www.ijemst.net

Using Worked Examples with Active Learning in a Large lecture College Algebra Course

David Allan Miller
West Virginia University, United States

Matthew Schraeder
West Virginia University, United States

To cite this article:

Miller, D. A., & Schraeder, M. (2022). Using worked examples with active learning in a large lecture college algebra course. *International Journal of Education in Mathematics, Science, and Technology (IJEMST), 10*(1), 1-23. https://doi.org/10.46328/ijemst.2011

The International Journal of Education in Mathematics, Science, and Technology (IJEMST) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

International Journal of Education in Mathematics, Science and Technology

2022, Vol. 10, No. 1, 1-23 https://doi.org/10.46328/ijemst.2011

Using Worked Examples with Active Learning in a Large lecture College Algebra Course

David Allan Miller, Matthew Schraeder

Article Info

Article History

Received:
30 June 2021
Accepted:
18 November 2021

Keywords

College algebra
Worked examples
Cognitive Load Theory
Large lecture
Supplemental sessions

Abstract

At a research university near the east coast, a college algebra class has been restructured into two large lectures a week, an active recitation size laboratory once a week, and an extra day devoted to active group work called Supplemental Practice (SP). SP was added as an extra day of class where the SP leader has students work in groups on a worksheet of examples and problems, based on worked example research, that were covered in the previous week's class material. Two sections of the course were randomly chosen to be the experimental group and the other section was the control group. The experimental group was given the SP worksheets and the control group was given a question-and-answer session. The experimental group significantly outperformed the control on a variety of components in the course, particularly in light of the number of days students attended SP.

Introduction

A Commitment to America's Future: Responding to the Crisis in Mathematics and Science Education (Business, 2005) stated that "nationally, 22% of all college freshman fail to meet the performance levels required for entry level mathematics courses and must begin their college experience in remedial courses" (p. 6). The enrollment in College Algebra has grown recently to the point that there are estimated 650,000 to 750,000 students per year in the United States (Haver, 2007), which has surpassed the enrollment in Calculus. Although there are almost three-fourths of 1 million students enrolled in College Algebra, it is estimated conservatively that 40% of these students receive a grade of D or F, and that number can reach above 60% at some colleges (Haver, 2007). To address this lack of success, faculty members teaching College Algebra at a large research university in the eastern part of the United States have implemented a new structure in the course that emphasizes active learning through a day of instruction called Supplemental Practice.

Background on Supplemental Practice Schedule

Supplemental Practice (SP) at our institution was implemented during the Fall 2004 semester and was originally

modeled after Supplemental Instruction (Arendale, 1994; SI Staff, 1997). The College Algebra course that consisted of three lectures each week morphed into a structure of two lectures each week in a large lecture room (Mondays and Fridays), an active laboratory class on Tuesdays in computer classrooms (where students met in smaller groups), and a day of SP for lower-achieving students on Wednesdays. The SP days were required for students who scored lower than an 80% on a placement exam or lower than a 70% on any regular test. Starting in the Fall 2006 semester, the SP sessions changed into active problem-session days modeled after cognitive science "worked-out example" research. The worked-out example research, henceforth denoted as worked examples, requires students to study a worked example for a particular topic, ask questions about anything in the example that they do not understand, and work a similar example without reference to the worked example, nor other outside sources (Cooper & Sweller, 1985; Ward & Sweller, 1990; Zhu & Simon, 1987; Carroll, 1994; Tarmizi & Sweller, 1988).

The SP sessions and worksheets were developed based on worked example research because it allows the instructor to work in the large lecture class to assist students as they learn the material and it helps students: 1) to be actively engaged with the material in a setting where they can get feedback and assistance as they solve problems, 2) in transferring information from working (short-term) memory to long-term memory, 3) to regulate their learning and build confidence that they can work problems, and 4) with homework, quizzes, and studying for tests outside of class.

Theoretical Background

Cognitive load theory describes learning in terms of processing information using long-term memory, which "effectively stores all of our knowledge and skills on a more-or-less permanent basis and working memory, which performs the intellectual tasks associated with consciousness" (Cooper, 1998, p. 1). Information can be placed into long-term memory after it is processed by working memory (Cooper, 1998). However, working memory is limited and can only hold a certain amount of information for a limited amount of time (Cooper, 1998). The information in working memory is constantly being replaced by more relevant, useful, and necessary details based on what we are thinking, doing, and experiencing.

Simply put, cognitive load is the strain placed on one's working memory when completing a mental task. A low cognitive load does not require much thought, while a high cognitive load necessitates a great deal of brain function. At any given time, the human brain can only withstand a certain amount of activity. When learning new material, the higher the cognitive load, the more stress and strain are placed upon an individual, resulting in more difficulty in learning that new material.

Based on the borrowing principle, knowledge that is gained by an individual is borrowed from the knowledge of another, whether that information is passed on verbally, in a written format, or simply through observation (Sweller, 2006). Various sources of information are combined through the borrowing principle to create new knowledge, but new information is only truly created through the problem-solving process (Sweller, 2006). Information from long-term memory may be used to solve problems, which is an instance of the borrowing

principle (Sweller, 2006). However, when sufficient knowledge is not available, one must engage in some form of trial-and-error to derive a solution, which is known as the randomness through genesis principle (Sweller, 2006). This process is similar to how evolution occurs through genetic mutations: successful problem-solving attempts become part of long-term memory, which become new knowledge, which can then be passed on to others through the borrowing principle (Sweller, 2006).

Although working memory has limitations with respect to new information (based on cognitive load), there are no such restrictions when handling information that has been stored in long-term memory (Sweller, 2006). The key is to shift new information from working memory to long-term memory so that it can be categorized and used in the future. Information from long-term memory is recalled to working memory to deal with situations and solve problems, which can create new information that is added to long-term memory (Sweller, 2006). The cycle repeats as new knowledge is gained by individuals and is then passed on to others.

Research has identified three categories of cognitive load: intrinsic, extrinsic, and germane. Each is affected by different aspects of learning. DeLeeuw and Mayer (2008) found that the correlations among these three measures of cognitive load were not very high, but stated that some measures are more sensitive to one type of change in cognitive load than to others.

Intrinsic cognitive load refers to the "intrinsic nature of the materials to be learned" (Ayres, 2006, p. 287) and is influenced by the complexity of the material (known as element interactivity) and students' prior knowledge (Sweller & Chandler, 1994). Cierniak, Scheiter, and Gerjets (2009) surmised that intrinsic cognitive load is only affected by the nature of the content and not the instructional design. Große and Renkl (2006) found that high intrinsic cognitive load can inhibit learning.

Extraneous cognitive load is impacted by the mode and conditions in which the instruction takes place (Ayers, 2006), also known as the instructional design (Cierniak et al., 2009). It is caused by processing unnecessary information, which impedes learning because it does not lead to the development of productive problem-solving schema (Cierniak et al., 2009).

Like extraneous cognitive load, germane cognitive load is affected by instructional design (Cierniak et al., 2009). Unlike extraneous cognitive load, germane cognitive load benefits learning, rather than inhibiting it (Cierniak et al., 2009) because it promotes schema acquisition (Ayres, 2006; Sweller, Van Merrienboer, & Paas, 1998). Where extraneous cognitive load deals with unnecessary information (distractions), germane cognitive load directs students' attention to relevant learning processes (Sweller et al., 1998).

Sweller (2006) observed that germane cognitive load is not often affected directly. Rather, it is impacted "indirectly by reducing extraneous cognitive load and so freeing working memory capacity for an increase in germane cognitive load" (Sweller, 2006, p. 168). Cierniak et al. (2009) also advocated for reducing extraneous cognitive load (information processing that hinders learning) and increasing germane cognitive load (information processing that supports learning). However, increasing germane cognitive load without sufficient

working memory will likely inhibit learning rather than promote it (Sweller, 2006), so one must be cautious when attempting to accomplish this.

