7.73% (14.06)

*Covariate significant at $p < .05$ level. **Covariate significant at $p < .10$ level.

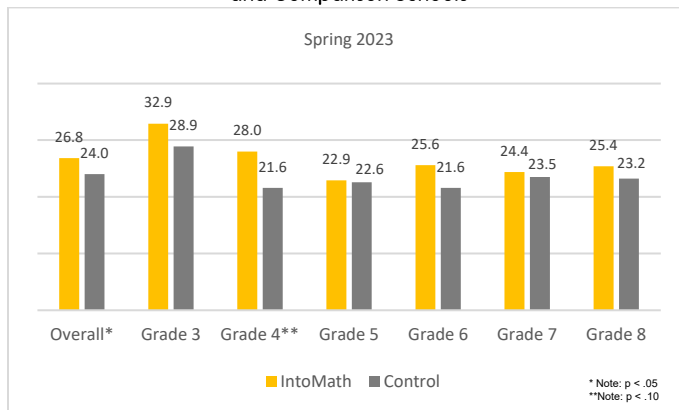
IMPACT OF INTO MATH

Math performance of schools that used the *Into Math* program were compared to closely matched schools using existing state assessment data from Spring 2023, while controlling for 2019 (pre-COVID) math performance.

Overall analyses of main effects show that schools using *Into Math* demonstrated a significantly higher proficiency rate than non-users, $p < .05$. The percent of students proficient in math was 2.0% points higher among *Into Math* schools as compared to control schools.

Grade level analysis revealed a similar trend in which *Into Math* users demonstrated higher proficiency rates than comparison schools in Grades 3–8. This finding was marginally significant for Grade 4, $p < .10$.

Figure 1. Spring 2023 Overall Estimated Marginal Means for *Into Math* and Comparison Schools



Subgroup analyses were also performed to determine if *Into Math* had a differential impact on certain student populations. Positive trends were observed for all subgroups except African Americans. Marginally

significant results were obtained for Female, White, and Economically Disadvantaged students (see Table 3), $p < .10$.

Table 3. Spring 2023 Estimated Marginal Means for <i>Into Math</i> and Comparison Schools as a Function of Subgroup				
Subgroup	<i>Into Math</i>	Comparison	Difference	Sign
Male	28.7	27.8	0.9	--
Female	25.6	22.2	3.4	**
White	30.4	27.5	2.9	**
African American	6.4	8.6	-2.2	--
Hispanic	18.7	16.5	2.2	--
Disadvantaged	16.5	14.0	2.5	**
ELL	8.7	6.0	2.7	--

Note: ** $p < .10$

CONCLUSION

Results from the *Into Math* quasi-experimental study showed that *Into Math* usage is associated with positive effects in student math performance in Illinois.

Into Math schools demonstrated higher proficiency rates than control schools, with statistically significant differences observed, $p < .05$.

Examination of results within each grade level also showed a similar positive pattern of results. In general, *Into Math* users demonstrated higher proficiency rates within Grades 3–8 in Year 3. Males, Females, Whites, Hispanics, Economically Disadvantaged, and English Language Learner *Into Math* students also demonstrated higher proficiency rates as compared to non-users but differences were only marginally significant for Female, White, and Disadvantaged students.

While the results are encouraging, the study has several limitations that should be considered when interpreting results. The IL State Education Department masks data for students in subgroups with less than 10 students; as such, the data from *Into Math* schools do not represent all impacted students.

Further, the study uses assessment data from before and after the COVID-19 pandemic. As such, results from this study should be considered in conjunction with other education impact analyses which took place around this time to get a better understanding of how COVID-19 might have influenced standardized assessment scores.

Researchers were unable to obtain implementation data on how well teachers implemented the program and their level of fidelity to the learning model. Such data could reveal stronger effects as program implementation is a known mediator when conducting curricula research.

Lastly, quasi-experimental studies like these rely on existing users. Thus, the analysis cannot completely rule out other potential external factors that could be leading to these results. However, the statistical procedures put into place, and the consistency of the findings do increase the confidence that usage of *Into Math* was the underlying mechanism leading to the significant differences between treatment and comparison.

Despite these limitations, the results from this quasi-experimental study using IL state assessment data provides evidence that *Into Math* is an effective math program that can help accelerate student learning.