Dynamic Concept Categorization: A Systematic Approach to Improving Classroom Performance

A THESIS

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

APRIL 26, 2012

By____

Thomas W. Hancock, Ph. D **Committee Chairperson**

____

Caleb Lack, Ph. D **Committee Member**

____

Linda Rittner, Ph. D **Committee Member**

Dynamic Concept Categorization:

A Systematic Approach to Improving Classroom Performance

William B. Huffman

University of Central Oklahoma

Author Note

William Huffman, Department of Psychology, University of Central Oklahoma.

This thesis was completed under the direction of Dr. Thomas Hancock.

Correspondence concerning this article should be addressed to William Huffman, Department of Psychology, University of Central Oklahoma, Edmond, OK 73034. E-mail: whuffman@UCO.edu

Abstract

A study technique requiring participants to form unique combination of categories was implemented to explore its impact on test scores. Findings suggest that study aids in general seem to be more beneficial for college senior level courses than freshmen. The paradigm incorporated flash cards as a study aid prior to a psychology exam. The shallow condition allowed participants to use flash cards in an unstructured form while the elaborative condition was instructed to organize the flash cards into a small number of meaningful categories. Test performance declined for the freshman level classes regardless of condition. It was found that the dynamic concept categorization task (elaborative) was ineffective when compared to a self-driven control group (shallow). The trend of the data found from this experiment suggests that attending a review session and participating in a study aid task is more beneficial than not participating in one at all.

Keywords: memory, encoding, retrieval, levels of processing, education, study aids

Dynamic Concept Categorization:

A Systematic Approach to Improving Classroom Performance

Over the last two decades United States classroom performance has been falling in comparison with other countries. According to the Organization for Economic Co-operation and Development (OECD) Programme for International Student Assessment (2010), students in the United States around the age of fifteen ranked 14^{th} in reading, 17^{th} in science, and 25^{th} in mathematics internationally. This problem affects us all if we wish to keep up with other nation's academic competency. One possible solution to this problem is to improve students' study habits. Improved methods of study would allow instructors to more easily teach their students and aid students so that they may learn the material to a higher level of proficiency. The current project discusses the beneficial effects of elaborative encoding, the effective study techniques that use elaborative encoding strategies, and then propose and test a method to improve encoding in order to enhance memory retrieval.

The act of studying for a test involves several memory processes, including enhancing encoding (elaborative encoding) and enhancing retrieval (improving an individual's familiarity with a given topic). One way to implement these memory-enhancing techniques is with study aids (Thomas & Rohwer, 1986). Many effective study aids attempt to use elaborative encoding so that the to-be-learned items may be better committed to memory. Individuals that have used this type of encoding technique are often able to recall information more effectively in academic settings than compared to other study techniques (Papinczak, 2009). As such, findings suggest that when students use these elaborative encoding techniques it leads to higher test scores.

What is elaborative encoding and retrieval? Elaborative encoding is the method of committing to-be-learned material to memory in a meaningful and organized fashion (Craik &

Tulving, 1975). Memory retrieval is the ability to search our memories in order to find a previously learned concept or idea (Neely, 1976). Multiple retrieval cues make memory retrieval more likely to occur when we need to retrieve target information. When applied to education, elaborative encoding often enhances the learner's ability to retrieve information. As the ease of information retrieval is increased, so should test scores increase as well.

Elaborative Encoding

Elaborative encoding is often preferred over maintenance rehearsal due to benefits of more accurate memory retrieval (Craik & Watkins, 1973). Maintenance rehearsal tasks encompass rote memorization techniques that are often ineffective (Dickson & Bauer, 2008). Likewise, learners that implement elaborative encoding techniques often perform better in academic settings than those that do not (Lahitenen, Lonka, & Lindblom-Ylänne, 1997; Papinczak, 2009).