Literature on Worked Example Research and Cognitive Load

The discipline of cognitive science deals with the mental processes of learning, memory, and problem solving. Worked example research was developed based on Sweller's (1988) cognitive load theory. The total load on working memory at any moment in time is referred as the cognitive load. Most people can retain about seven "chunks" of information in their working memories and when they exceed that limit at any moment in time, there will be a loss of information in the working memory (Sweller, 1988). In other words, there is an overflow of information in the working memory and cognitive overload. Cognitive overload can be thwarted if one limits information so that it does not exceed the students' working memory. One way this can be done is to transfer information from working memory to long-term memory as information is being processed (or soon after). According to Sweller (1988), optimum learning occurs in humans when one minimizes the load on working memory, which, in turn, facilitates changes in long-term memory.

It has been suggested that worked examples reduce the cognitive load on a student and might optimize schema acquisition (Sweller & Owen, 1989; Sweller & Cooper, 1985). In addition, worked examples have been researched (and used) in a variety of subjects: mathematics (Cooper & Sweller, 1985; Zhu & Simon, 1987), engineering (Chi, Bassok, Lewis, Reimann, & Glaser, 1989), physics (Ward & Sweller, 1990), computer science (Catrambone & Yuasa, 2006), chemistry (Crippen & Boyd, 2007), and education (Hilbert, Schworm, & Renkl, 2004).

Sweller and Cooper (1985) conducted one of the first studies on worked examples in high school Algebra. Through five experiments, they examined the use of worked examples as a substitute for problem solving and consistently found that the worked example group outperformed the control group (which worked conventional homework problems instead of studying worked examples) and had less time on task. Zhu and Simon (1987) demonstrated the feasibility and effectiveness of teaching mathematical skills through chosen sequences of worked examples and problems in a Chinese middle school's Algebra and Geometry curriculum without lectures or other direct instruction. Ward and Sweller (1990) established that worked examples formatted to reduce the need for students to mentally integrate multiple sources of information resulted in test performances superior to either conventional problems or worked examples that required students to split their attention.

According to Sweller (2006), change (and the subsequent learning) is most effective when employing the borrowing principle, so using worked examples should be an efficient method of learning for students. To be clear, the need for randomness as genesis can never be eliminated or erased, since it will always be needed to solve new problems or deal with unique situations. However, working memory is designed to handle situations that require the use of the borrowing principle more quickly and efficiently than it can handle randomness as genesis (Sweller, 2006). After all, it is easier to remember how to solve a problem than it is to solve a problem for the first time, so using worked examples is a more efficient use of instruction time than just problem solving.

The idea that learning from examples is more favorable than learning through problem solving when acquiring initial cognitive skills is known as the worked-example effect (Schwonke, Renkl, Krieg, Wittwer, Aleven, & Salden, 2009). However, Catrambone and Yuasa (2006) found that worked examples are not sufficient for learning; further action is necessary. After learning from a worked example (sometimes more than one), students must practice their own problem (or problems), which provides motivation to utilize what was learned (Catrambone & Yuasa, 2006). The ultimate goal is for students to recognize the situations in which specific problem-solving techniques are appropriate, as well as the results and consequences of using those techniques (Catrambone & Yuasa, 2006). Worked examples provide the first step towards reaching this goal.

Worked examples can be presented in two basic formats: example-problem and problem-example. The example-problem sequence presents worked examples, followed by problems for the students to complete. Reisslein, Atkinson, Seeling, and Reisslein (2006) found that the example-problem structure was more effective for students with lower prior knowledge. The problem-example sequence, which was more effective for students with higher prior knowledge (Reisslein et al., 2006), allows students to work on some problems and provides a worked example that can be referenced only when necessary. In short, the stronger students were able to use prior knowledge and problem-solving techniques to successfully find solutions with minimal assistance from the worked examples, but weaker students required more-structured help to succeed.

When students utilize worked examples, they frequently use self-explanation: an attempt to rationalize the steps in the worked example to make the process understandable. Chi et al. (1989) showed that beyond merely studying worked examples, “good” students generally monitored their own understanding and misunderstanding through self-explanations, which the “poor” students did not do. Große and Renkl (2006) found that self-explanations can produce a germane cognitive load that is likely to promote learning. However, Sweller (2006) interjected that any technique that directly impacts germane cognitive load while leaving intrinsic and extraneous cognitive load unchanged may potentially increase learning time, even if overall learning is not affected.

The amount of information provided to students is a vital component to worked examples. Extraneous information, particularly information that students already know, can confuse students and inhibit learning (Sweller, 2006). While weaker students need additional explanation to understand the process of solving a problem, stronger students can be impeded by being exposed to that same information. This phenomenon is known as the Expertise Reversal Effect, in which one strategy works for novices and another is more effective as students acquire knowledge and skill (Kalyuga, Ayres, Chandler, & Sweller, 2003).

In a classroom with students of mixed abilities, either a delicate balance must be struck when providing guidance in the form of worked examples or considerations must be taken to provide the proper amount of help to allow each group of students to succeed. Schwonke et al. (2009) had success in this realm when utilizing faded worked examples. By gradually reducing the amount of support provided on the worked examples, students required less time to learn material and acquired a deeper conceptual understanding than with problem

solving supported by feedback and scaffolding (Schwonke et al., 2009). Sweller (2010) reported similar results by providing partially worked examples, with students tasked to complete the problems and find the solutions. Students were first given fully worked examples, followed by partially worked examples to be completed, and then unworked problems to solve without any assistance. This process facilitated germane cognitive load (Sweller, 2010).

Ayres (2006) investigated isolated vs. integrated worked examples. Isolated worked examples break the solution into individual elements, focusing on the change in one element at a time. The integrated worked examples change several expressions simultaneously, accomplishing in one step what the isolated examples take several steps to complete. The error rates and cognitive load were the lowest for the group that used the isolated worked examples, which agrees with "the prediction that a single-calculation strategy (isolated-elements) reduces element interactivity and consequently overall cognitive load" (Ayres, 2006, p, 291). As described by Sweller (2010), the isolated-interacting elements effect is controlled by changes in intrinsic cognitive load. At times, the element interactivity can lead to extremely high cognitive load, which lower-ability students cannot process in working memory (Sweller, 2010).

Ayres' (2006) results echoed those of Sweller (2006) and Reisslein et al. (2006) in that students of different abilities require different types of support to succeed. Although the results suggested that the isolated worked examples resulted in fewer errors and a low cognitive load for the lower-ability students and for the overall group, students with higher ability fared better with the integrated worked examples (Ayres, 2006). The supposition is that the isolated worked examples were too simple for the higher-ability group and not engaging enough to hold the students' interest, but those same isolated examples lowered the intrinsic cognitive load enough for the lower-ability group to be effective (Ayres, 2006). The key to understanding the integrated examples is for students to learn to focus on a single element when presented with integrated worked examples and mentally isolate it, along with any changes to it (Sweller, 2010). In other words, the approach is one of divide-and-conquer in which students must look at each part separately without trying to do everything at once. Stronger students can do this, but weaker students are overwhelmed by the complexity of the task.

Another contrast of styles in the format of worked examples is molar view vs. modular view. Molar view is a holistic approach that categorizes problems and defines their associated solution procedures (Gerjets, Scheiter, & Catrambone, 2004). Conversely, modular view divides problems into basic elements that can be assessed and understood separately (Gerjets et al., 2004). Molar examples tend to result in high cognitive load levels, which can hinder "higher-level processes like elaborations and comparisons" (Gerjets et al., 2004, p. 55), but modular examples can reduce overall cognitive load (Gerjets et al., 2004). The change in example style from molar to modular can reduce intrinsic cognitive load (Sweller, 2006).

Besides the reduction in cognitive load, Gerjets et al. (2004) found a variety of additional benefits when using modular view over molar view, regardless of how similar or different the new problems are from the worked examples. Students using the modular view required less study time to achieve greater success when solving problems, regardless of the number of topics addressed, the type of task completed, or the amount of

instructional explanation provided (Gerjets et al., 2004). These results were seen with learners of all ability levels, as well (Gerjets et al., 2004).

