Title
Integrating Cognitive Principles to Redesign a Middle School Math Curriculum

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Davenport, Jodi
Kao, Yvonne
Schneider, Steven

Publication Date
2013

Peer reviewed

Integrating Cognitive Principles to Redesign a Middle School Math Curriculum

Jodi L. Davenport (jdavenp@wested.org)
WestEd STEM Program, 300 Lakeside Drive, 25th Floor
Oakland, CA 94612 USA

Yvonne S. Kao (ykao@wested.org)
WestEd STEM Program, 400 Seaport Court, Suite 222
Redwood City, CA 94063 USA

Steven A. Schneider (sschnei@wested.org)
WestEd STEM Program, 400 Seaport Court, Suite 222
Redwood City, CA 94063 USA

Abstract

Does a middle school mathematics curriculum that is redesigned using principles based in cognitive research improve student outcomes? To test whether research can be effectively translated into practice, the *Connected Mathematics Project 2* (CMP2) curriculum was revised according to four principles 1) integrating visual with verbal information, 2) prompting for self-explanation of correct and incorrect worked examples, 3) spacing learning over time, and 4) using formative assessment. This study of 6th grade and 8th grade mathematics education addresses the research question: "Do students who are exposed to specific redesigned CMP2 curriculum modules (treatment) exhibit greater improvements in mathematics performance in the module-specific content area than their counterparts exposed to the regular CMP2 curriculum (control)?" Preliminary analyses show statistically significant effects of the redesigned CMP2 units in three of the four curricular units in this study.

Keywords: cognitive psychology; mathematics; math education; education; spaced learning; formative assessment; worked examples; visual representations

Introduction

Lab-based research in cognitive and learning sciences has led to a number of recommendations for improving learning and instruction (e.g., Pashler et al., 2007). Tightly controlled experiments have shown that learning can be enhanced with strategies such as mapping between visual representations, prompting for explanation of worked examples, using quizzing to promote learning and spacing practice opportunities over time. The vast majority of studies focus on specific strategies in isolation rather than how principles may be combined. If these research findings are to be meaningfully applied to classrooms, the synergistic effects of the strategies must be tested in real-world settings. In the current paper, we describe a large-scale effort of the National Center on Cognition and Mathematics Instruction (Math Center) in the United States to bridge research and practice by applying cognitive principles to redesign an existing mathematics curriculum and testing the efficacy of these materials.

To test the synergistic effects of research-based instructional strategies, the Math Center applied four principles to redesign *Connected Mathematics Project 2* (CMP2), a widely-used middle school (grades 6-8) mathematics curriculum. The Math Center team selected cognitive-based principles shown to improve student learning: 1) integrating visual with verbal information to promote the integration of concepts, 2) prompting for self-explanation of correct and incorrect worked examples, 3) carefully spacing the learning of critical content and skills over time, and 4) using quizzes to provide focused feedback and adjust instruction to the needs of students.

The CMP2 curriculum is an NSF-funded, research-based curriculum for grades 6-8 that covers topics emphasized in both national and state standards and aligns well with key ideas from the NCTM (2006) *Focal Points.* Key features of the curriculum are that it (1) is organized around important mathematics ideas and processes, e.g., number sense, symbolic reasoning, and probability, (2) is problem-centered, and (3) builds and connects concepts across problems, units, and grades. Each year of the curriculum is divided into eight units; each unit includes a student booklet and accompanying teacher materials to support instructional practice.

Applying the principles to revise instructional materials (e.g., the print curriculum) and instructional practice (e.g., what happens in the classroom) required expertise across many fields. Teams devoted to cognition research, 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. The iterative, multi-layered design process that we have developed for integrating the cognitive principles with the CMP2 curriculum applies not only in the context of mathematics instruction, but also to bridging research with instructional design across content areas.

The Principles

The following four principles were selected as they have demonstrated effectiveness in student learning, have broad applicability to instruction, and can be readily implemented in a range of curricular materials.