Simpson, Olejnik, Tam, and Supattathum (1994) examined the benefits of elaborative encoding on memory performance. Struggling students were placed into two groups: an elaborative rehearsal group and rote memorization group. Students in the elaborative encoding group were exposed to generalization about the texts, gave meaningful examples of key concepts, and were given information in an organized fashion that made sense. Individuals that used elaborative processes outperformed the rote memorization group in nearly every verbal test that was administered. Thus, when learning newer information it seems beneficial to use elaborative encoding when available.

In a similar vein, Dickson and Bauer (2008) demonstrated the detrimental effects of shallow processing on learning. During lecture, students were allowed to take notes. Upon initial examination, students were not allowed to use their notes. For the second phase of the experiment, students were allowed to retake the test using their notes. They found that students performed better when they had their notes, indicating that they did not learn the material by making them. This data suggests that shallow processing skills (such as rote note reproduction) are an ineffective aid in terms of test preparation. The researchers also argued that the students could have used their notes as a crutch instead of thoroughly learning the material ahead of time.

The aforementioned study does not suggest that note taking is useless, but instead indicates many students may use poor note taking techniques. Lahitnen, Lonka, and Lindblom-Ylänne (1997) investigated several different forms of note taking. Students who took no notes performed the worst followed by individuals who took notes verbatim. However, individuals that paraphrased or made concept maps outperformed all other methods of note taking. Paraphrasing forces individuals to take in the information and rewrite it in their own words; this type of task evokes deep processing while simple note reproduction instead causes shallow processing. Concept mapping on the other hand, is a method of deep processing in which the student is forced to think about, manipulate, and organize information. Individuals that use methods of deep processing perform better on examination than those that use shallow processing.

Further evidence of the effectiveness of elaborative encoding was found by Papinczak (2009). Studying individuals who already knew how to apply methods of deep processing showed that they performed at a higher performance rate academically than those who did not. Individuals that used deep processing also tended to have fewer problems with understanding new information then those who did not apply depth of processing tactics.

Elaborative encoding can be directly implemented as novel information is being given to the individual. For instance, Makany, Kemp, and Dror (2009) had students use a spontaneous note taking task using a program called SmartWisdom. In order to use the SmartWisdom method appropriately students had to take notes in a non-linear fashion that was essentially the same to producing a concept map. What is unique about this idea is that the SmartWisdom software enabled students to make a concept map in real time with novel information. Results yielded such that individuals who used the SmartWisdom software performed better on measures of story comprehension, memory, and complexity of mental representations than individuals that did not use it. This study further emphasizes how effective encoding tasks can improve examination performance.

Further inquiry into the idea of concept mapping finds that concept maps can help individuals that are familiar with the material just as much as individuals that are unfamiliar with the to-be-learned material (Lee & Nelson, 2005). In other words, concept mapping that leads to elaborative encoding allows the learner to better understand the material regardless of how much they already know about the topic. Lee and Nelson (2005) categorized learners into high and low levels of prior knowledge of a concept and had them complete one of two types of concept maps (generative or completive). In the generative condition, participants were asked to create a concept map by creating their own connections. Within the completive concept map condition, participants were given an already constructed concept map. Results from this experiment concluded that individuals that were in the generative concept mapping condition out performed learners in the completive mapping condition. Also, as expected individuals with higher previous knowledge of the concept outperformed individuals with lower previous knowledge of the concept. What is key about these results are that individuals in the generative condition outperformed individuals of similar history of prior knowledge that were in the completive condition.

Elaborative encoding allows students to be more efficient with their limited short-term memory. Work by Miller (1956) has shown that our short-term memory has a capacity of approximately 7 ± 2 concepts. These memory stores can be utilized efficiently or not. For example, short-term memory can be used to temporarily store independent items such as single digits or letters. However, short-term capacity can be greatly increased when related meaningful information is processed (e.g. 9 independent: F I C A B S U I A vs. 3 meaningful: FBI USA CIA). Elaborative encoding allows for chucking of information so that individuals may better understand associations that connect otherwise unrelated concepts (Wickelgren, 1981).