The research on worked examples in mathematics cited above has been conducted in smaller classes or a laboratory setting, in which students volunteered to be part of the study that was not conducted in a particular course in a high school or college. The research in the current study was conducted in a large lecture classroom setting and concentrated on determining whether worked examples helped to promote success in the course. In addition, past worked example research in mathematics has dealt very little with college mathematics courses, classes in a large lecture setting, or implementing an extra day of class to focus on working with students to master material (Sweller & Cooper, 1985; Zhu & Simon, 1987; Ward & Sweller, 1990). The research could be valuable to other researchers who are working to promote student success in large lecture classes. The research question that will be addressed in this study is: "Do students in the experimental group who attend most (or all) of the Supplemental Practice sessions earn significantly different grades on course components than students in the control group?"

Prior Work on Supplemental Sessions

Miller and Schraeder (2015) investigated the differences between an experimental group that consisted of students who attended the SP day each week where worked examples and group work were implemented and a control group of students who received a question-and-answer session. They found that the experimental group significantly outperformed the control group on Test 3 and the Final Exam, and almost significantly on Test 2 and the quizzes. There was no significant difference on total points in the course, Test 1, or Test 4.

When Miller and Schraeder (2015) investigated the experimental and control groups based on scores on a retired ACT representing prior knowledge, in which they separated the control and experimental group into a high, middle, and low prior knowledge group, they found significant differences occurred with the middle prior knowledge experimental and control groups, with the experimental group significantly outperforming the control group on nearly every comparison. Miller and Schraeder (2015) argued that the low prior knowledge group had too many mathematical obstacles to overcome; whereas, the high prior knowledge group would perform well no matter what interventions were available in the class. However, the middle prior group benefited from the worked examples, since it helped to build their knowledge-regulation component and they became more motivated as they had success in the course. All of this fed into helping them increase their learning of College Algebra. This study will focus on examining what effect attendance of SP days had on students' performances on course components in College Algebra.

Method

Course Description

Three different types of College Algebra courses were taught at the university. Three-day large lecture College Algebra was comprised of two lectures each week in a large lecture setting and one day each week in the math

lab where students actively worked in small groups. The 4-day College Algebra course had the same format as the 3-day College Algebra course, except the 4th day was spent in SP. Five-day College Algebra had five lectures each week with a class size of approximately 40 students. The 5-day College Algebra class administered all quizzes and tests by pencil and paper and did not have a laboratory component.

All College Algebra courses required specific placement test scores, with 3-day College Algebra requiring the highest placement score and the 5-day College Algebra needing the lowest. All sections of the 4-day College Algebra course were coordinated by one course coordinator to ensure that the course structure, labs, quizzes, and tests were the same from section to section. The same instructor taught all three sections during the semester the study was conducted.

The quizzes were given online outside of class and students were given up to three attempts on each quiz. Laboratories were completed in groups of two or three students when the class met in the math lab. The labs had interactive applets (mostly grapher applets) that students utilized to complete the lab. A grapher applet could be used on the tests in place of a graphing calculator. The tests and Final Exam were common online tests consisting of 20 multiple-choice questions and were administered through the eCampus management system during the lab once every three to four weeks. Students were given a list of suggested practice problems, but no homework was collected. The quizzes and tests were based on these practice problems. The course consisted of 100 points for attendance, 200 points for eight labs, 100 points for six quizzes, 100 points for each of four tests, and a 200-point comprehensive Final Exam.

Participants and Setting

The setting for the research was a large lecture 4-day College Algebra course with an annual enrollment of around 1000 students at a research university in the eastern part of the United States. The participants for this study were students in the three sections of the 4-day College Algebra course. One section was taught in the morning (experimental) and the other two in the early afternoon (one experimental and one control). The Monday class was the SP session when students worked in small groups on the worked example worksheets or had a question-and-answer session, depending on which class they were in. College Algebra is a general elective course and many majors require it for their degrees, so students from a variety of majors took the class.

Worked Example Worksheets

During the SP days, worked example worksheets were distributed to the students to complete collaboratively. Since the class was still in the large lecture classroom setting with a theatre-seating structure, students formed groups with other students near them as they saw fit. Usually, students worked with one to three other students seated close to them, but some chose to work independently. The worked example worksheets consisted of an expert solution of a College Algebra problem followed by a problem for the students to complete. Examples of several worked examples from worksheets are shown in the work by Miller and Schraeder (2015) to give the reader some idea of their structure. The worksheet was always given to the students as one sheet (front and

back) in a two-column format with headings for all worked examples, followed by the corresponding section in the textbook (Sullivan & Sullivan, 2006).

There were approximately 8 to 12 worked examples and problems on each worksheet, with material consisting of topics covered during the previous week's lectures. No new material was ever presented on the worksheets, which included problems directly from or derived from the problems in the textbook. The worksheets were never developed while referencing material directly from tests, quizzes, or labs. However, most of the questions from the tests and quizzes were similar to the problems in the book.

The worksheets were modeled after worked example research in that they presented an expert's solution to a problem followed by a problem for the students to work out. The only difference is that it is not plausible to ask the students to not reference the worked examples while working problems, so this was never attempted. Furthermore, most studies on worked examples state that the student should be given a similar problem to complete (nearly identical in some cases), but in SP, the problems sometimes varied from the worked examples. However, Gerjets et al. (2004) had success with different problems when using a modular format, which the worksheets utilized.

Experiment

The researchers randomly designated one of the course sections as the control group (n = 177) and the other two sections as the experimental group (n = 320). Students in the experimental group were given a worked example worksheet at the beginning of each of the 13 SP days and asked to work in groups to complete the worksheet. Two to three undergraduate class assistants and a graduate student circulated around the room to answer any questions. In the control group, a graduate student conducted a question-and-answer session during the extra day instead of giving a worksheet to the students. Students were able to get any question answered, but the graduate student only answered questions from the students and did not generate additional questions.

Students in the experimental group who attended 8+ SP sessions were grouped into the 8+ experimental group (n = 279), and students in the experimental group who attended all 13 SP sessions were grouped into the 13 experimental group (n = 117), so the 8+ experimental group included the members of the 13 experimental group. We will denote the 8+ experimental and 13 experimental groups as 8+ Exp and 13 Exp, respectively. Quantitative data (scores on tests and quizzes, supplemental days attended, class attendance, total points, etc.) were collected for each student in the control, 8+ Exp, and 13 Exp groups, and analyzed at the end of the semester. Both the control and experimental groups had similar demographics.

Findings

The 8+ Exp and Control Groups

Data from the experimental and control groups were compared on a variety of levels by using a t-test. An F-test determined whether the variances were equal or unequal and the corresponding t-test was administered. The

experimental and control groups had similar levels of retention (number of students who completed the course) at 80.5% and 84%, respectively. At the beginning and end of the semester, all students were given a retired ACT math test that consisted of 60 questions. Based on the number of problems answered correctly, students were given up to 10 extra credit points for the ACT. These extra credit points were excluded from all grade calculations for this study.

Table 1 shows the topics for the questions on the ACT, which section in the textbook covered those topics, and the number of questions related to those topics.

Table 1. Breakdown of the ACT

Chapter of the Textbook/Other Topics	Topics from Chapter on ACT	Number of Questions
Review	Number Systems, Evaluating Algebraic Expressions, Using Laws of Exponents, Perimeter	27
1	Linear Equations, Application using Linear Equations, Absolute Value Inequalities, Inequalities, Midpoint, Application of Slope, Absolute Value, Basic Graphing	11
2	Lines	6
3	Properties of Functions	1
6	Composite Functions, Logarithms, and Exponential and Logistic Growth and Decay Models	3
12	System of Equations	2
Probability and Statistics	Average, Probability, Median, Weighted Average	5
Trigonometry	Definition & Graphs of Trigonometric Functions, Trigonometric Identities, Angles, Arc Length	5

Twenty-seven of 60 questions were on review topics that are covered in a typical high school algebra class in the United States, while an additional 23 questions would likely be covered in high school algebra (linear equations, inequalities, absolute value problems, distance, midpoint, quadratics, concept of functions, composition of functions, systems of equations). Therefore, around 83% of the ACT questions were on concepts likely covered in high school algebra and, thus, the ACT was a good measure of students' prior knowledge entering the class. The ACT questions on Trigonometry were near the end of the test and few students completed the entire ACT. Finally, the probability questions (not central to this study) were fairly basic and scattered throughout the test.

