Macrothink Institute™

Journal of Studies in Education
ISSN 2162-6952
2013, Vol. 3, No. 1

Is Learning Enhanced by Conceptual Integration During Practice?

Maura Pilotti (Corresponding author)

Ashford University

Linda LaGrange, Tyler Broderick, Edward Martinez, Salif Mahamane & Jennifer Almand

New Mexico Highlands University

Received: Sep. 23, 2012 Accepted: November 14, 2012 Published: February 1, 2013

doi:10.5296/jse.v3i1.2441 URL: http://dx.doi.org/10.5296/jse.v3i1.2441

Abstract

In this study, we examined whether retrieval practice promoting conceptual integration can improve test performance and self-assessment. Prior to testing, introductory psychology students received retrieval practice through their weekly lab sessions. Short-answer questions for practice exercises emphasized either single concepts or conceptual integration, whereas those for tests concerned individual concepts. During the first half of the semester, test performance was higher following practice focused on concepts than practice focused on conceptual integration. No difference existed during the second half. Performance was correlated with estimates before testing when practice emphasized individual concepts, but not conceptual integration. The transfer-appropriate processing framework offers an account for these data by suggesting that emphasis on conceptual integration might have fostered discrepancy between conceptual operations used for practice and those utilized for testing, affecting test performance.

Keywords: Learning, Memory, Practice, Retrieval, Field Study

Acquisition of knowledge in introductory college courses is especially arduous for two main reasons: learners must not only develop internal representations of newly acquired information, but also connect these representations with one another and with pre-existing knowledge in long-term memory (Dagenbach, Horst, & Carr, 1990). Research suggests that these links, which are essential to the preservation of new information in long-term memory and its retrieval when circumstances require, are initially weak (Carr & Dagenbach, 1994; Dagenbach et al., 1990). In laboratory experiments, conceptual integration has been found to strengthen connectivity among memory representations, thereby facilitating retrieval from long-term memory and analytical activities on the retrieved information (see Anderson, 1987). Conceptual integration, however, is not a passive process for which mere repeated exposure to a domain of knowledge during learning is sufficient to create a network of robust connections among the domain's concepts. Rather, it is an effortful activity that relies heavily on the learner's cooperation (Conway, Gardiner, Perfect, Anderson, & Cohen, 1997).

Evidence collected from simulated classroom environments indicates that practice tests can improve students' subsequent test performance relative to reviewing (Pilotti, Chodorow, & Petrov, 2009). Practice tests have also been found to improve performance even when test questions probe material that is not tested during practice but is conceptually related to it (Anderson & Biddle, 1975; Chan, McDermott, & Roediger, 2006; Pilotti et al., 2009; Roediger, & Karpicke, 2006a). Therefore, existing evidence suggests that testing, which has been traditionally used for assessment purposes, can serve as both memory enhancer and learning tool. The act of answering practice questions can serve as memory enhancer largely because it induces the retrieval of newly acquired information (i.e., offers retrieval practice opportunities), thereby consolidating memory traces (see Karpicke, 2012; Karpicke & Smith, 2012; Mulligan, 2002; Slamecka & Graf, 1978). Answering practice questions can also activate conceptually related information, consolidating the links between tested and related information, and thus serving as learning tool (Hamaker, 1986).

Practice tests can also improve test performance by serving as self-assessment tools. Self-assessment during learning allows students to determine what has been mastered and what needs additional study (Balch, 1998; Dunning, Johnson, Ehrlinger, & Kruger, 2003; Hacker, Bol, Horgan, & Rakow, 2000; Maki & Serra, 1992; Zechmeister & Shaughnessy, 1980). Research has shown that students overestimate their test preparedness (Dunning, Heath, & Suls, 2004) and that overestimation tends to be magnified in students with low or average academic abilities (Balch, 1992; Zabrucky, Agler, & Moore, 2009). Students' tendency to rely on their subjective familiarity with a domain of knowledge is generally believed to result in poor self-assessment and, subsequently, in less-than-stellar test performance (Balch, 1998; Glenberg & Epstein, 1987; Glenberg, Sanocki, Epstein, & Morris, 1987; Maki & Serra, 1992). Practice tests have been found to enhance self-assessment accuracy by reducing students' emphasis on domain familiarity (Maki & Serra, 1992; Maki, Foley, Kajer, Thompson, & Willert, 1990).