As demonstrated by Dickson and Bauer (2008), shallow processing and maintenance rehearsal procedures often result in ineffective memory aids. When possible, elaborative encoding should be used in order to enhance the encoding of to-be learned information (Lahitenen, Lonka, & Lindblom-Ylänne, 1997; Papinczak, 2009). Elaborative encoding and deep processing strategies enable the student to store and retrieve information from memory more easily so that they may better perform in academic settings.

Memory Retrieval

Meaningful organization of information, such as that done by using elaborative rehearsal, aids memory retrieval (Roediger, 1980). When encoding information to be retrieved at a later date it is more effective to use elaborative rehearsal tactics to maximize the probability of accurate recall in the future (Simpson et al., 1994). Because it is unknown to the student exactly how retrieval cues will appear during the test, it is important that the individual be able to make meaningful connections between the given concepts of interest in an effort to more easily retrieve the desired information. Our memories can be like dominoes, if they are organized correctly, they can knock each other down until the target information is recalled (Wickelgren, 1981).

The current project focused on memory retrieval agents that are evoked from questions (e.g. academic examination). In other words, when an individual is asked a question on a test, memory retrieval refers to the ability of that student to be able to sift through their memory in order to find the relevant information to be recalled (Bekerien & Dritschel, 1992). Individuals have a vast amount of information that has been acquired throughout their lives and if that information is not encoded into memory in a meaningful fashion, it is harder to retrieve a specific piece of information when it is needed. Through elaborative encoding tasks such as paraphrasing and concept mapping, learners can form meaningful associations and knowledge structures that will make retrieval less effortful.

Current Aims and Goals

The benefits of elaborative encoding are evident for maximizing test performance (e.g. Lahitenen, Lonka, & Lindblom-Ylänne, 1997; Makany, Kemp, & Dror, 2009; Papinczak, 2009). From the findings of Lee and Nelson (2005), the idea of generative concept mapping is likely to enhance learning through elaborative encoding and processing (having the learner create the concept map themselves was more effective than having them study a completed concept map). The current project aims to enhance learning with elaborative encoding through a dynamic, but systematic, process. Participants followed a procedure conceptually similar to making a concept map in order to enhance the encoding of to-be-learned information. Instead of drawing a map participants attempted several times to create meaningful relationships between concepts. On each iteration they were to try and reduce the number of categories they made from the previous trial. It was hypothesized that individuals that were given the dynamic concept mapping task would outperform control groups in academic testing. As such, it was expected that construction of fewer and fewer categories would lead to an improved understanding of how concepts relate to one another and thus create rich elaboration.

Method

Participants

Participants included 162 students from the University of Central Oklahoma. The student pool came from four psychology courses that were all taught by the same instructor; two introduction to psychology classes (freshmen level with some non-majors) and two theories of learning and cognition classes (senior level with all psychology majors). The study intended to benefit the students by serving as a review prior to a psychology exam.

Materials

Materials used included 1600 3 x 5 flash cards, 320 sheets of paper, and 3200 2 x 4 address labels. On the back of each flash card, there was a brief description of the respective concept (e.g. a description for the concept *cognitive psychology* would read the study of memory, learning, and attention). In addition, each note card had a code for that respective concept or idea (e.g. 321 for learning, 476 for attention, 689 for memory). These three digit numbers were generated randomly from random.org. Each group of note cards contained key concepts from previous lectures. These key concepts were taken from previous lectures and were determined using a list randomizer from random.org. Each note card was constructed prior to the experiment (see Appendix A and B for examples). Each group of note cards was randomly shuffled prior to distribution to the students using the Latin Square sorting procedure.

Design

This study implemented a 2 (test) X 3 (review type) mixed group design for both senior level and freshman level classes. Review type was a between subjects condition and consisted of an experimental condition (dynamic categorization review), a control condition (self-driven review), and a no review condition. Test type (within) consisted of the first test and second test given during the semester. Three dependent variables were used; Total exam score, multiple choice score, and short answer score. Multiple choice answers made up 80% of all the tests while short answer made up 20% of all tests.

