ISSN: 2690-0726

Research Article

Journal of Educational and Psychological Research

Bridging Cognitive Science and Teaching Practice: A Proof of Concept Study of the Cognitive Instructional Techniques (CIT) Observation Instrument

Elida V Laski*

Associate Professor, Lynch School of Education & Human Development, Boston College, United States

***Corresponding Author**
Elida V Laski, Associate Professor, Lynch School of Education & Human Development, Boston College, United States. E-mail: elida.laski@bc.edu

Submitted: 05 Jan 2023 ; **Accepted**: 07 Jan 2023 ; **Published**: 07 Feb 2023

Citation: *Laski, E. V. (2023). Bridging Cognitive Science and Teaching Practice: A Proof of Concept Study of the Cognitive Instructional Techniques (CIT) Observation Instrument. J Edu Psyc Res, 5(1), 594-600.*

Abstract

This paper presents the Cognitive Instructional Techniques (CIT) observation instrument – a novel tool to assess the degree to which teaching practice is congruent with instructional techniques that have emerged from cognitive science. The paper includes a description of CIT, the process by which it was developed, and initial findings related to its use. CIT includes a 0-4 rating scale (0 = not observed to 4 = high congruence) for eight instructional techniques: Anchoring, Guided Instruction, Multiple Representations, Quizzing, Self-Explanation, Signalling, Variable Practice, and Worked Examples. The instrument was used to code randomly selected videotaped 4th and 5th grade mathematics lessons in the USA (N = 42) selected from the Measures of Effective Teaching Project database (Bill and Melinda Gates Foundation, 2018). Results of this proof-of-concept study indicated that the CIT instrument captures variability as well as stability in teaching practice. Overall, teaching practice was found to be congruent with some instructional techniques advocated for by cognitive researchers (e.g., testing), but not others (e.g., self-explanation). For instance, half of the instructional lessons in the sample did not use multiple external representations to explain math concepts and less than a quarter demonstrated the use of multiple representations in a manner consistent with what research has identified as most effective. This initial study indicated the potential usefulness of the CIT instrument in determining the extent to which instructional practice is congruent with ideas from cognitive science. The ways in which CIT could create further intersections between cognitive science and education are discussed.

Keywords: Learning Sciences, Mathematics Education, Multiple Representations, Fractions, Teacher Education

Introduction

Increasing the academic achievement of students is critical for a nations' economy and individual citizens' college, career, and economic opportunities [1, 2]. Given that teachers' instructional practice is one of the most important factors contributing to students' achievement, identifying the extent to which it includes effective strategies is central to determining targets for improvement [3-5].

Over the past two decades, there has been increased interest in the application of research about learning processes to teaching practice [6-9]. A central goal of cognitive research is to identify the situations and processes that lead to more efficient, robust, and durable learning. Learning scientists have built upon basic research about human learning and memory to identify and test specific approaches for facilitating learning [10-12]. These approaches describe ways of structuring and presenting information as well as ways to promote students' cognitive processing to enhance learning.

While numerous review articles and books have been written describing the range of cognitive learning strategies that have been found to facilitate learning in domains ranging from reading, science, and mathematics, the extent to which they are presently used in teaching practice is largely unknown [12-16]. A survey of teacher educators suggested that some essential ideas related to cognitive development and the developmental progression of skills are commonly part of conversations in both fields, but other basic research findings from experimental studies of learning and memory are less often incorporated into teacher preparation [17]. Thus, the present study aimed to develop a classroom observation instrument—Cognitive Instructional Techniques (CIT)—that could be used to examine this issue.

Cognitive Instructional Techniques (CIT) Observation Instrument

The goal of our work was to develop an instrument grounded in the results of empirical studies of learning; thus, we employed primarily an etic—top-down—approach in which the literature from cognitive psychology drove the development of categories and descriptions of the instructional techniques. The entire process, however, was a multi-step and iterative process that also involved consideration of the teaching practices themselves.