The key premise upon which the present investigation relies is that learning does not merely necessitate that recently acquired information be passively added to long-term memory, but rather that this information be actively integrated (Dagenbach et al., 1990). This investigation

asks whether retrieval practice (Chan et al., 2006; Karpicke & Smith, 2012) explicitly promoting conceptual integration (i.e., connectivity among memory representations, including recently acquired concepts and pre-existing knowledge) can improve learning over retrieval practice focused on individual concepts. It also asks whether retrieval practice can improve self-assessment (i.e., the ability to appraise what one knows about to-be-tested material), another critical aspect of learning (Dunning et al., 2004).

If answering practice questions activates associated information via automatic spreading activation, then practice tests can be expected to promote conceptual integration when questions directly focus students' attention on the relationships between/among newly acquired concepts. This prediction is supported by evidence illustrating that spreading activation is sensitive to attention (Naccache, Blandin, & Dehaene, 2002). It is reasonable to predict that if practice tests promote conceptual integration by focusing students' attention on associations in long-term memory, then practice tests may offer students reasonably accurate insights into their ability to retrieve such concepts during actual tests, thereby improving self-assessment before tests. Whether retrieval practice focused on conceptual integration improves test performance relative to retrieval practice focused on individual concepts may depend on the extent to which attention to memory associations improves retrieval by increasing the accessibility of concepts above and beyond the level induced by automatic spreading activation.

Method

This investigation involved two phases chronologically organized. Phase 1 was devoted to the development of short-answer questions that not only covered either isolated concepts or relationships between concepts, but also were related to students' pre-existing knowledge. The choice of short-answer questions was based on evidence suggesting that they can be effective retention tools when used for retrieval practice purposes (Butler & Roediger, 2007; Kang, McDermott, & Roediger, 2007), while their scoring is only moderately labor intensive, a property useful in large introductory courses. Questions were tailored to students' pre-existing knowledge (including their sociocultural, economic, and physical environments) to ensure that the content of the questions could be linked to familiar information in long-term memory. Two semesters were devoted to developing, testing, and revising pre-test and test materials, all designed to refer to the content of a textbook (Myers, 2009).

Phase 2 was devoted to the implementation of pedagogical changes and the evaluation of outcomes. In this phase, participants were 200 students who enrolled in one of two sections of an introductory psychology course and in one of the 10 labs associated with it. The course involved two 50-min lectures, during which individual concepts from a textbook chapter were illustrated and examples offered, and one 50-min lab during which retrieval practice was implemented. Labs were assigned to one of five teaching assistants, each covering a lab with instruction focused on retrieval of single concepts (SC) and a lab with instruction devoted to retrieval of concepts and their links (LC). Consistency dictated that lectures be taught by one instructor using the same textbook (Myers, 2009). For retrieval practice to serve as a learning and self-assessment tool, feedback regarding students' answers was provided during labs immediately and approximately one week after responses to homework and practice questions

had been scored (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008; Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). During Phase 2, four activities were undertaken for each of the topics (chapters) covered by the course:

(1) *Assessment of pre-existing knowledge*, which consisted of pre-tests assessing conceptual knowledge prior to in-class coverage.

(2) *Coverage of topic* in class (i.e., lectures).

