What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

October 2012

WWC Review of the Report "Mastery Learning and Student Teams: A Factorial Experiment in Urban General Mathematics Classes"[1]

The findings from this review do not reflect the full body of research evidence on *mastery learning* or *student team learning*.

What is this study about?

The study examined the effects of *mastery learning* and *student team learning* on the math achievement of high school students. The analysis included 588 ninth-grade students in 16 urban Philadelphia high schools.

Within each school, teachers were randomly assigned to one of four conditions:

1. *Focused Instruction:* a comparison condition that consisted of teacher instruction, student worksheets, and quizzes
2. *Mastery:* the same components of focused instruction plus formative tests, corrective instruction to help students improve performance, and enrichment activities
3. *Teams:* the same components of focused instruction, except that students completed worksheets as teams and received rewards based on team scores
4. *Teams-and-Mastery:* the combined elements of the focused instruction, mastery, and teams conditions

The study assessed the effectiveness of the different conditions after one year by comparing student performance on a shortened version of the Comprehensive Test of Basic Skills (CTBS), Mathematics Computations and Concepts and Applications subscales.[2]

WWC Rating

The research described in this report meets WWC evidence standards with reservations

Strengths: This study is a well-implemented randomized controlled trial.

Cautions: There was high attrition of students from this study; however, the authors were able to demonstrate the equivalence of the analytic samples at baseline.

Features of *Mastery Learning* and *Student Team Learning*

Mastery learning is a method of teaching that requires that students master a skill before moving on to the next skill. Students must demonstrate mastery of a skill by scoring sufficiently high on a formative assessment (e.g., 80%–90% correct) before moving to the next topic.

Student team learning places students into four-member heterogeneous teams. They are encouraged to help each other to understand and master material, rather than working independently. To foster teamwork, teams receive incentives (such as being recognized in a weekly class newsletter) based on the average performance of their team on individually administered assessments.

What did the study find?

The study found, and the WWC confirmed, a statistically significant positive effect on math achievement for students in *team learning* (i.e., the combined

Teams and *Teams-and-Mastery* groups) compared with those not in *team learning* (i.e., the combined *Focused Instruction* and *Mastery* groups). The study found, and the WWC confirmed, no statistically significant effects on math achievement for students using *mastery learning* (i.e., the combined *Mastery* and *Teams-and-Mastery* groups) compared with those not using *mastery learning* (i.e., the combined *Focused Instruction* and *Teams* groups). Finally, the study showed no significant differences among any of the individual groups in the study (for example, *Teams* only vs. *Teams-and-Mastery*).

Appendix A: Study details

Slavin, R. E., & Karweit, N. L. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal, 12*(4), 725–736.

Setting

The study took place in ninth-grade classrooms in 16 junior and senior high schools in Philadelphia over the course of a school year.

Study sample

Teachers were stratified by school and randomly assigned to one of four conditions: (1) *Focused Instruction* (the comparison condition), (2) *Mastery*, (3) *Teams*, or (4) *Teams-and-Mastery*. The initial sample included 1,092 students in 44 classrooms in 16 schools. The analysis sample (students with valid pre- and posttest data) included 588 students from those 44 classrooms in 16 schools. Of those, 76% of the research sample were Black, 19% were White, 6% were Hispanic, and 1% were Asian American.

Intervention group

There were three intervention conditions assessed in this study: *Teams*, *Mastery*, and the combination of the two (*Teams-and-Mastery*).

Study authors reported that students in the *Teams* condition experienced a sequence of teaching, worksheet study in heterogeneous four-member teams, and weekly formative assessments. To encourage peer teaching in groups, the group with the highest team average performance on weekly formative assessments was recognized in a weekly class newsletter.

Students in the *Mastery* condition experienced the same teaching content that was provided to the *Teams* condition, but worked individually on their worksheets. Like the students in the *Teams* condition, students took weekly formative quizzes (to assess whether students achieved the 80% mastery criterion) but did not receive incentives for performance. Students who had not achieved mastery received corrective instruction, and those students who achieved mastery participated in enrichment activities.

The *Teams-and-Mastery* group combined the aspects of both the *Teams* and *Mastery* conditions. Students in all four conditions (intervention and comparison groups) used the same schedule and curriculum materials, adapted from *Mathematics for Today* (Tobin, 1975).