Because it is representative of the material covered in an algebra class and the fact that the ACT is often used for college admission, we believe the ACT provided a good assessment of prior knowledge and is a valid assessment for this purpose. In addition, Butler, Pyzdrowski, Walker, and Butler (2012) have the used this same retired ACT (giving it both at the beginning and end of the semester) to assess gains in knowledge in College Algebra. Miller and Schraeder (2015) also used this ACT as a baseline for knowledge students have at the beginning of the College Algebra course. As with Butler et al. (2012), the ACT was administered twice in this study: during the first week of classes (Pre-ACT) and during the last week of classes (Post-ACT).

The control and 8+ Exp groups' mean Pre-ACT scores were 28.40 and 26.88 with standard deviations of 6.41 and 6.91, respectively. The control group significantly outperformed the experimental group ($p = 0.01$) on the Pre-ACT. The control and 8+ Exp groups earned mean Post-ACT scores of 32.81 and 32.35, with standard deviations of 6.46 and 7.13, respectively. There was no significant difference between the mean Post-ACT scores of the two groups, meaning the 8+ Exp group eliminated the advantage that the control group initially possessed.

The two groups were compared with respect to each test, the Final Exam, quizzes, total points in the course (excluding points for attendance and extra credit), and course grade point average (see Table 2). Course GPA was calculated by assigning a quantitative score for the final grade that each student earned in the course (A = 4, B = 3, C = 2, D = 1, and F = 0). The 8+ Exp group significantly outperformed the control group on nearly every measure: Test 2 ($p < 0.05$), Test 3 ($p < 0.001$), Test 4 ($p < 0.01$), quizzes ($p < 0.001$), Final Exam ($p < 0.0001$), total points ($p < 0.001$), and course GPA ($p < 0.001$).

Table 2. Means (Standard Deviations) - Control and 8+ Experimental Groups

	Test 1	Test 2	Test 3	Test 4	Final Exam	Quizzes	Total Points	Course GPA
Control Group (n=177)	68.84 (16.21)	66.86 (19.93)	66.58 (19.61)	67.23 (22.67)	56.10 (24.02)	71.52 (18.89)	605.88 (140.25)	1.97 (1.17)
8+ Experimental Group (n=279)	68.62 (15.97)	70.86 (17.03)	72.24 (17.81)	72.35 (20.02)	64.61 (20.04)	78.09 (17.62)	652.69 (113.88)	2.34 (1.16)

The 8+ Experimental and Control Groups with Respect to Prior Mathematical Knowledge Level

The Expertise Reversal Effect (Kalyuga et al., 2004) describes how different explanations can work better for certain groups of students with different ability levels. To investigate this, the researchers compared the control and 8+ Exp groups based on levels of prior knowledge. The high prior mathematical knowledge groups (high 8+ Exp and high control) consisted of students who scored 31 or higher score (out of 60) on the Pre-ACT.

The members of the middle prior mathematical knowledge groups (middle 8+ Exp and middle control) got 26 to 30 correct, and the students in the low prior mathematical knowledge groups (low 8+ Exp and low control) correctly answered 25 or fewer questions. These ranges in scores were chosen in an effort to trisect both the

control and 8+ Exp groups. The n values in each group would change quite dramatically if the ranges were changed. While not perfectly divided, this split was the most equitable. The researchers could have looked at quartiles, but simplified things by examining only three groups.

No significant differences existed between corresponding prior mathematical knowledge levels form the 8+ Exp and control groups on either the Pre-ACT or Post-ACT, so both high groups had similar results, as well as with the two middle groups and two low groups. Notice that the control group significantly outperformed the experimental group on the Pre-ACT, but none of the control subgroups significantly outperformed their experimental counterparts. The low 8+ Exp group was larger than the high 8+ Exp and middle 8+ Exp groups, but the three control groups had more similar sizes, so the low 8+ Exp group contributed to a lower overall mean for the 8+ Exp group on the Pre-ACT.

The bold p-values in Table 3 show which comparisons between corresponding groups were not significant (using $p = 0.05$ for significance).

Table 3. Means (Standard Deviations) with respect to Prior Knowledge Levels

	High Control (n=54)	High 8+ Exp (n=78)	p-value	Middle Control (n=59)	Middle 8+ Exp (n=79)	p-value	Low Control (n=49)	Low 8+ Exp (n=104)	p-value
Total Points	678.77 (88.04)	709.92 (86.15)	0.023	578.59 (160.41)	657.26 (101.60)	0.0003	574.42 (109.51)	614.83 (124.73)	0.022
Test 1	75.83 (12.99)	76.09 (12.55)	**0.455**	67.88 (15.18)	70.38 (14.20)	**0.164**	64.80 (14.25)	62.16 (16.92)	**0.159**
Test 2	77.69 (15.47)	79.29 (12.58)	**0.264**	64.41 (22.32)	69.87 (15.23)	0.045	59.08 (16.06)	65.58 (19.26)	0.015
Test 3	74.17 (14.59)	79.36 (14.65)	0.024	66.19 (19.94)	74.56 (15.36)	0.003	61.12 (19.02)	66.63 (19.16)	0.049
Test 4	75.37 (16.25)	79.87 (14.85)	**0.054**	63.47 (24.48)	73.73 (19.29)	0.003	64.18 (20.52)	66.83 (22.17)	**0.235**
Final Exam	139.26 (31.85)	149.87 (28.76)	0.026	100.17 (54.63)	126.20 (36.70)	0.0005	101.84 938.98)	117.98 (43.50)	0.012
Quizzes	77.72 (14.30)	82.56 (12.89)	0.024	69.21 (20.09)	79.54 (17.51)	0.001	70.82 (16.85)	75.56 (19.64)	**0.063**
Course GPA	2.52 (0.99)	2.91 (0.97)	0.013	1.85 (1.26)	2.38 (1.11)	0.006	1.61 (0.91)	1.98 (1.18)	0.028

As can be seen, the 8+ Exp subgroups outperformed their control counterparts nearly universally, with most of those comparisons being significantly different. Of note is the dominance by the middle 8+ Exp group over the middle control group.

The 13 Experimental Group versus the 8+ Experimental Group

Gerjets et al. (2004) found that a modular form of worked examples produced significantly better results that a molar form regardless of the number of topics addressed, the type of task completed, or the amount of instructional explanation provided. This study did not compare modular and molar forms, but the researchers decided to investigate whether these phenomena occurred when comparing students who attending 8+ SP sessions vs. those who attended all 13. The intent was to gauge the impact of attending all of the SP session as opposed to just most of them, exposing the 13 Exp students to more topics, more types of tasks, and potentially more instruction time. Table 4 shows the mean scores, standard deviations, and p-values for total points (excluding attendance and extra credit), tests, Final Exam, quizzes, and course GPA (with standard deviations in parentheses).

Table 4. Means (Standard Deviations) - 8+ and 13 Experimental Groups

	8+ Experimental (n=279)	13 Experimental (n=117)	p –value
Total Points	652.69 (113.88)	693.99 (85.59)	0.00005
Test 1	68.62 (15.97)	73.59 (13.61)	0.00093
Test 2	70.86 (17.03)	74.36 (15.39)	0.02337
Test 3	72.24 (17.82)	78.25 (14.66)	0.00029
Test 4	72.35 (20.02)	79.49 (13.36)	0.00002
Final Exam	129.29 (40.15)	137.52 (32.72)	0.01620
Quizzes	78.09 (17.62)	85.07 (11.29)	0.000002
Course GPA	2.344 (1.16)	2.752 (0.955)	0.00017

The 13 Exp group significantly outperformed the 8+ Exp group in every facet of the course (using $p = 0.05$ for significance). The differences would have been greater if the 8+ Exp group only included the students who attended 8 to 12 supplemental sessions (n = 162), rather than the encompassing the 13 Exp group.

Analyzing Prior Knowledge for the Control, 8+ Exp, and 13 Exp Groups

Using the same requirements for creating the high, middle, and low prior knowledge subgroups as before, the

control, 8+ Exp, and 13 Exp groups were divided and compared. The low, middle, and high control groups all had a higher mean on the Pre-ACT than the corresponding 13 Exp groups, but only the middle group was significantly higher ($p < 0.0001$). However, there were no significant differences on the Post-ACT.