(3) *Retrieval practice* by means of lab activities administered after the coverage of a given topic was completed in the lecture portion of the course. Labs were randomly assigned to one of two retrieval practice regiments: In the SC condition, practice questions during labs and homework questions focused on students' knowledge of specific concepts, whereas in the LC condition, questions focused on students' understanding of the relationships between concepts. For instance, within the topic of memory, a SC question would present a concrete example of either recall or recognition memory, and then ask students to name and define the concept illustrated in the question (e.g., 'Luis loves to tell stories about his experiences. When people ask him about his past, he is able to remember many stories without having to look at old pictures or read his diary. Name and define the type of retrieval he uses.' or 'Abigail is a scuba diver. She has a hard time remembering the names of fish without actually seeing them in the water. Pictures or videos can also help her remember the correct names. Indeed, she can name a fish almost immediately after she sees it. Name and explain the type of retrieval she uses'). In contrast, a LC question would offer a concrete illustration of both concepts and then ask students to name the concepts and explain their similarities and differences (e.g., Mega and Paul learned many Spanish vocabulary words years ago in high school. Last week, they met a couple of tourists from Spain. It was hard at first, but Paul and Megan joined forces and spent a pleasant evening with the couple. Paul was able to understand most of the words the tourists spoke. Megan managed to think of some of the words she learned in high school and then used them in sentences. Identify recall and recognition in this example and define both types of retrieval by mentioning what they have in common and what makes them different').

(4) *Knowledge assessment and self-assessment after practice* by means of in-class tests administered after activities 1-3. Each test included questions not previously used. Immediately before and after testing, students were asked to estimate their performance on a scale from 0 to 100. It is important to note that each test included half of the concepts covered during the retrieval practice phase of the investigation.

Results

The study entailed four dependent measures for each topic covered in the course: pre-test scores, test scores, and estimates before and after testing. Answers to questions were rated on a scale from 0 to 10. Scores were organized into quarters (1st, 2nd, 3rd, and 4th), each involving approximately three weeks and three topics. Percentages were then used to express pre-test and test scores for individual quarters. Estimates of test performance were articulated on a similar scale. All results reported below are significant at the .05 level unless otherwise indicated. Analyses involving either percentage scores or arcsine transformed scores yielded

equivalent results. For simplicity, analyses displayed below involve percentage scores. Degrees of freedom may differ between analyses because data include only values from regularly scheduled assessment activities without the inclusion of make-up tests. Figure 1 illustrates pre-test and test scores, and pre- and post-test assessment values across the four time periods.

Performance Measures

The first question asked whether pre-test and test performance differed between instruction conditions across time. Although there were no significant pre-test differences between instruction conditions [Fs $\leq$ 1.78, *ns*], test performance during the first half of the semester was higher for students who had been exposed to SC retrieval practice than to LC retrieval practice [1st quarter: $F(1, 179) = 4.79$, $MSE = 484.86$, $p = .030$, and 2nd quarter: $F(1, 165) = 4.02$, $MSE = 399.19$, $p = .047$]. Test performance did not differ between groups during the second half of the semester [Fs < 1].

Self-Evaluation Measures

It is important to note that estimates before testing describe a student's ability to predict his/her preparation for an upcoming test, whereas estimates after testing illustrate the student's post-facto appreciation of his/her completed performance. Most estimates tended to be over-estimations before and after testing for all quarters [Fs $\geq$ 4.48] with the exception of estimates before testing produced in the 3rd quarter which were rather accurate [$F = 3.00$, *ns*]. In general, discrepancies between test performance and estimates before and after testing did not vary between types of retrieval practice [Fs $\leq$ 3.86, *ns*] with the exception of estimates after testing during the 4th quarter, which were greater following SC retrieval than LC retrieval. Interestingly, contrary to the expectation that overestimations would increase as academic abilities declined (see Balch, 1992), in both retrieval practice conditions, overestimations before and after testing tended to increase with test scores [$r_s \geq 270$].

Following SC retrieval practice, test performance was positively correlated with estimates before testing during the first half of the semester [1st quarter: $r_s = .377$; 2nd quarter: $r_s = .372$] and with estimates after testing across the entire semester [$r_s \geq .405$]. No correlations existed with estimates before testing during the second half of the semester [3rd and 4th quarters: $r_s \leq .183$, *ns*].