Comparison group

The *Focused Instruction* group included the following schedule: teaching (the same content provided across the intervention conditions), individual worksheets, and weekly quizzes. Students in this condition worked individually, did not receive corrective instruction, and did not receive incentives for performance.

Outcomes and measurement

Every third item from the Mathematics Computations and Concepts and Applications subscales of the CTBS, Level 2, Form S was selected as an item for an assessment that could be completed in a single class period. This test was used as the pretest and posttest. For a more detailed description of this outcome measure, see Appendix B.

Support for implementation

Teachers were trained by school district staff development personnel.

Reason for review

This study was identified for review by the WWC because it was cited as evidence in an Investing in Innovation (i3) grant proposal.

Appendix B: Outcome measure for the mathematics achievement domain

Mathematics achievement	
Comprehensive Test of Basic Skills (CTBS), Mathematics Computations and Concepts and Applications subscales, Level 2, Form S	Researchers created a 30-item test that could be completed during a class period by selecting every third item from the CTBS assessment.

Appendix C: Study findings for the mathematics achievement domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
CTBS Mathematics Computations and Concepts and Applications subscales	*Mastery* vs. no *Mastery*	44 classrooms/ 588 students	16.71 (5.54)	16.63 (5.74)	0.08	0.01	+1	> 0.05
CTBS Mathematics Computations and Concepts and Applications subscales	*Teams* vs. no *Teams*	44 classrooms/ 588 students	17.32 (5.76)	16.15 (5.58)	1.17	0.21	+8	< 0.03

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. This study is characterized as having a statistically significant positive effect because univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. CTBS = Comprehensive Test of Basic Skills.

Study Notes: A correction for multiple comparisons was needed but did not affect significance levels. The *p*-values presented here were reported in the original study. The authors conducted a 2x2 random effects, nested analysis of covariance, and reported the F-tests of the *Mastery*, *Teams*, and *Teams-and-Mastery* interaction in the article. The *p*-values for the main effects are presented here, and the test of the interaction term was nonsignificant (contrasts for each individual group are presented in Appendix D). The authors provided adjusted means and pooled standard deviations for each group in an email correspondence.

Appendix D: Supplemental findings for the mathematics achievement domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
CTBS Mathematics Computations and Concepts and Applications subscales	*Mastery* vs. comparison	25 classrooms/ 325 students	16.18 (5.83)	16.10 (5.32)	0.08	0.01	+1	0.95
CTBS Mathematics Computations and Concepts and Applications subscales	*Teams-and-Mastery* vs. comparison	20 classrooms/ 285 students	17.32 (5.20)	16.10 (5.32)	1.22	0.23	+9	0.30
CTBS Mathematics Computations and Concepts and Applications subscales	*Teams-and-Mastery* vs. *Mastery*	23 classrooms/ 290 students	17.33 (5.20)	16.19 (5.83)	1.14	0.20	+8	0.31

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. CTBS = Comprehensive Test of Basic Skills.

Study Notes: Corrections for clustering and multiple comparisons were needed but did not affect significance levels. The authors did not conduct tests of the differences between individual groups in the article (the F-test of the *Teams-and-Mastery* interaction was nonsignificant, suggesting that none of the individual groups were significantly different from each other), and as such, the *p*-values presented here were calculated by the WWC. There were a total of six possible comparisons of the individual *Mastery* and *Team* intervention combinations; however, only the three contrasts presented here were eligible to meet WWC standards with reservations. The comparisons of *Teams-and-Mastery* vs. *Teams*, *Teams* vs. *Mastery*, and *Teams* vs. comparison required a statistical adjustment for differences in the pretest scores of the groups at baseline, which the authors did not conduct, and as such, these contrasts do not meet WWC standards. The WWC calculated the intervention group mean by adding the difference-in-differences adjusted estimate of the average impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttests means. Please see the *WWC Handbook* for more information.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the High School Mathematics review protocol, version 2.0. The WWC rating applies only to the results that were eligible under this topic area and met WWC standards without reservations or met WWC standards with reservations, and not necessarily to all results presented in the study.

[2] There were three individual contrasts that did not meet WWC standards because the groups were not shown to be equivalent at baseline: *Teams-and-Mastery* vs. *Teams*, *Teams* vs. *Mastery*, and *Teams* vs. *Focused Instruction*. These results are not included in this report.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, October). *WWC review of the report: Mastery learning and student teams: A factorial experiment in urban general mathematics classes.* Retrieved from http://whatworks.ed.gov.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.