The comparisons for the other course assessments are shown in Table 5. Non-significant results (using $p < 0.05$ for significance) are shown in bold. The 13 Exp groups achieved higher averages on all assessments than the corresponding control groups, with nearly every advantage being significant. The lack of significance for the low prior knowledge groups on Test 1 could be attributed to those students needing more help to overcome mathematical barriers and only having two SP sessions before taking that test. The low 13 Exp students also struggled to acclimate themselves to the format and expectations of the SP sessions more than the other students in the experimental group.

Table 5. Means (Standard Deviations) - with respect to Prior Knowledge Levels

	High Control (n=54)	High 13 Exp (n=38)	p-value	Middle Control (n=59)	Middle 13 Exp (n=34)	p-value	Low Control (n=49)	Low 13 Exp (n=42)	p-value
Total Points	678.77 (88.04)	742.75 (72.63)	0.0001	578.59 (160.41)	684.89 (84.90)	0.00003	574.42 (109.51)	656.27 (78.87)	0.00004
Test 1	75.83 (12.99)	81.18 (11.47)	0.02	67.88 (15.18)	73.09 (11.28)	0.032	64.80 (14.25)	67.02 (13.39)	**0.223**
Test 2	77.69 (15.47)	81.97 (12.50)	**0.073**	64.41 (22.32)	74.41 (13.19)	0.004	59.08 (16.06)	67.14 (16.61)	0.011
Test 3	74.17 (14.59)	81.97 (13.98)	0.006	66.19 (19.94)	79.26 (12.86)	0.0001	61.12 (19.02)	73.93 (16.02)	0.0004
Test 4	75.37 (16.25)	83.95 (12.69)	0.003	63.47 (24.48)	78.82 (14.88)	0.0002	64.18 (20.52)	75.83 (12.19)	0.0006
Final Exam	139.26 (31.85)	154.21 (26.27)	0.008	100.17 (54.63)	130.0 (34.20)	0.0008	101.84 938.98)	126.90 (31.660)	0.0005
Quizzes	77.72 (14.30)	86.89 (11.42)	0.0005	69.21 (20.09)	87.60 (9.06)	2E-08	70.82 (16.85)	81.98 (11.99)	0.0002
Course GPA	2.52 (0.99)	3.24 (0.73)	0.00002	1.85 (1.26)	2.62 (0.95)	0.0006	1.61 (0.91)	2.36 (0.93)	0.0001

Each 13 Exp subgroup outperformed the corresponding 8+ Exp group on every assessment (Table 6). Non-significant results are in bold, with $p = 0.05$ used for significance. However, the high 13 Exp group was significantly better for all comparisons and the middle 13 Exp group scored significantly higher for everything except the Final Exam, but the low 13 Exp group only faired significantly better for half of the comparisons.

One important aspect that can be easily overlooked is the comparison between the low 8+ Exp and low 13 Exp groups to the middle control group. Both the low 8+ Exp and 13 Exp groups outperformed the middle control

group on every measure other than Test 1. While the low 8+ Exp group was not significantly higher than the middle control group on any comparison, the 13 Exp group was significantly higher than the control group on quizzes ($p = 0.0008$), Test 3 ($p = 0.0338$), Test 4 ($p = 0.0040$), total points ($p = 0.0005$), and GPA ($p = 0.0059$).

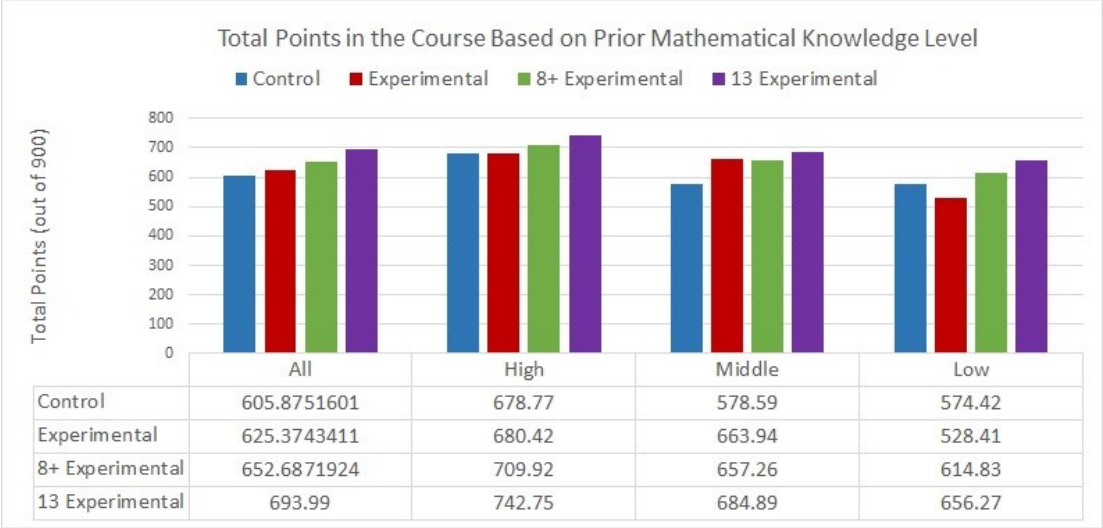
Table 6. Means (Standard Deviations) with respect to Prior Knowledge Levels

	High 8+ Exp (n=78)	High 13 Exp (n=38)	p-value	Middle 8+ Exp (n=79)	Middle 13 Exp (n=34)	p-value	Low 8+ Exp (n=104)	Low 13 Exp (n=42)	p-value
Total Points	709.92 (86.15)	742.75 (72.63)	6E -09	657.26 (101.60)	684.89 (84.90)	0.0039	614.83 (124.73)	656.27 (78.87)	0.0365
Test 1	76.09 (12.55)	81.18 (11.47)	7 E -08	70.38 (14.20)	73.09 (11.28)	0.0026	62.16 (16.92)	67.02 (13.39)	**0.2748**
Test 2	79.29 (12.58)	81.97 (12.50)	5 E -06	69.87 (15.23)	74.41 (13.19)	0.0220	65.58 (19.26)	67.14 (16.61)	**0.3739**
Test 3	79.36 (14.65)	81.97 (13.98)	0.0002	74.56 (15.36)	79.26 (12.86)	0.0007	66.63 (19.16)	73.93 (16.02)	**0.0581**
Test 4	79.87 (28.36)	83.95 (12.69)	5 E -06	73.73 (19.29)	78.82 (14.88)	0.0034	66.83 (22.17)	75.83 (12.19)	0.0025
Final Exam	149.87 (28.76)	154.21 (26.27)	2 E -06	126.20 (36.70)	130.0 (34.20)	**0.267**	117.98 (43.50)	126.90 (31.660)	**0.3038**
Quizzes	82.56 (12.89)	86.89 (11.42)	2 E -09	79.54 (17.51)	87.60 (9.06)	3 E -07	75.56 (19.64)	81.98 (11.99)	0.0012
Course GPA	2.91 (0.97)	3.24 (0.73)	0.0068	2.38 (1.11)	2.62 (0.95)	0.0221	1.98 (1.18)	2.36 (0.93)	0.072