Following LC retrieval practice, test performance during the first half of the semester was not correlated with estimates before testing [1st and 2nd quarters: $r_s \leq .126$, *ns*], whereas it was positively correlated with estimates after testing [1st quarter: $r_s = .573$; 2nd quarter: $r_s = .434$]. Estimates before testing were positively correlated with test performance during the 3rd quarter [$r_s = .279$], but not during the 4th quarter [$r_s = .154$, *ns*]. The same pattern applied to estimates after testing, which were positively correlated with test performance during the 3rd quarter [$r_s = .407$], but not during the 4th quarter [$r_s = .088$, *ns*].

Discussion

In the present investigation, we compared the effects on test performance of two types of retrieval practice, one focused on conceptual integration and the other focused on isolated concepts. Although the test questions involved isolated concepts, which were those upon

which students had the opportunity to practice, study and test questions were always different in form. Retrieval practice was operationally defined as the retrieval of individual concepts or isolated concepts during weekly lab meetings.

Although the results of the present investigation do not support the prediction that conceptual integration enhances test performance, they are consistent with the transfer-appropriate processing account of memory (Graf & Ryan, 1990; McCrudden, 2011; Mulligan & Picklesimer, 2012). According to the latter, the overlap between operations used during practice and those required by a test is essential to successful test performance. In the present investigation, retrieval practice with explicit emphasis on conceptual integration should have made the conceptual operations used for practice and those utilized for testing less likely to overlap. In this scenario, the transfer-appropriate processing account would predict test performance to suffer and its relation with estimates of preparation and actual performance to weaken following retrieval practice focused on conceptual integration. The influence of type of practice (at least on performance), however, should have diminished as students became more familiar with the test format. Indeed, this is the pattern of results observed in the present investigation. Interestingly, Pyc and Rawson (2009) found improved performance following retrieval practice with greater effort producing greater gains. If retrieval practice focused on conceptual integration is assumed to be more effortful than practice focused on individual concepts, then greater gains for practice focused on conceptual integration should have been obtained. Although questions requiring retrieval of two concepts and their links were judged by independent raters as harder than those requiring the retrieval of individual concepts, we found no evidence that effort would lead to superior performance.

To our knowledge, no research exists that has examined the effects of these two forms of retrieval practice on performance and self-assessment in a real classroom setting. Investigations of retrieval practice in the lab or in the field tend to compare the effects of retrieval practice on performance and self-assessment with those of reviewing, usually illustrating the benefits of retrieval practice on both measures (see Agarwal, Karpicke, Kang, Roediger, & Mc-Dermott, 2008; Pilotti et al, 2009; Roediger & Karpicke, 2006b). For instance, Pilotti et al. (2009), who examined retrieval practice and test of isolated concepts, reported an improvement in test performance and self-assessment as measured by responses to questions that covered concepts related to those of the practice questions. The results of the present investigation do not appear to support the notion that retrieval practice irrespective of attention could have enhanced automatic spreading activation, reinforcing connections among concepts in the textbook, and thereby nullifying the effect of attention specifically directed to such connections (as in the LC retrieval practice condition). If retrieval practice had this effect, then performance following the two types of retrieval practice should have been equivalent across the entire semester. Instead, retrieval practice focused on isolated concepts yielded higher performance than retrieval practice focused on conceptual integration during the first half of the semester.

Acknowledgement

This research was supported by a grant from the Spencer Foundation [No. 23282 23282 6151 170]. We would like to thank all NMHU students, research assistants and teaching assistants who graciously donated their time to make this study possible.

Correspondence concerning this article should be addressed to Maura Pilotti, College of Health, Human Services and Science, Ashford University. E-mail: Maura.pilotti@gmail.com.

References

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology, 22*(7), 861-876. http://dx.doi.org/10.1002/acp.1391

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review*, *94*(2), 192-210. http://dx.doi.org/10.1037/0033-295X.94.2.192

Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, 9*, 90-132. New York: Academic Press.