Procedure

In order to establish a baseline, this study was ran during student's second examination. For the second exam a coin flip determined which classes received what kind of review. One section from each class type served as a control condition (self-driven review) while the other section served as an experimental condition (dynamic categorization review). Individuals that did not attend their review session were placed into the no review condition. Prior to the beginning of the experiment participants were asked to sign a consent form. Both conditions lasted 50 minutes and were conducted during their normally scheduled class time prior to an exam. Attendance was optional and participants were not allowed to keep the flash cards.

Participants in the experimental condition were given a dynamic categorization review. Note cards with key terms from previous lectures were given to participants to serve as a review for a psychology exam. Participants were asked to categorize the ideas and concepts into as few meaningful categories as possible, stacking the cards on top of one another so that they can only see the prior card of that respective category. As the participants made their respective groups they were asked to write down their codes in order and groupings on the sheet provided (see

Appendix C). For example, all the cards in category one would be coded in column one on the code sheet. All the cards in category two would be coded in column two on the code sheet and so on. Therefore, it was expected to see the same amount of categories constructed as columns used on the code sheet. Once completed, participants were asked to reshuffle their note cards and repeat the process two more times. Between each iteration, participant were given a new sheet of paper in which to write the codes of the concepts down on.

Individuals in the control condition were given a self-driven review. This type of review was a typical question and answer session. In addition, participants were also given the flash cards used in the experimental condition.

At the end of the experiment, participants were thanked for their time. Individuals were debriefed about the experiment after their regularly scheduled exam. As with all course material, the findings from this research was kept confidential and locked in a file cabinet.

Results

The total number of participants in this study was reduced to 133 due to students dropping their respective class or rescheduling one of their tests to a different time. For those individuals that were in the dynamic categorization review, analysis on number of categorizations made did not change over the course of the experiment $F(1, 44) = 0.48$. It was thought that individuals would lower their number of categories made through repetition of the dynamic concept categorization task, this did not happen. Individuals seemed to be anchored on the number of categories they made during the first iteration.

A (3 review by 2 test for the total exam score) MANOVA F-test of variance was conducted on only the freshman level data. A main effect of test was found between the first ($M = 57.24$, $SE = 1.62$) and second test ($M = 53.86$, $SE = 1.52$) on total exam score $F(1, 79) = 8.59$, thus suggesting that this study negatively impacted freshmen level grades on the second exam (see table 1). Further analysis of the exam score was performed by examining multiple-choice and short answers. Analysis of the freshmen level produced a main effect between first (M = 47.8, SE = 1.12) and second test (M = 44.78, SE = 1.05) on multiple-choice questions $F(1, 79) = 10.1$, indicating that freshmen performed at a lower level on the second exam. A main effect was found between review type for the freshmen level data on short answers $F(2, 77) = 5.19$. A LSD post hoc test suggested that the self-driven review (M = 11.45, SE = 0.93) outperformed the no review condition (M = 7.16, SE = 0.96, $p < 0.01$). There was no statistical difference between the dynamic categorization review when compared to the self-driven and no review conditions in terms of short answers (see *figure 2*). For a breakdown of how each no review condition performed overall compared to the rest of the class in which they were in see Footnote 1.

A (3 condition by 2 test for the total exam score) MANOVA F-test of variance was also conducted on only the senior level data. A main effect was found for review type; dynamic categorization, self-driven, and no review on total exam score $F(2,50) = 3.30$ (see *figure 1*). A LSD post hoc test yielded that the self-driven review (M = 70.42, SE = 2.51) out performed both the dynamic categorization review (M = 63.47, SE = 2.83, $p < 0.05$) and the no review condition (M = 60.4, SE = 3.89, $p < 0.05$). No statistical differences were found between seniors multiple-choice questions. A main effect of review type was found on senior short answer questions $F(2,50) = 6.28$. A LSD post hoc yielded that that the self-driven review (M = 15.83, SE = 0.78) performed significantly better than the dynamic categorization review (M = 13.02, SE = 0.88, $p < 0.05$) and the no review (M = 11.1, SE = 1.2, $p < 0.01$) in terms of short answers (see *figure 2*).

