Bridging Research and Practice: Efficacy of a Research-Based Redesign of a Grade 7 Mathematics Curriculum

Yvonne Kao (ykao@wested.org)
Jodi Davenport (jdavenp@wested.org)
Bryan Matlen (bmatlen@wested.org)
Larry Thomas (lthomas@wested.org)
Steven Schneider (sschnei@wested.org)
WestEd

Abstract

To test whether research can be effectively translated into practice, the 7th grade *Connected Mathematics Project* 2 curriculum was redesigned according to three research-based principles. The efficacy of the redesigned curriculum was tested against the original curriculum in a two-year cluster-randomized trial. Schools (n = 114) were randomly assigned to use the redesigned curriculum (treatment) or the original curriculum (control). Treatment students scored higher than control students on six of eight unit post-tests. For two of these units, differences between treatment and control were large enough to be considered "substantively important" and one was statistically significant. In addition, moderator analyses suggested a global trend for the treatment to be more beneficial for traditionally underperforming students.

Citation

Kao, Y., Davenport, J., Matlen, B., Thomas, L., & Schneider, A. (2017). *Bridging research and practice: Efficacy of a research-based redesign of a grade 7 mathematics curriculum*. Paper presented at the Annual Meeting of the American Educational Research Association, San Antonio, TX.

Bridging Research and Practice: Efficacy of a Research-Based Redesign of a Grade 7 Mathematics Curriculum

Yvonne Kao (ykao@wested.org)
Jodi Davenport (jdavenp@wested.org)
Bryan Matlen (bmatlen@wested.org)
Larry Thomas (lthomas@wested.org)
Steven Schneider (sschnei@wested.org)
WestEd

Objective

Lab-based research in cognitive and learning sciences provides many recommendations for improving learning and instruction (e.g., Pashler et al., 2007). Tightly-controlled experiments demonstrate that many strategies improve learning: mapping between visual representations, prompting for explanation of worked examples, using quizzing to promote learning, and spacing practice opportunities over time (e.g., Cooper & Sweller (1987); Cepeda, Pashler, Vul, et al. 2006; Clark & Mayer, 2003; Larkin & Simon, 1987; Mayer 2001; Kalyuga, Chandler, and Sweller (2001); Paas & Van Merrienboer, 1994; Rohrer & Taylor, 2006). Though promising, the vast majority of studies focus on specific strategies in isolation, use researcher-created instructional materials, and test effects of brief instructional manipulations. For cognitive science research to make an impact on classroom learning, the synergistic effects of the strategies must be tested in real-world settings with real curricula over durations that are educationally meaningful (i.e., learning on the order of weeks and months rather than hours and days). In this paper, we describe a large-scale effort to bridge research and practice by applying cognitive principles to redesign an existing mathematics curriculum and testing the efficacy of these materials.

This project brought together researchers and practitioners from six institutions. The starting point of this work was the practice guide, *Organizing Instruction and Study to Improve Student Learning*, distributed by the Institute of Education Sciences (Pashler et al., 2007). To test the synergistic effects of research-based instructional strategies, the research team translated these recommendations into three design principles: 1) prompt for self-explanation of correct and incorrect worked examples, 2) visually support mapping between representations, 3) carefully space practice of critical content and skills over time by using quizzes formatively. These principles have strong research backing, are broadly applicable to instruction, and can be readily implemented in a range of curricular materials.