Integrating Visual and Verbal Information Combining visual information with verbal descriptions serves two important functions in mathematics instruction: 1) ensuring that text for instruction and problem-solving are perceived and understood and 2) promoting fluency in mapping between representations (e.g., equations, diagrams, graphs, or tables). To maximize learning benefits, research suggests that visual and verbal information should be integrated (e.g., Clark & Mayer, 2003; Larkin & Simon, 1987; Moreno & Mayer, 1999) and task-irrelevant information should be removed (e.g., Harp & Mayer, 1998). Visual cues such as color, proximity and grouping can support integration. Removing "seductive details;" that is, representations that are engaging but only tangentially related to the topic of instruction or the problem at hand (e.g., Harp & Mayer, 1998), helps learners focus on relevant information. To apply the visual mapping principle, researchers removed irrelevant images, added visual cues (e.g., color), and modified existing images to facilitate mapping.

Worked Examples In mathematics, students must learn to fluently carry out procedures across a variety of problem types. Interleaving problems to solve with worked examples of how to solve a problem improves student learning (Zhu & Simon, 1987; Clark & Mayer, 2003). Prompting students to explain worked examples further increases learning by facilitating the integration of new information. (Chi, 2000; Roy & Chi, 2005). In worked example exercises, students see complete or partially worked out solutions (which can be correct or incorrect) and explain the rationale behind problem solving steps or the error that was made in an incorrect example. Positive effects of interleaving worked examples have been reported in a variety of courses (Clark & Mayer, 2003; Paas & Van Merrienboer, 1994; Sweller & Cooper, 1985). Worked examples are more effective and more efficient for learning and transfer because they allow students to spend limited cognitive resources on understanding the ideas underlying the solutions rather than on generating solutions (Sweller, 1999). Further, explaining both correct and incorrect worked examples promotes greater learning than correct examples alone (Siegler, 2002; Siegler & Chen, 2008; Rittle-Johnson 2006). To apply the worked examples principle, researchers modified existing homework activities to include worked examples that prompt for self-explanation of problem solving steps.

Spaced Learning and Formative Assessment Extensive research in cognitive psychology has demonstrated large retention advantages when learners have multiple opportunities over time to practice key facts, concepts, and knowledge rather than few instances of "massed" practice, a phenomenon called the spacing effect (Cepeda et al., 2006; Rohrer & Taylor, 2007). When learners practice recalling and applying relevant information through quizzing, they are more likely to retain that knowledge for a greater period of time. Spacing instruction and practice reinforces connections between key ideas and promotes transfer.

Periodic testing provides students with opportunities to practice retrieving knowledge, reflect on the state of their knowledge, and transfer knowledge to new problems (Butler & Roediger, 2007; Roediger & Karpicke, 2006; Rohrer, 2009). Cycles of feedback and reflection that allow for revision and knowledge updating can help learners master targeted concepts and skills (e.g., Pavlik et al., 2007). Evidence from classroom learning contexts shows that the *formative* use of assessment can enhance instructional effectiveness (e.g., Black & Wiliam, 1998); here, formative assessment is defined as a process used by teachers and students that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes. In the revised materials, teachers were provided with quizzes and instruction on how to use feedback formatively in the classroom.

Method

The design of this study is a within-teacher cluster-randomized trial. The primary research question of this study is: "Do 6th and 8th grade students who are exposed to a redesigned curricular unit (treatment) show greater pre-to-post test improvements in mathematics scores than students exposed to the unmodified curricular unit (control)?"

Participants Researchers collected data from 64 6th grade teachers (1270 students at 45 schools) and 56 8th grade teachers (1180 students at 42 schools). Teachers had prior experience with the CMP2 curriculum and came from a diversity of schools across seventeen states in the United States. Background characteristics of participating teachers and demographic characteristics of their students are presented respectively in Table 1 and Table 2.

Table 1: Professional background of participating teachers.