The first step in the development process was to conduct a literature review to identify instructional techniques for which there was robust empirical evidence from cognitive research. To select the strategies that would be included in the observation tool we considered (1) amount of empirical evidence in support of the approach and extent to which experts have made arguments calling for its application to practice, (2) likelihood that an approach would be observable in isolated classroom lessons, and (3) likelihood instances of the technique would be possible to identify and evaluate. Based on our own review of the literature as well as of influential reviews written by experts in the field of cognitive science, we selected to include eight instructional techniques: anchoring; guided instruction; multiple representations; variable practice; quizzing, self-explanation; signaling; and worked examples [10-12]. Three—guided instruction, quizzing, and signaling—were comprised of two sub-techniques each. Table 1 lists the instructional techniques, presents their operational definitions, and includes examples of empirical work that provides evidence for the technique.

The second step involved developing descriptions of qualitative differences in the extent to which a technique might be reflected in instruction in terms of four ordinal categories from no congruence to high congruence. For each technique, we first outlined the no congruence and high congruence categories. The no congruence category described instruction in which the approach was completely absent. The high congruence category described instruction that included the primary aspect of the technique as well as any parameters that have been found to lead to the greatest learning in lab studies. The remaining categories – low and moderate congruence – required the greatest inferences. For these categories, we inferred from the literature which parameters and what aspects of those parameters were most important and then used a logical sequence.

The final steps in the development process involved iterative refinement of the instrument. First, the instrument was vetted by four experts in the field of cognitive science. They provided comments about the extent to which the instrument captured the essence of each technique and reflected the empirical evidence. Based on their comments, slight modifications were made to the operational definitions and qualitative descriptors. Second, three raters used the instrument to code between two and three instructional lessons at a time, met to discuss disagreements, and revised the qualitative descriptors based on the discussions, as necessary. The process proceeded for several cycles until the inter-rater agreement reached κ = .83.

Table 1: Overview of Cognitive Instructional Techniques (CIT) Observation Instrument

Instructional Technique	***Operational Definition***	***Illustrative Supporting Evidence***
Anchoring a. Familiar knowledge b. Real-world problems	Concepts and skills are embedded in a practical "real world" problem, familiar situation, or are connected to familiar knowledge a. *Familiar Knowledge to Support Understanding* – making connections to the real world or to prior knowledge for the purpose of explaining a math concept b. Solving Real-World Problems – embedding content in real-world contexts or problems	Barron & Darling-Hammond (2010) Cognition and Technology Group at Vanderbilt (1992) [18, 19]
Guided Instruction a. Whole group b. Individual	Guidance (e.g., modeling, coaching, scaffolding) is provided which elicits, extends, and clarifies student thinking a. *Whole Group Guidance* – attempts to engage the whole class and support students' understanding b. Individual Guidance – attempts to support individual students' understanding through one-to-one interactions	Kirschner et al. (2006) Lee & Anderson (2013) [20, 21]
Multiple Representations	Use of multiple representations to explain and illustrate concepts or problem-solving	Ainsworth (1999) Cleaves (2015) [22, 23]
Quizzing a. Testing b. Spacing	Use of a quiz or questions to promote students' retrieval of information a. *Testing* – involves how accountable all students are for retrieving information b. *Spacing* – involves the amount of time between when students learned the material and are asked to retrieve it	Kang et al. (2007) Rohrer & Taylor, 2007 [24, 25]
Self-Explanation	Opportunities are provided for students to explain their thinking and strategies (either correct or incorrect)	Berry (1983) Chi et al. (1994) [26, 27]
Signalling a. Cueing structure b. Guiding attention	Cues are provided that guide attention to relevant information and support the organization of information a. *Cueing Structure and Key Concepts* – Involves highlighting key ideas and concepts discussed in the lesson b. *Guiding Attention* – Involves verbal and visual cues that direct students' attention to important aspects of the lesson	Anderson (1995) Mautone & Mayer (2001) Meyer & Rice (1981) [28, 29, 30]
Variable Practice	Interleaves content, problem types, or strategies rather than blocking them	de Croock & van Merriënboer (2007) Taylor & Rohrer (2010) [31, 32]
Worked Examples	Includes worked examples that demonstrate problem solution paths	Catrambone (1995) Atkinson, R. Derry, S., Renkl, A. & Wortham, D., 2000 [33, 34]

Table 2 provides an example of the final coding criteria for each congruence level for the Multiple Representations technique.