To ensure the application of research-based principles is ecologically valid, the team chose to make revisions to a widely-used 7th grade math curriculum, *Connected Mathematics Project 2* (CMP2). The curriculum is divided into eight units:

1. **Variables and Patterns (VP)**: Introducing Algebra
2. **Stretching and Shrinking (SS)**: Similarity
3. **Comparing and Scaling (CS)**: Ratio, Proportion, and Percent
4. **Accentuate the Negative (AN)**: Positive and Negative Numbers
5. **Moving Straight Ahead (MS)**: Linear Relationships

6. **Filling and Wrapping (FW)**: Three-Dimensional Measurement
7. **What Do You Expect? (WE)**: Probability and Expected Value
8. **Data Distributions (DD)**: Describing Variability and Comparing Groups

Application of Cognitive Theories

Applying the principles to revise instructional materials (i.e., the print curriculum) and instructional practice (i.e., what happens in the classroom) required expertise across many fields. Three teams of cognitive researchers were formed to revise CMP2 using the three research-based principles. Additional teams devoted to mathematics, professional development, and production collaborated to ensure that the revised materials were grounded in the research findings, were mathematically accurate and appropriate (in terms of student development and curriculum standards), were clearly specified for teachers, and were produced with a high level of technical quality. Below we summarize the nature of the revisions to the CMP2 curriculum. Figure 1 shows a page from the original curriculum and its redesigned counterpart.

Worked Examples

To revise CMP2 according the *Worked Examples* principle, the first instance of a particular problem type was turned into a fully worked (correct or incorrect) example paired with a self-explanation prompt that requires the student to explain a key feature in the problem or explain why a step would not be correct. The next instance of a problem type was converted to a faded worked example, that requires the student to complete a partially-worked problem. Subsequent instances are left for the students to solve independently.

Visual-Verbal Mapping

Next, the *Visual-Verbal Mapping* team made edits, and indicated which irrelevant graphics were to be removed, where graphics should be created or modified, and when text should be changed to make the mapping between the text and visual information more salient. New or modified graphics supported the integration of representations by maintaining common colors across multiple representations and using proximity to group relevant information (Harp & Mayer, 1998; Moreno & Mayer, 1999). Text changes included adding text that offered a spatial directive (e.g., “below”) to indicate which visual was linked to the text and adding labels to diagrams.

Spacing and Formative Assessment

Finally, the *Spacing* and *Formative Assessment* teams redesigned CMP2’s *Teacher’s Guide* for each unit rather than the student materials. Focusing on the *Teacher’s Guide* allowed changes to be made to the spacing of practice and assessment of skills acquired earlier in the curriculum. Teacher materials were modified to include charts showing which skills should have been mastered before the start of each unit, when students should have mastered the skill (according to CMP2), a prototypical example of each skill in the mastered unit, and practice problems for each skill from CMP2’s *Additional Practice and Skills Workbook*. Teachers were also provided with a Prior Skills Assessment to evaluate students’ mastery of the prior skills. Assessment items were drawn from selected problems in the *Additional Practice and Skills Workbook*. The Prior Skills Assessment was designed to be used formatively. Teachers may have decided to review one or more prior skills before beginning the unit based on students’ performance.

Research Questions

1. Do 7th grade students who are exposed to the redesigned curriculum (treatment) show greater learning than students exposed to the original curriculum (control)?
2. Does the effect of the redesigned curriculum differ for traditionally lower-performing students?

Methods

Design and Procedure

The efficacy of the redesigned curriculum was evaluated using a two-year cluster-randomized trial. The purpose of the two-year design was to allow teachers to learn and practice with the research-based principles during the first year and only measure the impact of the redesign in the second year. One-hundred and fourteen schools (containing 181 teachers recruited in two cohorts and 2,596 students) that were already using the CMP2 curriculum participated in the study; schools were randomly assigned to continue using the original CMP2 curriculum (control group) or to adopt the redesigned curriculum (treatment group) for the study period. Treatment teachers participated in a 2-day workshop at the start of the study, to introduce them to the research-based principles and their implications for practice. Treatment teachers also participated in regular follow-up workshops throughout the study period.

Teachers completed a baseline measure of mathematics pedagogical content knowledge and administered a pre-test to students at the beginning of the second study year. The pre-test measure was the pre-algebra readiness diagnostic developed by the Mathematics Diagnostic Testing Project (MDTP). Teachers then taught the CMP2 curriculum following their school pacing guide, administering post-tests for each of the eight curriculum units immediately upon finishing each unit.