Discussion

In theory, elaborative encoding techniques should be more beneficial than shallow processing techniques in terms of exam performance (Lahitenen, Lonka, & Lindblom-Ylänne, 1997; Papinczak, 2009). The trend of the data suggests that the self-driven condition performed the best, however these findings were not significant for the freshmen level classes. These findings are counter intuitive because the self-driven condition was administered as a shallow level technique. While the dynamic concept categorization task is considered to be a deeper level procedure; a technique that should have been more beneficial for the students overall test performance. Because, individuals likely became anchored on the number of categories made, it could be the case that no new formations of the concepts were made. Thus, the dynamic categorization task actually turned out to function similarly to a shallow level processing task. Regardless of which review session was attended, it seemed better than not attending one (however, in the case of the dynamic categorization review these findings were not significant).

Although reviews seem beneficial for everyone, this project seemed more beneficial to the senior level class than the freshman. These findings fall in line with previous research (Lee & Nelson, 2005). Similarly, looking between first and second exams at the freshman level, a 3.5% reduction was seen on overall test performance. These findings suggest that other methods of study may be more beneficial for individuals with lower levels of prior knowledge on a given subject, such as the task presented by Makany, Kemp, & Dror, (2009). Tasks such as non-linear note taking evoke deep processing upon initial encoding that would likely lead to improved exam performance with freshmen.

This study was designed with keeping external validity high in order to see how this procedure would affect exam performance in an applied educational setting. As such, this study had several potential confounds. Each of the four exams had a 20% short answer section that was graded by four different teaching assistants. The differences seen between the short answer conditions could be due of the different teaching assistants that graded the tests. However, this will not affect the variance noticed between tests one and two (because the teaching assistants were kept consistent across exams). The difficulty of each individual exam was not taken into consideration. Furthermore, one senior and one freshman level class were on Monday, Wednesday, and Friday while one senior and one freshman level class were on Tuesday and Thursday.

One possible solution to improving future projects like this could be to build off of the shallow processing technique that already seems to be more effective (in this case the self-driven condition). Again, in the self-driven control review individuals were simply given the flash cards and were told they could study them. Perhaps monitoring what these individuals are doing correctly could lead to improved study aid designs.

Another possibility could be to improve the dynamic concept categorization task. Administer the iterations over multiple class sessions could beneficial. This idea may help eliminate the students functional fixedness in the categorization task. This change to the procedure would also provide a spacing effect that is often seen as beneficial to memory related tasks (Glenberg, 1977). Ultimately, it may also be better to automate the dynamic concept categorization task such that the students would not have to worry about the flash card code sheet. This would make the flow of the entire task run smoother yet not lose the manipulation check of keeping track of how many categories were made.

In academic settings, it is uncertain which retrieval cues will be available during test. Construction of multiple paths to these target memories would allow for use of available retrieval cues during test. Elaborative encoding allows for construction of these paths towards a target memory. As such, future projects should further inspect the effectiveness of elaborative encoding and study aids in academic settings.

References

Bekerian, D. A., & Dritschel, B. H. (1992). Autobiographical remembering: An integrative approach. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspective on autobiograpcal memory* (pp. 135-150). Dordrecht, Netherlands: Kluwer.

Chiou, C. (2009). Effects of concept mapping strategy on learning performance in business and economics statistics. *Teaching in Higher Education*, *14*(1), 55-69.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268-294.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning & Verbal Behavior, 12*(6), 599-607.

Dickson, K. L., & Bauer, J. J. (2008). Do students learn course material during crib sheet construction? *Teaching of Psychology*, *35*, 117-120.

Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory*, *3*(3), 282-294.

Gurlitt, J., & Renkl, A. (2010). Prior knowledge activation: how different concept mapping tasks lead to substantial differences in cognitive processes, learning outcomes, and perceived self-efficacy. *Instructional Science*, *38*, 417-433.