Characteristic	6th Grade	8th Grade
Majored in math or math education	27%	43%
Advanced degree	64%	66%
Mean years of teaching experience	12.3 (SD = 8.2)	13.7 (SD = 7.6)

Table 2: Demographic characteristics of participating students.

Characteristic	6th Grade	8th Grade
Socioeconomically disadvantaged	41%	43%
Ethnicity		
White	67%	60%
Black	10%	13%
Hispanic	14%	13%
Other	9%	14%

Materials Two 6th grade units and two 8th grade units from the CMP2 curriculum were revised according to the cognitive principles described above. The 6th grade units used in this study were *Bits and Pieces III* (decimals and percents) and *Covering and Surrounding* (area and perimeter). The 8th grade units were *Shapes of Algebra* (linear equations and coordinate geometry) and *Say it with Symbols* (expressions and equations). Teams of researchers were formed for each of the principles. The cognitive research teams developed rubrics to identify whether the existing materials aligned with the cognitive design principles, and if not, to specify how the materials would be altered to be in compliance. Next each team made sequential revisions to the CMP2 materials. Changes that overlapped with other principles were discussed and resolved in biweekly meetings.

1. Coney Island Park wants a bumper-car ride with 24 square meters of floor space and 22 meters of rail section.

a. Sketch some floor plans for this request.

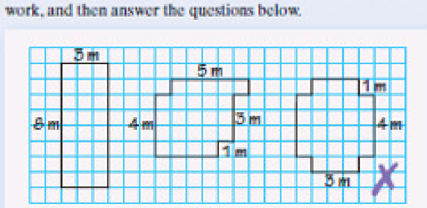
b. Describe the bumper-car ride in terms of its area and perimeter. Report what each measure tells you about the ride.

Figure 1: A problem from the original Covering and Surrounding unit.

1. Coney Island Park wants a bumper-car ride with 24 square meters of floor space and 22 meters of rail section.

a. Sketch some floor plans for this request.

Dominick completed the first two floor plans in a correct way, but his third plan does not meet the requirements. Look at his work, and then answer the questions below.



Dominick's third floor plan meets one requirement, but not the other. Which one does it fail to meet? How can you tell?

b. Describe the bumper-car ride in terms of its area and perimeter. Report what each measure tells you about the ride.

Figure 2: The revised version of the problem in Figure 1. A worked example has been incorporated into part a and the park photograph has been removed.

The mathematics team reviewed the revised curricular materials to ensure mathematical accuracy and appropriateness. Finally, the production team worked with the cognitive and math content teams to clarify design decisions as necessary. Examples of the original and revised curriculum materials are shown in Figures 1 and 2. Concurrent with the production of the materials, the professional development team met to develop measures of fidelity of implementation and to identify effective ways to communicate the underlying rationale and practical implementation of the cognitive design principles to the participating teachers.

Design This study used a within-teacher design: each teacher provided data from two units of CMP2, one revised and one control. Whether a given unit was used in its original or redesigned format was counterbalanced across participants. Teachers were randomly assigned to one of two groups, A and B, as depicted in Table 1 below. Group A served as the experimental group for one of the curriculum units and Group B served as the experimental group for the other. When multiple teachers taught at the same grade level in the same school, half the teachers at the school were assigned to group A and half to group B.

Table 1: Assignment of teachers to group.

Group	Treatment Unit	Control Unit
	6th Grade	
A	Bits and Pieces III	Covering and Surrounding
B	Covering and Surrounding	Bits and Pieces III
	8th Grade	
A	Say it with Symbols	Shapes of Algebra
B	Shapes of Algebra	Say it with Symbols

Procedure All teachers attended a two-day, online, professional development workshop to introduce them to the research-based principles and implications for instructional materials and practice. During these sessions, teachers worked as groups and in pairs to plan instruction for the treatment units. Teachers administered pre-tests for both study units immediately following the professional development. Teachers then taught CMP2 in their normal curriculum order, administering post-tests immediately upon completion of each study unit, treatment and control. Teachers completed weekly instructional logs for both the treatment and control units, in which they described their implementation of the unit, including any application of the research-based principles. This enabled researchers to measure fidelity of implementation and estimate the achieved relative strength (Hulleman & Cordray, 2009) of the treatment intervention by comparing the degree to which teachers implemented the research-based principles in their treatment vs. their control units.