Measures

The outcome measures reported in this paper are project-developed unit post-tests. One assessment was developed for each of the eight units in the curriculum. The final assessments included approximately 16 multiple-choice items. Approximately half of the items were derived from the original CMP2 curriculum materials and the remaining items were sourced from state, national, and international standardized tests. The final assessments also included one performance task from the Balanced Assessment of Mathematics. The assessments were designed this way to balance proximity to the CMP2 curriculum and comparability to norm-referenced standardized exams.

Analysis

To address the aforementioned research questions, analyses were conducted using hierarchical linear models (HLM) to estimate the impact of the treatment on the performance of grade 7 CMP unit tests.

Overall treatment impact for the seven unit tests were estimated using the following three-level HLM:

$$Y_{ijk} = \gamma_{000} + \gamma_{001}SchTX_k + \sum_1^p \gamma_{p00}\, STU_{ijk} + \sum_1^p \gamma_{0p0}\, TEA_{jk}$$
$$+ \sum_2^p \gamma_{00p}\, SCH_k + \nu_{00k} + u_{0jk} + \epsilon_{ijk} \qquad [1]$$

where Y_{ijk} is the post-test scale score for a given unit test outcome for the ith student of the jth teacher in the kth school. γ_{000} is the adjusted grand mean of the post-test score, $\gamma_{001}$is the effect of being assigned to the treatment condition, ν_{00} and u_{0jk} are random effect terms for school- and teacher-level clustering, and ϵ_{ijk} is a student residual term. $\sum_1^p \gamma_{p00}\, STU_{ijk}$ represents a vector of student-level covariates and their regression coefficients. Student-level covariates adjusted for in model [1] include the MDTP pre-algebra readiness pre-test scale scores, student gender, ethnicity, English language learner (ELL) status, and special education status. $\sum_1^p \gamma_{0p0}\, TEA_{jk}$ represents a vector of teacher-level covariates and their regression coefficients. Teacher-level covariates adjusted for in model [1] include classroom-averaged MDTP pre-test scale scores, cohort assignment, baseline teacher pedagogical content knowledge scores, and years of teaching experience. Lastly, $\sum_2^p \gamma_{00}\ \ SCH_k$ represents a vector of school-level covariates and their regression coefficients. School-level covariates adjusted for in model [1] include urbanicity, percent of students eligible for free or reduced price lunch, percent of students who are underrepresented ethnicities in STEM, and student pass rates on the state standards-based mathematics test.

Differential treatment impacts for student subgroups were evaluated using a three-level HLM similar to model [1] with the inclusion of specific treatment and student-level covariate interactions. Student subgroups of interest include under-performing students on pre-algebra readiness pretest, gender, ELLs, and underrepresented STEM minorities.

Results

The remainder of the paper reports analyses of the first seven units in the CMP2 curriculum—many participating schools omitted the last unit, Data Distributions, from their pacing guides and so very few post-tests for that unit were returned to the research team.

Impact of treatment assignment on student learning

Results of the overall average treatment impact are summarized in Table 1. Across the seven units evaluated, treatment effects are generally positive in favor of students who received the modified CMP2 units. Only one unit, *Accentuate the Negative*, showed a negative (but near-zero) treatment effect. Two units, *Comparing and Scaling* and *What Do You Expect?* have effect size estimates large enough to be considered substantively important by the What Works Clearinghouse (WWC) Standards (2014), but only *Comparing and Scaling* showed a statistically significant positive effect.