Lahtinen, V., Lonka, K., & Lindblom-Yänne, S. (1997). Spontaneous study strategies and the quality of knowledge construction. *British Journal of Educational Technology*, *67*, 13-24.

Makany, T., Kemp, J., & Dror, I. E. (2009). Optimizing the use of note-taking as an external cognitive aid for increasing learning. *British Journal of Educational Technology*, *40*(4), 619-635.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97.

Neely, J. H., & Keefe, D. E. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4*, 648-654

Papinczak, T. (2009). Are deep strategic learners better suited to pbl? A preliminary study. *Advance in Health Science Education*, *14*, 337-353.

Reisberg, D. (2006). *Cognition: exploring the science of the mind*. New York, NY: W. W. Norton & Company, Inc.

Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology; Human Learning and Memory*, *6*(5), 558-567.

Simpson, M. L., Olejnik, S., Tam, A. Y., & Supattathum, S. (1994). Elaborative verbal rehearsals and college students' cognitive performance. *Journal of Educational Psychology*, *86*(2), 267-278.

Thomas, J. W., & Rohwer, W. D. (1986). Academic studying: the role of learning strategies. *Educational Psychologist*, *21*(1), 19-41.

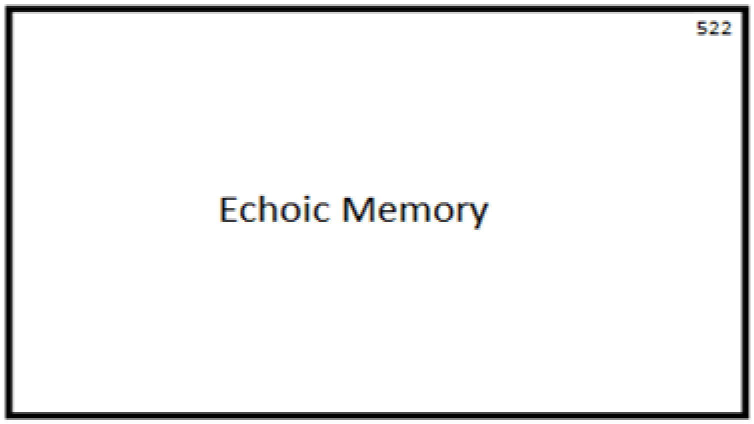
Wickelgren, W. A. (1981). Human learning and memory. *Annual Review of Psychology*, *32*, 21-52.

Footnotes

[1] These analyses were computed to see how the no review condition performed compared to the class they were in (e.g. the no review group was made up of individuals from both the dynamic review and the self-driven review). Note that the number of participants in some of these analyses are insufficient. At the freshmen level when collapsed across classes, the no review (M = 53.22, SE = 2.63, n = 9) did not significantly differ from the self-driven review (M = 57.11, SE = 2.86, n = 28) and the no review (M = 47.35, SE = 2.72, n = 17) did not significantly differ from the dynamic review (M = 54.85, SE = 2.75, n = 26) by overall exam performance. At the senior level a similar pattern was seen, the no review (M = 57, SE = 3.61, n = 3) did not significantly differ from the self-driven review (M = 70.42, SE = 2.76, n = 24) and the no review (M = 61.86, SE = 3.84, n = 7) did not significantly differ from the dynamic review (M = 63.47, SE = 2.77, n = 19) by overall exam performance. Even though none of these differences were significant, the no review group scored lower across all class and condition levels in terms of overall exam performance.