Measures

Researcher-developed assessments were used to evaluate student learning. The content of each curriculum unit was carefully mapped in order to assess the content areas, skills, and contexts presented to students. The same mapping was performed on the assessments to ensure they were well-aligned to the curriculum unit. All items were field-tested to establish reliability. Assessments included approximately 16 multiple-choice items and two open-ended items. Approximately half of the items were derived from existing CMP2 materials, and the remaining items were taken from state, national and international standardized tests.

For each unit, two test forms were created with approximately half of the multiple-choice and both open-ended items as linking items. Test forms were randomly assigned by class such that half of the classes took form A for pretest and form B for posttest, and the other half of the classes took form B for pretest and form A for posttest. Open-ended items were scored by trained raters using a standardized holistic rubric. Researchers computed weighted kappas to measure both intra-rater and inter-rater reliability. Intra-rater reliability ranged from 0.90 to 0.99. Inter-rater reliability ranged from 0.83 to 0.94.

Data Analysis

Item response theory (IRT) was used to equate the test scores across forms (Cook & Eignor, 1991). A partial credit model was used to generate item parameters, scale scores[1] for students, and assessment reliabilities, which ranged from 0.55-0.74 on pre-test and 0.77-0.82 on post-test.

ANCOVA models were used to estimate the treatment effects, controlling for pre-test scale scores and socioeconomic status. ANCOVAs for each unit were performed on students with complete demographic information and who completed both the pre-test and the post-test for that unit. The ANCOVA sample for each unit is shown in Table 2—the ANCOVA samples do not differ statistically from the full sample in their demographic makeup.

Table 2: ANCOVA sample for each unit.

Unit	Control	Treatment
	6th Grade	
Covering and Surrounding	481	384
Bits and Pieces III	431	496
	8th Grade	
Shapes of Algebra	349	371
Say it with Symbols	386	435

Results

6th Grade

To provide context for the IRT scale scores, traditional descriptive statistics for the overall change in students' performance from pre-test to post-test are shown in Table 3. Students made meaningful gains from pre-test to post-test on both units.

Table 3: Mean 6th grade assessment performance, all students

Test section	Pre-test	Post-test
	Covering and Surrounding	
Multiple-choice % correct	41.2% (SD = 16.1%)	61.0% (SD = 21.1%)
Open-ended out of 7 points	1.3 (SD = 0.9)	2.2 (SD = 1.3)
	Bits and Pieces III	
Multiple-choice % correct	47.2% (SD = 20.9%)	65.1% (SD = 23.1%)
Open-ended out of 8 points	1.6 (SD = 1.7)	2.7 (SD = 2.1)

Post-test scale scores for both 6th grade units, holding pre-test scores constant, are shown in Figure 3.

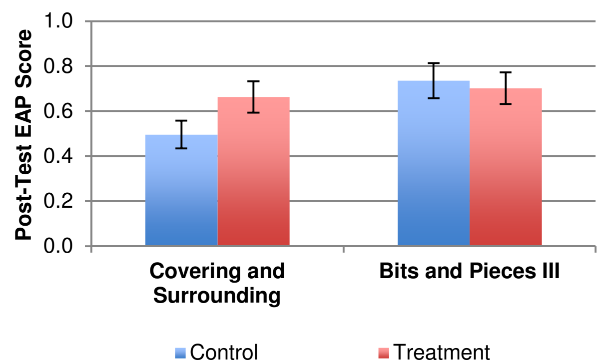


Figure 3: Post-test IRT scale scores for the 6th grade units. Error bars represent ±2 standard error.