Interactions between treatment and student characteristics on learning

To evaluate if the use of the modified CMP2 curriculum had differential impacts on student subgroups, cross-level treatment and student covariate interactions were included in model [1]. Table 2 shows the treatment-by-pre-test interaction effects. Following the procedures by Bauer and Curran (2005) and Tate (2004) to visualize the differential effect, Figure 2 shows how expected treatment effects on each unit test varied as a function of student pretest. Students who performed below zero on the pre-test (zero corresponds to mean performance) are associated with more positive treatment effects while student who performed above zero associated with more negative treatment effects. However, the interaction is only marginally significant for two units: *Variables and Patterns* and *What Do You Expect?*

A similar pattern emerges when modeling interactions between treatment assignment and student demographic covariates. The treatment tended to be more effective for students who are members of traditionally underperforming subgroups in mathematics: females, English language learners, and underrepresented ethnicities in mathematics (i.e., not white or Asian). However, the interactions are not consistently statistically significant. Table 3 shows the treatment-by-subgroup interaction effects for each unit.

Discussion

Overall, the redesigned curriculum exhibited trends for positive impact on six curriculum unit. Two units (*Comparing and Scaling* and *What Do You Expect?*) showed substantively important positive effect sizes (i.e., greater than 0.25; WWC, 2014). However, only one unit showed a statistically significant positive effect. The lack of significance may be due in part to low power as a result of high attrition observed over the two-year study.

Moderator analyses showed that the impact of the treatment interacted with baseline student characteristics, but these interactions tended to vary. While very few interactions were statistically significant, there is nonetheless a globally consistent trend suggesting that the treatment may be more effective for traditionally underperforming subgroups.

These findings suggest that purposefully engineering curricula based on learning sciences research is productive for improving student learning and may increase equity in educational outcomes. At the same time, the observed effects from this study were small and variable. Further analyses of project data will explore how the content of individual units and teacher practice may have mediated learning outcomes.

References

Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. Multivariate Behavioral Research, 40, 373–400.

Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., and Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. Psychological Bulletin, 132, 354-380.

Clark, R.C., and Mayer, R.E. (2003). e-Learning and the science of instruction: Proven guidelines for consumers and designers of multimedia Learning. San Francisco: Jossey-Bass.

Cooper, G., and Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. Journal of Educational Psychology, 79, 347–362.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of educational psychology*, *90*(3), 414.

Kalyuga, S., Chandler, P., and Sweller, J. (2001). Learner experience and efficiency of instructional guidance. Educational Psychology, 21, 5–23.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words.*Cognitive science*, *11*(1), 65-100.

Mayer, R.E. (2001). Multimedia learning. New York: Cambridge University Press.

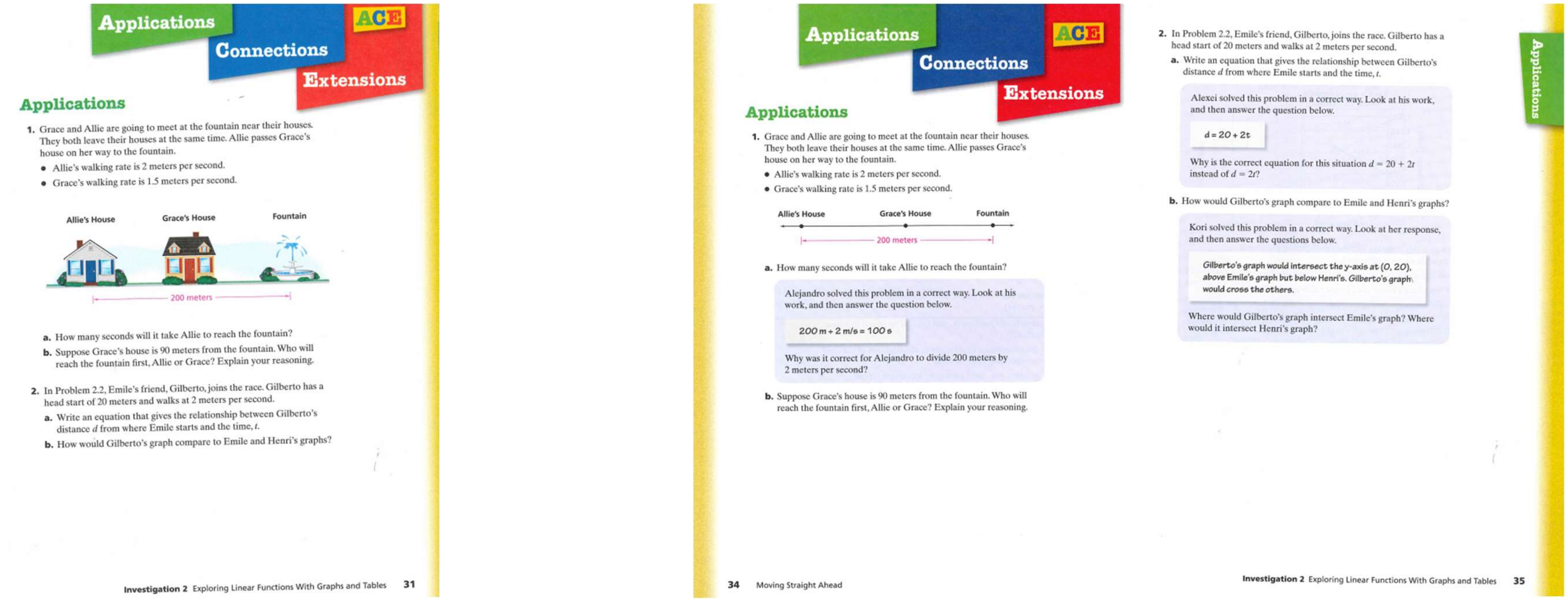
Moreno, R., and Mayer, R.C. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. Journal of Educational Psychology, 91, 358-368.