Comparing the Control, Experimental, 8+ Experimental, and 13 Experimental Groups

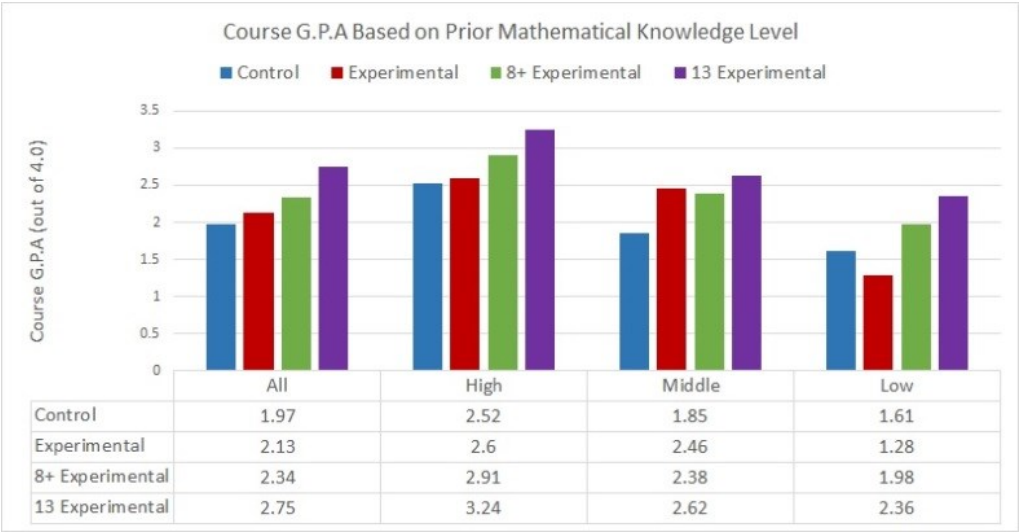
The overall experimental group was not used in the earlier comparisons because attending the SP sessions was optional, so the overall group included students who had minimal or no exposure to the worked examples. Including those students would skew the data because they did not receive the experimental treatment. However, including those students here allows for a better comparison overall, even without measuring significance. The question-and-answer sessions for the control group were not optional, so all of those students attended at least some to them.

Figure 1 shows the comparison for total points in the course (excluding attendance and extra credit) for each of the four groups overall and for each of the prior mathematical knowledge levels (high, middle, and low) for the four groups. Figure 2 compares the course grade point average using the same groupings as Figure 1. Figure 3 compares each test, the Final Exam, and quizzes for the four main groups. The comparisons in Figure 3 are replicated for the high (Figure 4), middle (Figure 5), and low (Figure 6) prior mathematical knowledge groups in each of the four main groups.



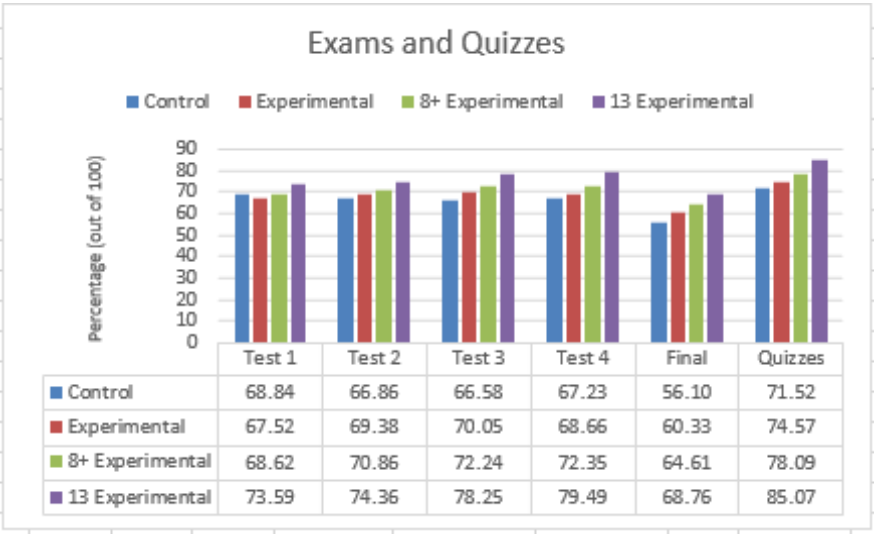
	All	High	Middle	Low
Control	605.8751601	678.77	578.59	574.42
Experimental	625.3743411	680.42	663.94	528.41
8+ Experimental	652.6871924	709.92	657.26	614.83
13 Experimental	693.99	742.75	684.89	656.27

Figure 1. Total Points - All Students & Prior Mathematical Knowledge Groups



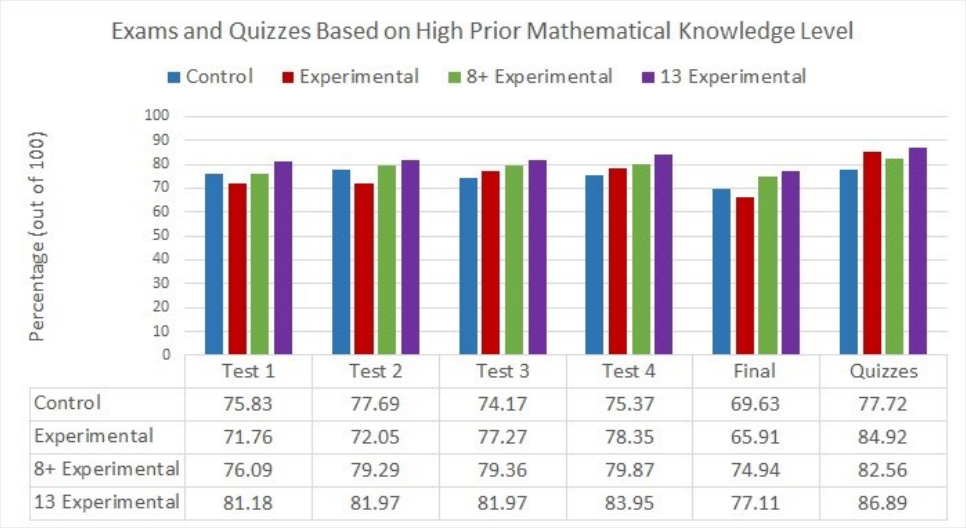
	All	High	Middle	Low
Control	1.97	2.52	1.85	1.61
Experimental	2.13	2.6	2.46	1.28
8+ Experimental	2.34	2.91	2.38	1.98
13 Experimental	2.75	3.24	2.62	2.36

Figure 2. Course GPA - All Students and Prior Mathematical Knowledge Groups



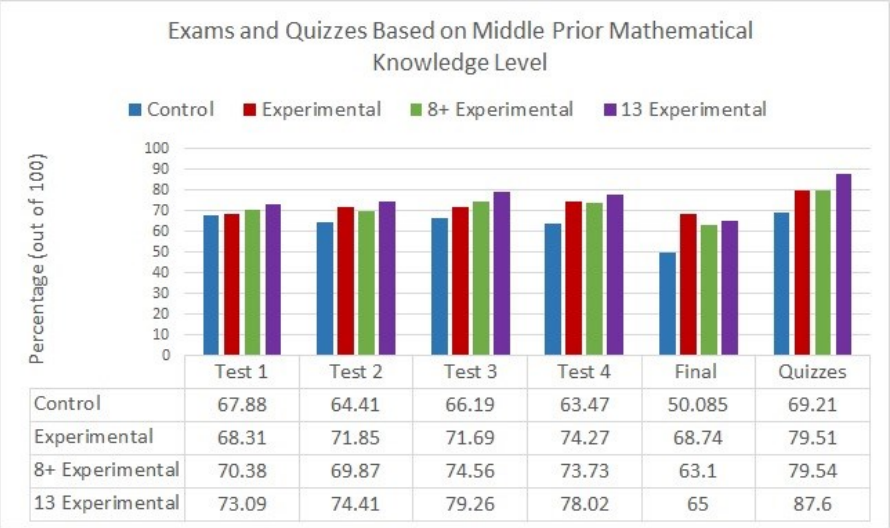
	Test 1	Test 2	Test 3	Test 4	Final	Quizzes
Control	68.84	66.86	66.58	67.23	56.10	71.52
Experimental	67.52	69.38	70.05	68.66	60.33	74.57
8+ Experimental	68.62	70.86	72.24	72.35	64.61	78.09
13 Experimental	73.59	74.36	78.25	79.49	68.76	85.07