Table 2: Example of Congruence Level Descriptors

Multiple Representations Technique			
No Observable Congruence [Score = 0]	*Low Congruence* [Score = 1]	*Moderate Congruence* [Score = 2]	*High Congruence* [Score = 3]
• Lesson uses only a single external representation (e.g., numerical symbols) to present or explain a concept	• Lesson involves two or more external representations to present or explain a concept (e.g., fractions & decimals; fractions & pie chart) • No conceptual explanation is provided about the relationship between the representations	• Lesson uses two or more external representations to present or explain a concept AND… • Explanation is provided about the relationship between the representations (e.g., both are used to explain the same underlying concept, how they are similar/different, how they complement each other)	• Lesson uses three or more external representations to present or explain a concept AND… • Explanation is provided about the relationship among the representations (e.g., both are used to explain the same underlying concept, how they are similar/different, how they complement each other) • Representations are presented simultaneously to allow immediate comparisons between them

The Present Study Goals

This study served as a proof-of-concept before formally evaluating the instrument's psychometric properties. It had two purposes. The first purpose was to examine whether the instrument captured variability in teaching practices across similar lessons. For this initial investigation, we constrained our sample to a narrow grade range (4th and 5th grade) and a particular content area (fractions and decimal lessons). The second purpose was to examine whether the instrument captured stability in teaching practices within individual teachers. For a subset of teachers included in the sample, we selected two videotaped lessons. This characteristic of our sample allowed us to compare these teachers' scores across the two lessons. Given that the format of lessons can vary depending on the content and students' familiarity with that content, it seemed unreasonable to expect a perfect correlation (i.e., mean difference score of 0) between any given teacher's lessons. However, we expected to find some stability in the congruence of teachers' instruction to the cognitive instructional techniques, which would be indicative of their general familiarity with the instructional technique.

Methodology

Data Source

The videos of teaching practice included in the present study were obtained from the Measures of Effective Teaching (MET) project. The MET data set includes videos of lessons that were recorded during the 2009-2010 and 2010-2011 school years. The lessons were conducted by a self-selected sample of teachers from six large school districts in the United States: Charlotte-Mecklenburg Schools, NC; Dallas Independent School District, TX; Denver Public Schools, CO; Hillsborough County Public Schools, FL.; Memphis City Schools, TN; and New York City Department of Education. While the data set is not nationally representative, it includes examples of teaching practice across a range of contexts.

Video Selection and Coding Procedure

For this initial study, we constrained the sample to videotaped fourth- or fifth grade lessons that involved instruction about fractions or decimals. Within the sample of videos that met these parameters, a total of 42 lessons were rated for 32 teachers (ten teachers were rated twice, for two separate lessons).

Ratings were based on 25 minutes of the video: the first 15 minutes, and minutes 25-35 or, if the video was shorter than 25 minutes, the full length of the lesson. This approach helped to account for variability in the length of the lessons. It was also consistent with approaches used with other teaching observation instruments [35].

Each lesson was scored by both the master coder and at least one other rater. Before raters viewed the videos, a research assistant not involved in video coding transcribed each video. The raters were instructed to consider both the transcript and the video of the lesson in determining their ratings. Raters first read through the transcript, noting instances that reflected the use of any of the techniques, then watched the video of the lesson. For each video session, raters assigned a score between 0 and 3 (0 = no congruence; 1 = low congruence; 2 = moderate congruence; 3 = high congruence) for each technique or sub-technique. Thus, each lesson was assigned 12 scores. Any discrepancies in scores were discussed, and a final score was based upon consensus.

The complete instrument, including coding instructions and criteria for each technique can be seen by contacting the author.

Results

Variability Across Teacher Practice

To examine whether the CIT instrument captured variability in teaching practices, we conducted descriptive analyses of the distribution of lessons' (a) overall scores across all eight techniques and (b) scores on each technique separately.