ANCOVA results are presented in Table 4 (mean-square error is shown in parentheses). Pre-test was significantly associated with post-test scores in both units.

Table 4: 6th grade ANCOVA results

Source	*df*	*F*	*p*	Partial η^2
		Covering and Surrounding		
Pre-test	1	288.93	< .001	0.251
Socioec. disadv.	1	54.50	< .001	0.060
Treatment	1	12.78	< .001	0.015
Error	861	(0.46)		
		Bits and Pieces III		
Pre-test	1	352.13	< .001	0.276
Socioec. disadv.	1	80.31	< .001	0.080
Treatment	1	0.40	.528	< 0.001
Error	923	(0.62)		

There was a statistically significant main effect of socioeconomic status in both units, with students who are not socioeconomically disadvantaged performing better than students who are. There was also a statistically significant

[1] Ability estimates were generated using expected a posteriori scoring.

effect of treatment in *Covering and Surrounding*, with treatment out-performing control, but no statistically-different differences between groups for *Bits and Pieces III*.

8th Grade

Traditional descriptive statistics illustrating the overall change in students' performance from pre-test to post-test is shown in Table 5.

Table 5: Mean 8th grade assessment performance, all students

Test section	Pre-test	Post-test
	Shapes of Algebra	
Multiple-choice % correct	37.2% (SD = 15.1%)	51.4% (SD = 20.6%)
Open-ended out of 8 points	1.00 (SD = 1.6)	2.8 (SD = 2.5)
	Say it with Symbols	
Multiple-choice % correct	43.2% (SD = 17.3%)	55.0% (SD = 21.4%)
Open-ended out of 8 points	1.5 (SD = 1.8)	2.7 (SD = 2.4)

Again, students made significant gains from pre-test to post-test on both units, although the 8th grade assessments were relatively more difficult than the 6th grade assessments. Post-test scale scores for both 8th grade units, holding pre-test scores constant, are shown in Figure 4.

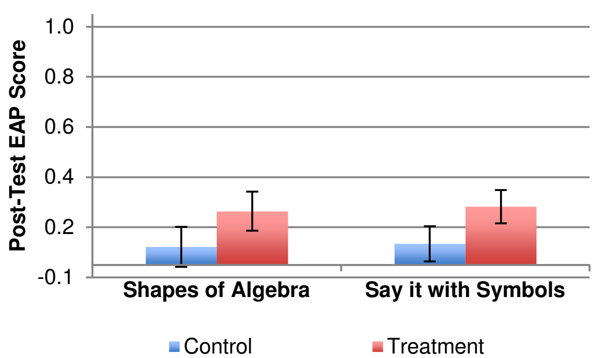


Figure 4: Post-test IRT scale scores for the 8th grade units. Error bars represent ±2 standard error.

ANCOVA results are presented in Table 6 (mean-square error is shown in parentheses). As in the 6th grade units, pre-test was significantly associated with post-test scores in both units and there was also a statistically significant main effect of socioeconomic status in both units, with students who are not socioeconomically disadvantaged performing better than students who are. Statistically significant effects of treatment were found for both units, with treatment out-performing control. Effect sizes for *Shapes of Algebra* and *Say it with Symbols* are similar.

Table 6: 8th grade ANCOVA results

Source	*df*	*F*	*p*	Partial η^2
		Shapes of Algebra		
Pre-test	1	157.57	< .001	0.180
Socioec. disadv.	1	34.09	< .001	0.045
Treatment	1	6.58	.011	0.009
Error	716	(0.55)		
		Say it with Symbols		
Pre-test	1	434.39	<.001	0.347
Socioec. disadv.	1	26.99	< .001	0.032
Treatment	1	9.72	.002	0.012
Error	817	(0.46)		