Balch, W. R. (1992). Effect of class standing on students' predictions of their final exam performance. *Teaching of Psychology, 19*(3), 136-141. http://dx.doi.org/10.1207/s15328023top1903_1

Balch, W. R. (1998). Practice versus review exams and final exam performance. *Teaching of Psychology, 25*(3), 181-185. http://dx.doi.org/10.1207/s15328023top2503_3

Butler, A. C., & Roediger, H. L. (2007). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory and Cognition, 36*(3), 604-616. http://dx.doi.org/10.3758/MC.36.3.604

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213-238. http://dx.doi.org/10.2307/1170535

Carr, T., & Dagenbach, D. (1994). Paying attention to newly learned semantic codes. *Canadian Psychology, 35*(1), 109-110. http://dx.doi.org/10.1037/h0084716

Chan, C. J. K., McDermott, K. B., & Roediger, H. L., III (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, *135*(4), 553-571. http://dx.doi.org/10.1037/0096-3445.135.4.553

Conway, M. A., Gardiner, J. M., Perfect, T. J., Anderson, S. J., & Cohen, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, *126*(4), 333-413. http://dx.doi.org/10.1037/0096-3445.126.4.393

Dagenbach, D., Horst, S., & Carr, T. (1990). Adding new information to semantic memory: How much learning is enough to produce automatic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(4), 518-519. http://dx.doi.org/10.1037/0278-7393.16.4.581

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69-106. http://dx.doi.org/10.1111/j.1529-1006.2004.00018.x

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions of Psychological Science*, *12*(3), 83-87. http://dx.doi.org/10.1111/1467-8721.01235

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory and Cognition, 15*(1), 84-93. http://dx.doi.org/10.3758/BF03197714

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General, 116*(2), 119-136. http://dx.doi.org/10.1037/0096-3445.116.2.119

Graf, P., & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6), 978-992. http://dx.doi.org/10.1037/0278-7393.16.6.978

Hacker, D. J., Bol, L., Horgan, D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(2), 160-170. http://dx.doi.org/10.1037/0022-0663.92.1.160

Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research*, *56*(2), 212-242. http://dx.doi.org/10.2307/1170376

Kang, S. H. K., McDermott, K. B., & Roediger, 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. http://dx.doi.org/10.1080/09541440601056620

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, *21*(3), 157-163. http://dx.doi.org/10.1177/0963721412443552

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, *67*(1), 17-29. http://dx.doi.org/10.1016/j.jml.2012.02.004

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 609-616. http://dx.doi.org/10.1037/0278-7393.16.4.609

Maki, R. H., & Serra, M. (1992). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology*, *84*(2), 200-210. http://dx.doi.org/10.1037/0022-0663.84.2.200

McCrudden, M. T. (2011). Do specific relevance instructions promote transfer appropriate processing? *Instructional Science*, *39*(6), 865-879. http://dx.doi.org/10.1007/s11251-010-9158-x

Mulligan, N. W. (2002). The effects of generation on conceptual implicit memory. *Journal of Memory and Language*, *47*(2), 327-342. http://dx.doi.org/10.1016/S0749-596X(02)00006-2

Mulligan, N. W., & Picklesimer, M. (2012). Levels of processing and the cue-dependent nature of recollection. *Journal of Memory and Language*, *66*(1), 79-92. http://dx.doi.org/10.1016/j.jml.2011.10.001

Myers, D. G. (2009). *Psychology in Everyday Life.* New York, NY: Worth Publishers.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*(5), 416-424. http://dx.doi.org/10.1111/1467-9280.00474

Pilotti, M., Chodorow, M., & Petrov, R. (2009). The usefulness of retrieval practice and review-only practice for answering conceptually related test questions. *The Journal of General Psychology, 136*(2), 179-204. http://dx.doi.org/10.3200/GENP.136.2.179-204

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, *60*(4), 437-447. http://dx.doi.org/10.1016/j.jml.2009.01.004

Roediger, H. L., III, & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181-210. http://dx.doi.org/10.1111/j.1745-6916.2006.00012.x

Roediger, H. L., III, & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255. http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4*(6), 592-604. http://dx.doi.org/10.1037/0278-7393.4.6.592

Zabrucky, K. M., Agler, L. L., & Moore, D. (2009). Metacognition in Taiwan: Students' calibration of comprehension and performance. *International Journal of Psychology*, *44*(4), 305-312. http://dx.doi.org/10.1080/00207590802315409

Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think you know but you don't. *Bulletin of the Psychonomic Society*, *15*(1), 41-44.