First, a total score for each lesson was calculated by summing across the 12 scores. The lowest possible score was 0 and the highest possible score was 36. Across the 42 lessons rated, the mean total CIT score was 15.86 (*SD* = 3.35), with the minimum score being 9 and the maximum score 23. An examination of the scores for each cognitive instructional technique indicated that this variability in total score was also reflected when examining teaching practices for individual techniques. The full range of congruence, from 0 (no congruence) to 3 (high congruence), was observed on half of the instructional techniques: Guided Instruction – Individual Guidance (*M* = 1.07, *SD* = .68), Multiple Representations (*M* = .74, *SD* = .86), Variable Practice (*M* = 1.24, *SD* = 85), Signalling – Structure and Key Concepts (*M* = 1.79, *SD* = .78), Signalling – Guiding Attention (*M* = 1.45, *SD* = .71), and Worked Examples (*M* = 1.62, *SD* = 62). For three of the cognitive instructional techniques, lesson's scores ranged between 1 and 3, with no lesson being rated as exhibiting no congruence: Guided Instruction – Whole Group (*M* = 1.36, *SD* = .62), Quizzing – Testing (*M* = 2.40, *SD* = .77), Quizzing – Spacing (*M* = 1.60, *SD* = .91). For two of the techniques, lessons' scores ranged between 0 and 2, with no lesson being rated as exhibiting high congruence: Anchoring – Familiar Knowledge (*M* = 1.07, *SD* = .68) and Self-Explanation (*M* = .69, *SD* = .78). On one of the strategies – Anchoring – Real World Problems – no teacher scored above low-congruence (*M* = .45, *SD* = .50).

Analysis of the percentage of individual lessons that fell within each congruence level provided further evidence of variability in teaching practice. In addition, this analysis also pointed to particular strengths and potential areas of improvement in teaching practice. For example, as shown in Table 3, the majority of lessons (57%) received a high congruence score on Quizzing – Testing; whereas, only 2% of lessons did so for Multiple Representations.

Table 3: Percentage of Lessons Scored at Each Congruence Level for Each Instructional Technique in the Cognitive Instructional Techniques (CIT) Instrument

	***No Observable Congruence* [Score= 0]**	***Low Congruence* [Score = 1]**	***Moderate Congruence* [Score = 2]**	***High Congruence* [Score = 3]**
Anchoring: Familiar Knowledge	19	55	26	0
Anchoring: Real World Problems	55	45	0	0
Guided Instruction: Whole Group	0	71	21	7
Guided Instruction: Individual Guidance	2	60	29	10
Multiple Representations	50	29	19	2
Quizzing: Testing Effect	0	17	26	57
Quizzing: Spacing Effect	0	69	02	29
Self-Explanation	50	31	19	0
Signalling: Structure & Key Concepts	5	29	50	17
Signalling: Guiding Attention	10	38	50	2
Variable Practice	19	45	29	7
Worked Examples	2	38	55	5

Stability Within Teachers

To examine whether the CIT instrument captures stability within the teaching practices of particular teachers, raters scored two lessons for ten different teachers. Lessons were scored at least two weeks apart to decrease rater bias. After calculating the total CIT score on each of the teachers' lessons, we calculated the mean absolute difference between the scores. For example, if a teacher received a CIT total score of 18 on one lesson and 22 on the second lesson the mean absolute difference in scores would be: |18-22| = 4. Finally, we conducted a one-sample *t*-test on the mean absolute difference between CIT scores, with the null hypothesis being a mean 2-point difference. Given that the format of lessons can vary depending on the content and student's familiarity with that content, it seemed unreasonable to expect a perfect correlation (i.e., mean difference score of 0) between lessons. The analysis indicated that on average, the mean difference between the scores on teachers' lessons was not greater than two points, indicating good test-retest reliability: $t(9) = 1.77$, $p = .111$.

Conclusions

Summary

For the most part, studies testing the efficacy of cognitive instructional techniques, such as the use of testing and spaced practice, have been conducted in laboratory settings and, while it is argued that these approaches are often not incorporated into teaching practice, there is little evidence of the veracity of this claim [36]. The Cognitive Instructional Techniques (CIT) instrument provides an operational definition of eight prominent instructional techniques that have emerged from the Learning Sciences as well as qualitative descriptors of various levels of congruence to those techniques. When applied to the observation of teaching practice, it can provide information about the presence of the cognitive instructional techniques in typical classroom instruction.