Paas, F., and van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problemsolving skills: A cognitive-load approach. Journal of Educational Psychology, 86, 122–133.

Rohrer, D., and Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. Applied Cognitive Psychology, 20, 1209-1224.

Tate, R. L. (2004). Interpreting hierarchical linear and hierarchical generalized linear models with slopes as outcomes. The Journal of Experimental Education, 73, 71–95.

Figure 1.



Applications Connections Extensions ACE

Applications

1. Grace and Allie are going to meet at the fountain near their houses. They both leave their houses at the same time. Allie passes Grace's house on her way to the fountain.
   - Allie's walking rate is 2 meters per second.
   - Grace's walking rate is 1.5 meters per second.

   a. How many seconds will it take Allie to reach the fountain?
   b. Suppose Grace's house is 90 meters from the fountain. Who will reach the fountain first, Allie or Grace? Explain your reasoning.

2. In Problem 2.2, Emile's friend, Gilberto, joins the race. Gilberto has a head start of 20 meters and walks at 2 meters per second.
   a. Write an equation that gives the relationship between Gilberto's distance d from where Emile starts and the time, t.
   b. How would Gilberto's graph compare to Emile and Henri's graphs?

Investigation 2 Exploring Linear Functions With Graphs and Tables 31

Applications Connections Extensions ACE

Applications

1. Grace and Allie are going to meet at the fountain near their houses. They both leave their houses at the same time. Allie passes Grace's house on her way to the fountain.
   - Allie's walking rate is 2 meters per second.
   - Grace's walking rate is 1.5 meters per second.

   a. How many seconds will it take Allie to reach the fountain?

   Alejandro solved this problem in a correct way. Look at his work, and then answer the question below.

   200 m ÷ 2 m/s = 100 s

   Why was it correct for Alejandro to divide 200 meters by 2 meters per second?

   b. Suppose Grace's house is 90 meters from the fountain. Who will reach the fountain first, Allie or Grace? Explain your reasoning.

34 Moving Straight Ahead

2. In Problem 2.2, Emile's friend, Gilberto, joins the race. Gilberto has a head start of 20 meters and walks at 2 meters per second.
   a. Write an equation that gives the relationship between Gilberto's distance d from where Emile starts and the time, t.

   Alexei solved this problem in a correct way. Look at his work, and then answer the question below.