Discussion

Students demonstrated large learning gains for each unit, suggesting both versions of the CMP2 curriculum were effective. Further, three of the four units in this study produced statistically significant effects of the treatment manipulation. That is, the treatment materials produced an additional boost to student learning over and above the existing materials. Why were some treatment units more effective than others? One possible explanation for this differential effect is that *Covering and Surrounding* and *Shapes of Algebra*, two of the three units showing a statistically-significant treatment effect, are both more spatially-oriented units. *Covering and Surrounding* addresses area and perimeter and *Shapes of Algebra* emphasizes coordinate geometry. While *Say it with Symbols* focuses on expressions and equations, students must link symbolic representations to graphs and other figures. In contrast, *Bits and Pieces III* more strongly emphasizes symbolic and tabular representations. The more figure-oriented units may allow for a more potent treatment, as the first cognitive principle directly relates to increasing the coherence in visual representations.

The current findings suggest that research-based instructional strategies can be applied synergistically to improve student outcomes in authentic classroom settings. These findings are of particular importance as the vast majority of existing research investigates design principles in highly controlled (and artificial) lab-based studies. Ongoing analyses will provide further insight into the nature of the treatment effects. We are currently analyzing teachers' instructional logs in order to better understand when and how they implemented the cognitive principles in their teaching practice, aside from using the revised student books. We would expect larger learning gains for students when teachers integrated the principles into classroom practice in addition to giving students the revised books. Additional studies are also being carried out at the sites of the partner institutions to investigate the effects of the additive effects of the principles. The Math Center team is also conducting a cluster-randomized trial of revisions to the entire 7th grade CMP curriculum, taking place during the 2012-2013 and 2013-2014 academic years. If the effects of the principles are cumulative throughout the school year, we would expect greater differences in performance

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C100024 to WestEd. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We acknowledge the extensive contributions of our research partners of the National Center for Cognition and Mathematics Instruction, particularly the co-PIs of the center: Martha Alibali, Julie Booth, Susan Goldman, Neil Heffernan, Ken Koedinger, Mitchell Nathan, James Pellegrino, and the leads of the math review, production, and data collection teams: Shandy Hauk, Kimberly Vivani, and Kathleen Lepori.

References

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, R. C., & Mayer, R. E. (2003). e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning. San Francisco, California: Jossey-Bass.

Cook, L. L. & Eignor, D. R. (1991). IRT equating methods. *Educational Measurement: Issues and Practice, 10*(3), 37-45.

Goldstein, H. (1987). Multilevel models in educational and social research. London: Griffin.

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), 414-434.

Hulleman, C. S. & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness, 2*, 88-110.

Karpicke, J. D., & Roediger, H. L., III. (2010). Is expanding retrieval a superior method for learning text materials? *Memory & Cognition, 38*, 116-124.

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

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358-368.

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

Paas, F. G. W. C. & Van Merrienboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122-133

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning: IES practice guide* (NCER 2007-2004). Washington, DC: National Center for Education Research.

Pavlik, P. I., Jr., Presson, N., Dozzi, G., Wu, S.-M., MacWhinney, B., & Koedinger, K. (2007). The FaCT (fact and concept) system: A new tool linking cognitive science with educators. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society*. Nashville, TN, USA.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development, 77(1),* 1–15.

Roediger, H.L., III, & Karpicke, J.D. (2006). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, Calif.: Sage Publications.

Rohrer, D. (2009). Avoidance of overlearning characterises the spacing effect. *European Journal of Cognitive Psychology, 21*(7), 1001-1012.

Roy, M., & Chi, M. T. H. (2005). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning*, (pp. 271–286). New York: Cambridge University Press.

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York: Cambridge University.

Siegler, R.S., & Chen, Z. (2008). Differentiation and integration: Guiding principles for analyzing cognitive change. *Developmental Science, 11,* 433–448.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59–89.

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychological Review, 68*(4), 229-247.

Zhu, X., & Simon, H.A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, *4,* 137-166.