The present results suggested that the instrument captures variability in teachers' practice as well as stability in teachers' practice from lesson to lesson. All of the instructional techniques were observed at either low or moderate congruence in half or more of the lessons. Across 42 lessons rated, there was a wide spread of scores, indicating that while some lessons demonstrate little congruence with cognitive instructional techniques others exhibit more. Further, teachers CIT scores across two lessons were not more than an average of two (out of a possible 36) points different. In other words, a teacher who received a higher CIT score on one lesson was also likely to receive a higher CIT score on another lesson. This pattern of results suggests that teachers may be receiving some training that promotes instruction congruent with some instructional techniques advocated for by cognitive researchers but that the integration of them into classroom practice varies across teachers.

Three of the instructional techniques were observed in all lessons, at low congruence or greater: Guided Instruction: Whole Group, Quizzing: Testing Effect, and Quizzing: Spaced Effect. This finding was noteworthy for the use of quizzing, both in terms of using testing to encourage students to retrieve information and spacing practice so that students have to retrieve information taught previously. The extent to which quizzing was observed in these elementary mathematics lessons to seems to contradict claims that testing is not used sufficiently in classroom practice [36]. Thus, the present results offer an example of where there may be greater overlap between cognitive science recommendations and teaching practice than some authors have posited.

On the other hand, none of the observed lessons exhibited the characteristics of high congruence with Anchoring and Self-Explanation. At present it is unclear how to interpret this result. It may be that it is unrealistically difficult to incorporate the instructional techniques as operationalized as the ideal in the cognitive literature in the context of a dynamic classroom setting. Alternatively, it may be that research about these techniques and the nuances related to high congruence have not been adequately communicated to teachers, thus, suggesting an area of potential teacher professional development. Further use of the instrument to observe a broader range of lessons and discussion with experts in the field of teacher education may help to discern between these possibilities.

Limitations and Future Directions

We believe the current proof-of-concept study suggests that the Cognitive Instructional Techniques observation instrument may be a useful tool for examining the integration of cognitive science instructional techniques in typical classroom practice. Its use could promote communication and collaboration between the fields of education and cognitive science. In terms of research, the instrument could prompt collaborations that examine which features of classroom dynamics promote or impede use of a given instructional technique. It could also provoke controlled studies of the effect of different levels of implementation of the instructional techniques during authentic teaching may have on student learning. In terms of educational practice, results might inform teacher preparation; teacher educators might consider these cognitive instructional techniques, particularly those on which a small percentage of lessons reflect high congruence, to a greater extent in their training of teachers. Further, examples of lessons that reflect varying levels of congruence with the cognitive instructional techniques could be used in teacher education to promote discussion about and analysis of practice.

Before moving toward widespread use of the CIT instrument, however, further research and development is needed. The current study coded only a small sample of lessons and conducted only cursory exploratory analyses. More extensive and formal studies of the instruments' psychometric properties ought to be conducted. Factor analyses should be conducted to examine relations between techniques and, perhaps, winnow the instrument. A Rasch analysis could be conducted to determine whether the instrument fits a unidimensional model, which would suggest that integration of the cognitive instructional techniques at a higher congruence may be a general construct for higher quality instruction. A generalizability study could be conducted to further refine the observation procedures and identify important parameters, such as the optimal number of observations and raters, that increase confidence in the scores [37]. These future directions would enhance the quality and utility of the Cognitive Instructional Techniques observation instrument. Given the promising results of this proof-of-concept study, we believe that continued effort and time toward developing CIT would be worthwhile investment [38-41].

Acknowledgments

Thanks to the various members of the Thinking & Learning Lab for assistance transcribing, coding, and hours spent discussing the instrument: Anna Ermakova, Karina Halloran, and Melissa Paz. Thanks also to the generosity of colleagues who vetted and provided comments on earlier versions of the CIT instrument: Julie Booth, Sharon Carver, David Daniel, and Nicole McNeil. This research was supported in part by a Boston College Research Incentive Grant awarded to the author.

References

1. Feinstein, L., & Bynner, J. (2004). The importance of cognitive Development in middle childhood for adulthood socioeconomic status, mental health, and problem behavior. *Child Development, 75(5),* 1329-1339.
2. Sadler, P. M., & Tai, R. H. (2007). The two high-school pillars supporting college science. *Science, 317(5837),* 457-458.
3. Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30(3),* 466-479.
4. Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica,*

73(2), 417-458.

5. Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record, 104(8),* 1525-1567.
6. National research Council, Cocking, R. R., Brown, A. L., & Bransford, J. D. (1999). How people learn: Brain, mind, experience, and school. *National Academy*.
7. Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education, 61(1-2)* 35-47.
8. National Research Council, & Up, A. I. (2001). Helping Children Learn Mathematics. Mathematics Learning Study Committee. J. Kilpatrick, J. Swafford, and B. Findell, eds. Washington, DC: *National Academy Press*.
9. Newcombe, N. S., Ambady, N., Eccles, J., Gomez, L., Klahr, D., Linn, M., ... & Mix, K. (2009). Psychology's role in mathematics and science education. *American Psychologist, 64(6*) 538-550.
10. Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14(1*) 4-58.
11. Graesser, A., Hakel, M., & Halpern, D. (2007). Life Long Learning at Work and at Home… and at the 19th Annual APS Convention. *APS Observer, 20.*
12. Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. *National Center for Education Research.*
13. Klahr, D., & Li, J. (2005). Cognitive research and elementary science instruction: From the laboratory, to the classroom, and back. *Journal of Science Education and Technology, 14(2),* 217-238.
14. National Research Council. (2005). How Students Learn: History, Mathematics and Science in the Classroom. Committee on How People Learn, A Targeted Report for Teachers. *Center for Studies on Behavior and Development, National Research Council.*
15. Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2(2),* 31-74.
16. Siegler, R. S. (2003). Implications of cognitive science research for mathematics education. In Kilpatrick, J., Martin, W. G., & Schifter, D. E. (Eds.), A Research Companion to Strategies and Standards for School Mathematics (pp. 119-233). Reston, VA: *National Council of Teachers of Mathematics.*
17. Laski, E. V., Reeves, T. D., Ganley, C. M., & Mitchell, R. (2013). Mathematics teacher educators' perceptions and use of cognitive research. *Mind, Brain, and Education, 7(1),* 63-74.
18. Barron, B., & Darling-Hammond, L. (2010). Prospects and challenges for inquiry-based approaches to learning. *The Nature of Learning: Using Research to Inspire Practice*, 199-225.
19. Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist, 27(3),* 291-315.
20. Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructuvist, Discovery, Problem-Based, Experimental, and Inquiry-Based Teaching. *Educational Psychologist, 42(2)*, 75-86.
21. Lee, H. S., & Anderson, J. R. (2013). Student learning: What has instruction got to do with it?. *Annual Review of Psychology, 64,* 445-469.
22. Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33(2-3),* 131-152.
23. Cleaves, W. P. (2008). Promoting mathematics accessibility through multiple representations jigsaws. *Mathematics Teaching in the Middle School, 13(8),* 446-452.
24. Kang, S. H., McDermott, K. B., & Roediger III, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19(4-5),* 528-558.
25. Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35(6),* 481-498.
26. Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *The Quarterly Journal of Experimental Psychology Section A, 35(1),* 39-49.
27. Chi, M. T., De Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18(3),* 439-477.
28. Anderson, J. R. (1995). Memory and Learning: An Integrated Approach.
29. Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93(2)* 377-389.
30. Meyer, B. J., & Rice, G. E. (1981). Information recalled from prose by young, middle, and old adult readers. *Experimental Aging Research, 7(3),* 253-268.
31. de Croock, M. B., & Van Merriënboer, J. J. (2007). Paradoxical effects of information presentation formats and contextual interference on transfer of a complex cognitive skill. *Computers in Human Behavior, 23(4),* 1740-1761.
32. Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24(6),* 837-848.
33. Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87(1),* 5-17.
34. Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70(2)* 181-214.
35. Kane, T. (2012). Capturing the dimensions of effective teaching: student achievement gains, student surveys, and classroom observations. *Education Next, 12(4),* 34-42.
36. Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39(5),* 406-412.
37. Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation

systems and a case for the generalizability study. *Educational Researcher, 41(2),* 56-64.

38. Hemsley-Brown, J., & Sharp, C. (2003). The use of research to improve professional practice: A systematic review of the literature. *Oxford Review of Education,* 29(4), 449-471.
39. Reif, F. (2008). Applying cognitive science to education: Thinking and learning in scientific and other complex domains. *MIT Press*.
40. Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17(3),* 249-255.
41. Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educational Psychology Review, 24(3),* 355-367.