Figure 3. Tests/Quiz Performance - All Prior Knowledge Groups



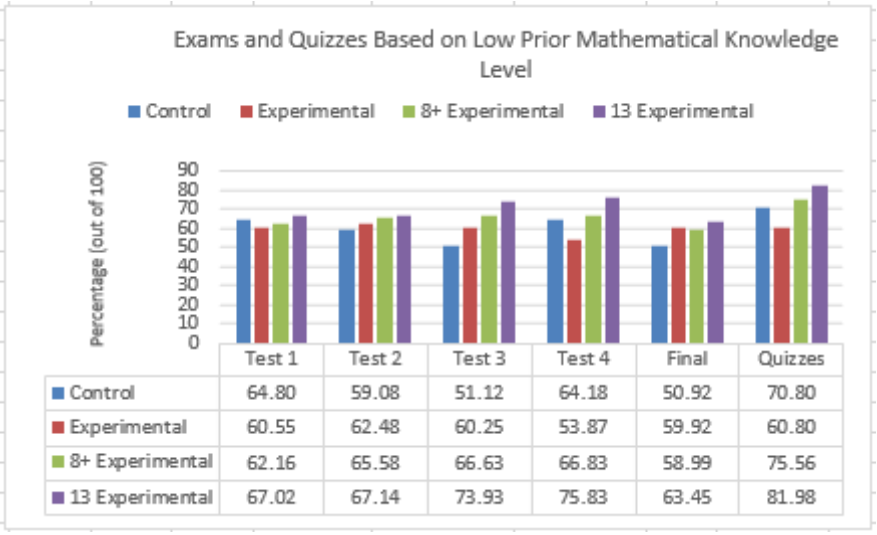
	Test 1	Test 2	Test 3	Test 4	Final	Quizzes
Control	75.83	77.69	74.17	75.37	69.63	77.72
Experimental	71.76	72.05	77.27	78.35	65.91	84.92
8+ Experimental	76.09	79.29	79.36	79.87	74.94	82.56
13 Experimental	81.18	81.97	81.97	83.95	77.11	86.89

Figure 4. Tests/Quizzes – All High Prior Mathematical Knowledge Groups



	Test 1	Test 2	Test 3	Test 4	Final	Quizzes
Control	67.88	64.41	66.19	63.47	50.085	69.21
Experimental	68.31	71.85	71.69	74.27	68.74	79.51
8+ Experimental	70.38	69.87	74.56	73.73	63.1	79.54
13 Experimental	73.09	74.41	79.26	78.02	65	87.6

Figure 5. Tests/Quizzes – All Middle Prior Mathematical Knowledge Groups



	Test 1	Test 2	Test 3	Test 4	Final	Quizzes
Control	64.80	59.08	51.12	64.18	50.92	70.80
Experimental	60.55	62.48	60.25	53.87	59.92	60.80
8+ Experimental	62.16	65.58	66.63	66.83	58.99	75.56
13 Experimental	67.02	67.14	73.93	75.83	63.45	81.98

Figure 6. Tests/Quizzes - All Low Prior Mathematical Knowledge Groups

Discussion

Cognitive load theory stipulates that reducing overall strain on working memory will likely result in better comprehension because it provides a chance for the information to be transferred to long-term memory (Sweller, 1988; Cooper, 1998). More specifically, intrinsic and extrinsic cognitive load should be minimized (Ayres, 2006; Cierniak et al., 2009; Große & Renkl, 2006), while germane cognitive load is kept high enough to keep students interested and engaged (Sweller, 2006; Cierniak et al., 2009; Ayres, 2006; Sweller et al., 1998). Worked examples (in various formats and presentations) have effectively accomplished this (Sweller & Owen, 1989; Sweller & Cooper, 1985; Zhu & Simon, 1987; Chi et al., 1989; Ward & Sweller, 1990; Catrambone & Yuasa, 2006; Crippen & Boyd, 2007; Hilbert et al., 2004; Schwonke et al., 2009; Reisslein et al., 2006; Miller and Schraeder, 2015).

The control group scored significantly higher than the 8+ Exp group on the Pre-ACT, but the averages were nearly equal on the Post-ACT, so the 8+ Exp group was able to overcome the advantage in prior knowledge that the control group had at the beginning of the study. The 8+ Exp group also scored significantly higher than the control group almost universally: Test 2, Test 3, Test 4, Final Exam, quizzes, total points, and course GPA. The only exception was Test 1, but the 8+ Exp group only had a few SP sessions before that assessment, so the students did not have a chance to benefit from the worked examples yet. The 13 Exp group was similarly able to outperform the 8+ Exp group, even on Test 1. However, not all students are created equally. Due to the Expertise Reversal Effect (Kalyuga et al., 2003), no single type of worked example will be the most effective for all students (Ayres, 2006, Reisslein et al., 2006). Modular view was superior to molar view for all students (Gerjets et al., 2004; Sweller, 2006), but the example-practice format worked better for weaker students, while the practice-example format benefited stronger students (Reisslein et al., 2006). Weaker students fared better with isolated worked examples, but stronger students had more success with integrated worked examples (Ayres, 2006). The current study showed that all students received an advantage from the worked examples, but some students received more of a benefit than others.

The high 8+ Exp, middle 8+ Exp, and low 8+ Exp prior knowledge groups all produced higher averages on every assessment than the corresponding subgroups from the control group. However, the middle 8+ Exp prior knowledge group saw more profound results than the high 8+ Exp and low 8+ Exp groups, which corroborates the findings of Miller and Schraeder (2013). One aspect that is easy to overlook is that the low 8+ Exp group produced higher averages than the middle control group on everything other than Test 1, and the middle 8+ Exp group had a higher average on the quizzes than the high control group. The results were more dramatic when comparing the 13 Exp subgroups to the control subgroups. The middle 13 Exp group produced significantly higher averages than the middle control group for every measure, and was higher than the high control group for total points, Test 3, Test 4, quizzes, and course GPA. The high 13 Exp group significantly outperformed the high control group on all but Test 2. The low 13 Exp group significantly outperformed the low control group on all but Test 1 and scored higher than the middle control group on all but Test 1 (sometimes significantly).

As with the other comparisons, the 13 Exp subgroups all performed better than the respective 8+ Exp

subgroups. However, the high 13 Exp group fared slightly better than the middle 13 Exp group in those comparisons. The high 13 Exp group scored significantly higher than the high 8+ group on all assessments; whereas, the middle 13 Exp group was not significantly higher than the middle 8+ group on the Final Exam, but significant on the rest. The low 13 Exp group was able to score higher than the middle 8+ group on Test 4, the Final Exam, and quizzes. The low 8+ Exp and low 13 Exp groups were even able to outperform the middle control group, the low 13 Exp group significantly so on several measures, indicating that the SP sessions can compensate for a lack of prior mathematical knowledge at times.

When comparing all four groups (control, experimental, 8+ Exp, and 13 Exp), the students' performances in the course improved as the number of SP days attended increased for many of the comparisons. The 8+ Exp and 13 Exp groups outperformed the control on total points (out of 900) by nearly 50 points and 90 points, respectively. On the tests and the Final Exam, the 8+ Exp and 13 Exp groups mainly outperformed the control group, with the difference in scores growing as the semester progressed. A similar pattern was observed when comparing the 8+ Exp and 13 Exp groups to just the experimental group. As a result of this better performance, the students earned better grades in the course (course GPA) as they attended more SP days. Similar results occurred when prior knowledge level was examined, except that the low control group often outperformed the low experimental group.

In summary, the 8+ Exp group primarily outperformed the control group. The 13 Exp group had higher averages than both the 8+ Exp and control groups. This was largely seen with the high, middle, and low prior knowledge subgroups, as well. However, the middle 8+ Exp group benefitted more than high 8+ Exp and low 8+ Exp groups when compared to the control prior knowledge subgroups. All three of the 13 Exp subgroups dominated the control subgroups. This dominance was also seen in the high 13 Exp and middle 13 Exp groups over their 8+ Exp counterparts, but the low 13 Exp group did not produce such a performance compared to the low 8+ group. Therefore, as found by Miller and Schraeder (2015), the worked examples had the greatest impact on the middle prior knowledge group in general. However, attending more SP sessions and having exposure to more worked examples seemed to benefit the high prior knowledge group, as seen with the comparisons between the high 13 Exp and high 8+ Exp groups. While the low 8+ Exp group's grades were not always significantly higher than the low control group and low 13 Exp group was similarly better (but not significantly better) than the low 8+ Exp group, the low 8+ Exp and low 13 Exp groups often outperformed the middle control groups, so those low groups were able to overcome disadvantages in prior knowledge.