   $d = 20 + 2t$

   Why is the correct equation for this situation $d = 20 + 2t$ instead of $d = 2t$?

   b. How would Gilberto's graph compare to Emile and Henri's graphs?

   Kori solved this problem in a correct way. Look at her response, and then answer the questions below.

   Gilberto's graph would intersect the y-axis at (0, 20), above Emile's graph but below Henri's. Gilberto's graph would cross the others.

   Where would Gilberto's graph intersect Emile's graph? Where would it intersect Henri's graph?

Applications

Investigation 2 Exploring Linear Functions With Graphs and Tables 35

A page from the original CMP2 curriculum (left) and the redesigned version (right). The redesigned version contains a revised figure and introduces worked examples into the problems.

Table 1.

Summary of treatment effects by curriculum unit.

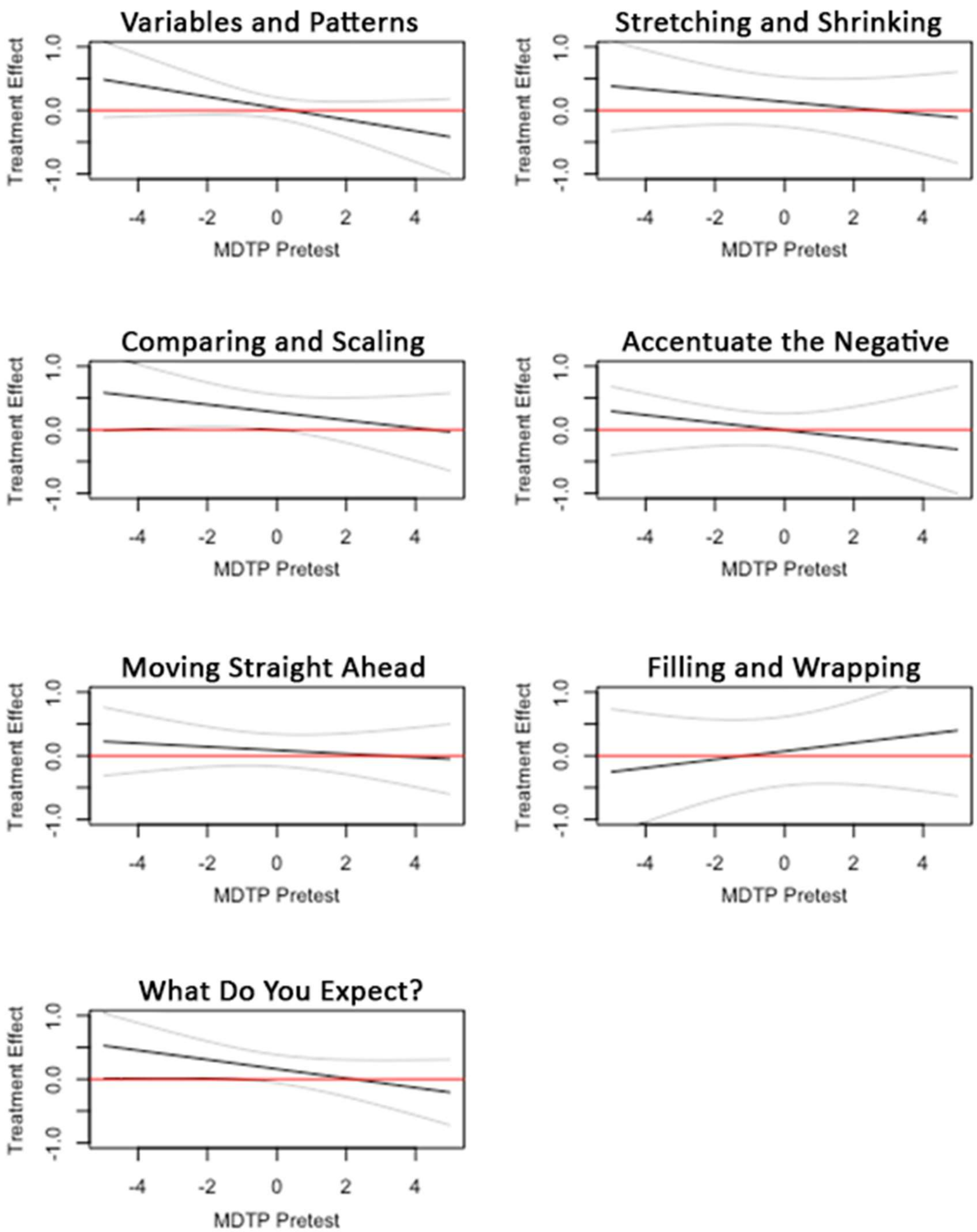
Unit Test	Coefficient	Standard Error	*p*	Effect Size
Variables and Patterns	0.04	0.09	0.69	0.07
Stretching and Shrinking	0.13	0.18	0.46	0.15
Comparing and Scaling	0.27	0.12	0.04	0.35
Accentuate the Negative	-0.01	0.12	0.91	-0.01
Moving Straight Ahead	0.08	0.11	0.47	0.13
Filling and Wrapping	0.09	0.23	0.70	0.12
What Do You Expect?	0.15	0.10	0.13	0.32