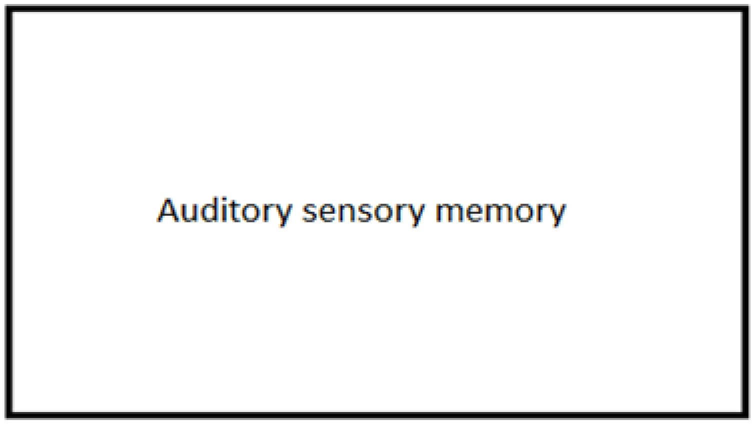
Appendix A

Flash Card (Front)



Appendix B

Flash Card (Back)



Appendix C

Flash Card Code Sheet

256	215	223	648				
568	654	947	314				
225	487	154					
659	547						
332	561						
	991						

Table 1

Descriptives of Overall Exam Performance

Class Level	Condition	Exam	Mean	SD	SE	N
Senior	Dynamic	First	62.37	9.44	2.16	.
		Second	63.47	12.08	2.77	19
	Self-driven	First	68.71	11.22	2.29	.
		Second	70.42	13.52	2.76	24
	No review	First	63.8	8.39	2.65	.
		Second	60.4	9.12	2.88	10
Freshman	Dynamic	First	57.12	16.06	3.15	.
		Second	54.85	14.02	2.75	26
	Self-driven	First	60.64	13.96	2.64	.
		Second	57.11	15.15	2.86	28
	No review	First	53.71	13.01	2.55	.
		Second	49.38	10.42	2.04	26