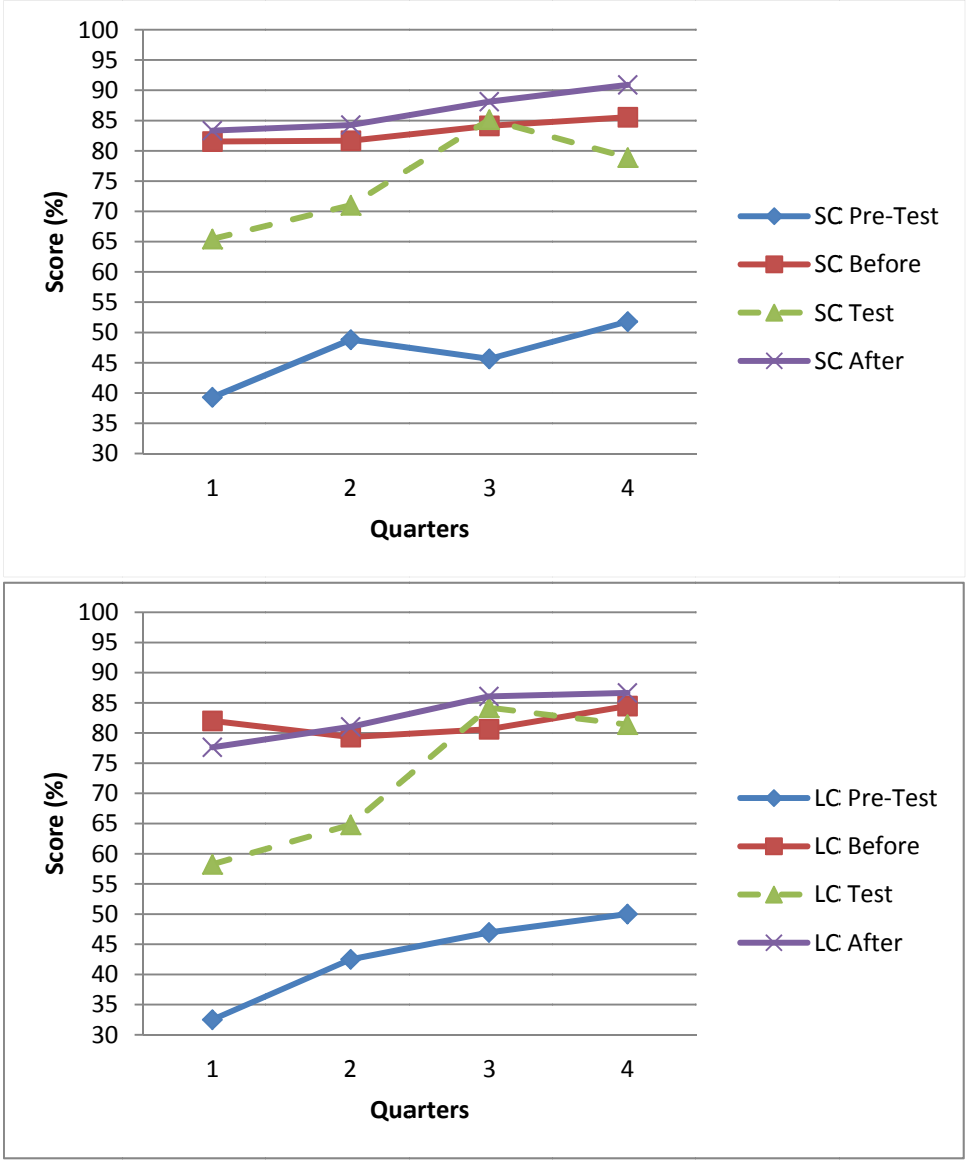

Figure 1. Pre-test performance, estimates before the test, test performance, and estimates after the test following retrieval practice focused either on single concepts (SC) or on linked concepts (LC) as a function of quarters across a semester.