Table 2.

Treatment-by-pretest interaction effects by unit. A negative coefficient indicates the treatment was more effective at improving learning outcomes for students who were lower-performing at pre-test.

Moderator	**Unit Test**	**Coefficient**	**Standard Error**	***p***
Pre-test scale score	Variables and Patterns	-0.09	0.05	0.06
	Stretching and Shrinking	-0.05	0.05	0.35
	Comparing and Scaling	-0.06	0.05	0.20
	Accentuate the Negative	-0.06	0.06	0.29
	Moving Straight Ahead	-0.03	0.04	0.51
	Filling and Wrapping	0.07	0.07	0.37
	What Do You Expect?	-0.07	0.04	0.06

Figure 2.



The black line on each graph shows the treatment effect for students at different levels of performance on the MDTP pre-test, with zero indicating average performance. The slope of the black line corresponds to the coefficient of the interaction term. When the black line is above the red line, the treatment effect is positive. Negative slopes indicate that the treatment was more effective for lower-performing students. The gray lines indicate the 95% confidence interval around the estimate of the treatment effect.

Table 3.

Influence of moderator variables on treatment effects by unit. A positive coefficient indicates the treatment was more effective at improving learning outcomes for students in the traditionally lower-performing subgroup.

Moderator	Unit Test	Coefficient	Standard Error	*p*
Gender	Variables and Patterns	0.04	0.07	0.60
	Stretching and Shrinking	0.06	0.08	0.42
	Comparing and Scaling	-0.06	0.07	0.40
	Accentuate the Negative	0.04	0.09	0.65
	Moving Straight Ahead	0.11	0.06	0.09
	Filling and Wrapping	-0.04	0.10	0.69
	What Do You Expect?	0.05	0.06	0.38
English language learner status	Variables and Patterns	0.06	0.28	0.83
	Stretching and Shrinking	-0.03	0.25	0.90
	Comparing and Scaling	-0.29	0.24	0.23
	Accentuate the Negative	0.61	0.29	0.03
	Moving Straight Ahead	0.16	0.20	0.41
	Filling and Wrapping	0.19	0.41	0.64
	What Do You Expect?	0.00	0.18	1.00
Underrepresented ethnicity in STEM	Variables and Patterns	-0.02	0.12	0.86
	Stretching and Shrinking	0.09	0.12	0.46
	Comparing and Scaling	-0.14	0.11	0.19
	Accentuate the Negative	-0.13	0.12	0.29
	Moving Straight Ahead	0.21	0.09	0.02
	Filling and Wrapping	0.16	0.15	0.28
	What Do You Expect?	0.11	0.09	0.21