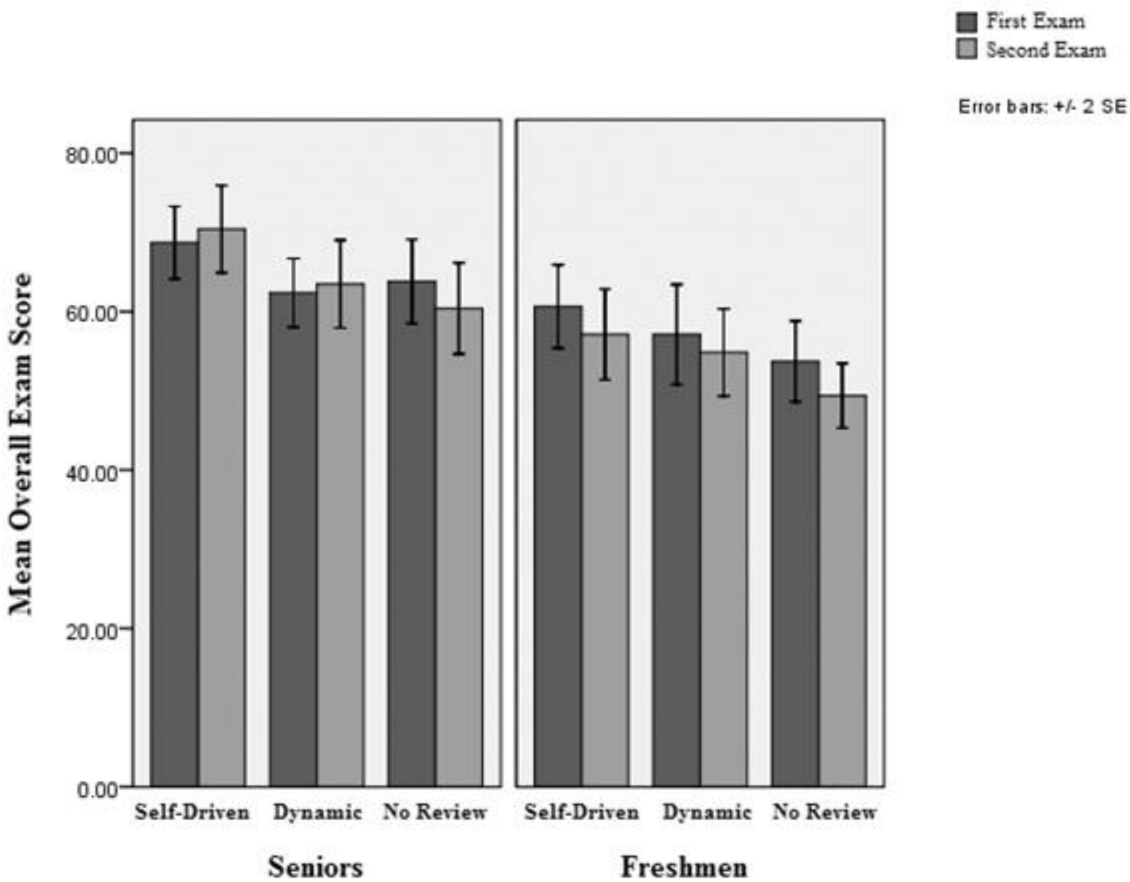


Figure 1. Mean overall exam score between first and second exam broken down by review study aid type and class level.

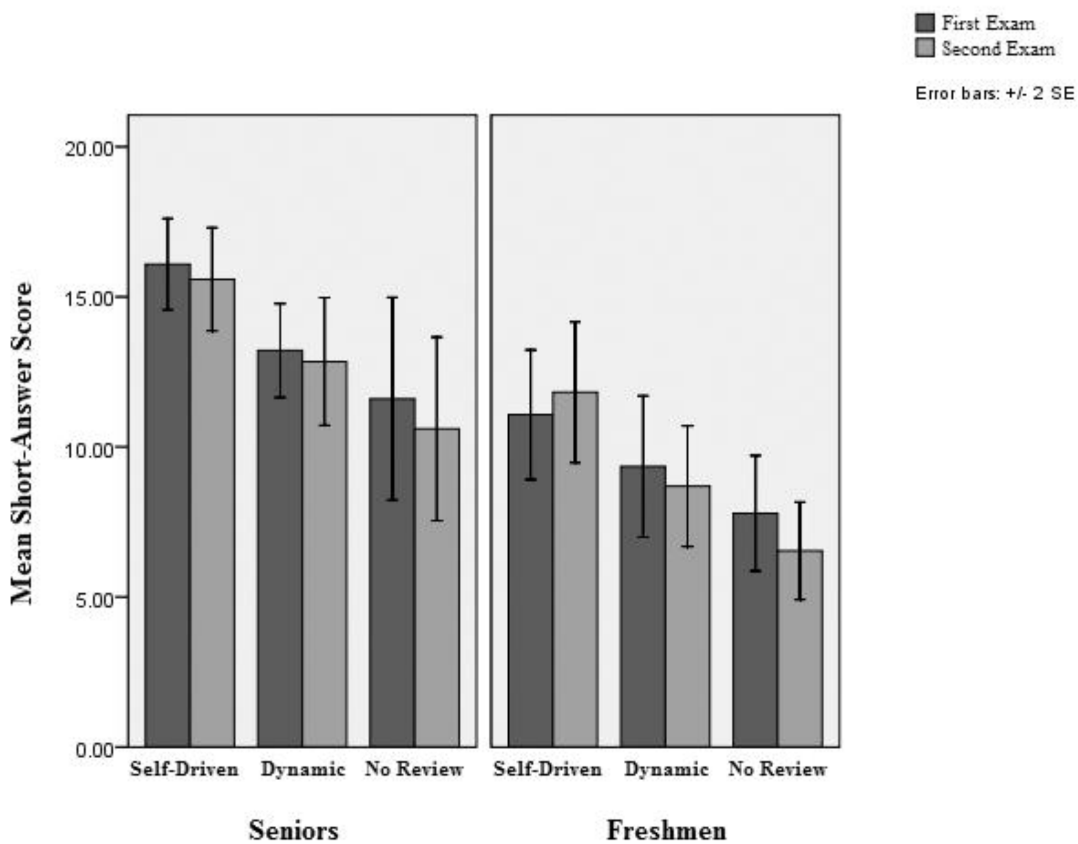


Figure 2. Mean short-answer score between first and second exam broken down by review study aid type and class level.