Conclusion

Many college instructors teach large lecture sections of introductory mathematics classes and struggle with high percentages of students who earn grades of D or F, or simply withdraw from the class (DFW rate). Recitations, out-of-class help sessions, and tutoring all required additional resources. These are familiar modes of intervention that colleges use to lower the DFW rate and, in general, help students be more successful in learning the material in a course. This study shows that carefully designed worksheets modeled after worked examples coupled with active group sessions in an extra day of class can be very beneficial in helping students

become more successful, regardless of prior knowledge in a course. An SP day each week provides an extra active-learning session in which students can work to comprehend the material while in groups. These SP days are like an extra day of class that emphasizes only the material that has been covered during prior lectures. The instructor is obligated to have a group of class assistants ready to help students with the material, but no new instruction is necessary.

This extra day of class each week is very important because, for the most part, students will show up for the extra day of class that is included in a class schedule for the semester and actively participate compared to an out-of-class session. Moreover, students are usually less likely to visit the instructor during office hours, or go to a mathematics tutoring center. For the instructor, SP days are the most efficient way to help many students at the same time and can be thought of as an office hour with the whole class. There would be no way for the instructor to help this many students during office visits and reduces the task of explaining material multiple times to different students if they do attend office hours. Students also review the worked example worksheets outside of class, which act like a tutor by presenting students with a number of examples and problems to practice. Furthermore, students who attended every extra day of active class were more successful in the class regardless of their level of prior knowledge. Instructors are often concerned that interventions only help certain groups of students, but this study has shown that all students were positively affected when they attended every SP day. Conversely, the middle prior knowledge group benefited more than the high and low groups among students who attended 8+ days of SP. Instructors can consider mandating that all students attend all SP days. It is in the opinion of the authors that one would see similar results from mandatory attendance, but not as profound because students who voluntarily attended every SP day did so of their own accord, so they were more likely to complete the work and exert the requisite effort.

This study can provide some insight into teaching using worked example worksheets embedded into a class as a form of active learning to help students become more successful learners. The authors believe that the worksheets provide a basis that helps students when studying and completing other assignments, which is supported by the students' comments on an end-of-the-semester survey. Because students with different abilities benefit from different types of instructional materials, designing course components that motivate all students is very important. One strategy is to provide more scaffolding for the weaker students in the form of additional interventions (perhaps mandatory) that are tailored towards those students. Specifically, activities could be designed to help students transition from dealing with isolated worked examples to integrated worked examples (Sweller, 2010), possibly by gradually reducing the amount of assistance provided on the worked examples from full examples to partial examples to full problems (Schwonke et al., 2009; Sweller, 2010). The desired result is to train students to monitor their own understanding through self-explanations (Chi et al., 1989; Große & Renkl, 2006). The stronger students will likely succeed with worked example worksheets that are integrated and will not need the additional worksheets. As the weaker students learn to isolate individual steps by themselves, they can be weaned off of the extra worksheets.

Finally, the authors believe that the worked example worksheets could be expanded into worksheets for each section covered in the class and used in a large lecture inquiry-learning-based class. Instructors might have to

teach a mini-lecture and use the majority of the class time to allow students to complete the worked example worksheets to build understanding. A "flipped" classroom could create a totally active learning environment in which the students watch short video lectures (mini-lectures) outside of class and work actively on worked example worksheets during class. This added support should also benefit weaker students.

Acknowledgements

We would like to acknowledge all the students and class assistants that have helped over the years with the active worked example sessions.

References

Alexander, J. M., Carr, M., & Schwanenflugel, P. J. (1995). Development of metacognition in gifted children: Directions for future research. *Developmental Review, 15,* 1-37.

Arendale, D.R. (1994, Winter). Understanding the supplemental instruction model. New Directions for Teaching and Learning. *Jossey-Bass Publishers, 60*, 11-21.

Ayres, P. (2006). Impact of reducing intrinsic cognitive load on learning in a mathematical domain. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(3), 287-298.

Brody, N. (1992). Intelligence. (2nd ed.). San Diego: Academic Press. Business-Higher Education Forum. *A Commitment to America's Future: Responding to the Crisis in Mathematics and Science Education*. Washington, DC: American Council on Education, 2005.

Business-Higher Education Forum. (2005). *A commitment to America's future: Responding to the crisis in mathematics & science education*. ERIC Clearinghouse.

Butler, M., Pyzdrowski, L., Walker, V., & Butler, F. (2012). The Effects of the Immediate Online Homework Feedback on Homework Attempts and Mathematics Performance, Chapter 7 in *Advancement in Online Education: International Perspectives and Practices*, Volume 2, published by Nova Science Publishers, New York, USA.

Carroll, W.M. (1994). Using Worked Examples as an Instructional Support in the Algebra Classroom. *Journal of Educational Psychology.* 86(3), pp. 360 – 367.

Catrambone, R., & Yuasa, M. (2006). Acquisition of procedures: The effects of example elaborations and active learning exercises. *Learning and Instruction*. 16, 139-153.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science.* 13, 145-182.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, *25*(2), 315-324.

Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW.

Crippen, K., & Boyd, E. (2007). The impact of Web-based Worked Examples and Self-Explanation on Performance, Problem Solving, and Self-efficacy. *Computers & Education.* 49(3), 809-821.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of educational psychology*, *100*(1), 223.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Garner, R. (1987). *Metacognition and reading comprehension.* Norwood, NJ: Ablex Publishing.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science*, *32*(1), 33-58.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition, 15*, 84-93.

Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner & R. C. Calfee (Eds.), *The handbook of educational psychology* (pp. 63-84). New York, NY: Macmillan.

Große, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction, 16*(2), 122-138.

Haver, W. (2007). *Renewal of College Algebra. In a MAA report Algebra: Gateway to a Technological Future*. Edited by Victor Katz. Published by the Mathematical Association of America. Available at http://www.maa.org/algebra-report/Algebra-Gateway-Tech-Future.pdf

Herrnstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Hilbert, T.S., Schworm, S., & Renkl, A. (2004). Learning From Worked-out Examples: The Transition from Instructional Explanations to Self-Explanation Prompts. In Gerjets, P., Kirschner, P.A., Elen, J., & Joiner, R. (Eds) (2004). *Instructional Design for Effective and Enjoyable Computer-supported Learning.* Proceedings of the EARLISIGs Instructional Design and Learning and Instruction with Computers (CD-ROM) Tuebingen: Knowledge Media Research Center.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23-31.

Miller, D., & Schraeder, M. (2013). Using Cognitive Science with Active Learning in a Large Lecture College Aglebra Course. In (Eds.) S. Brown, G. Karakok, K.H. Roh, and M. Oehrtman, *Proceedings of the 16th Annual Conference on Research in Undergraduate Mathematics Education*, Denver, CO.

Miller, D., & Schraeder, M. (2015). Research on Group Learning and Cognitive Science: A Study of Motivation, Knowledge, and Self-Regulation in a Large College Algebra Class. *Mathematics Educator, 24*(2), 27-55.

Model of Learning. http://dwb.unl.edu/Chau/CompMod.html (accessed 2/10/2011)

Perkins, D.N., J.W. Segal, J.F. Voss. (Eds.) (1987). *Informal reasoning and education* (pp. 83-106). Hillsdale, NJ: Erlbaum.

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*(2), 92-103.

Schraw, G.P., & Brooks, D.W. (1999). *Improving College Using Interactive, Compensatory*

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*, 359-382.

Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., & Salden, R. (2009). The worked-example effect: Not an artefact of lousy control conditions. *Computers in human behavior*, *25*(2), 258-266.

SI staff. (1997). *Description of the Supplemental Instruction Program. Review of Research Concerning the Effectiveness of SI from The University of Missouri- Kansas City and Other Institutions from Across the United States.*

Steinberg, R.J., & Wagner, R.K. (Eds.) (1988). *Mind in context: Interactionist perspectives on human intelligence.* Cambridge, UK: Cambridge University Press.

Sullivan, M., & Sullivan, M. (2006). *College Algebra and Trigonometry: Enhanced with Graphing Utilities*. (6th Ed.). Upper Saddle River, NJ, Prentice-Hall.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science,* 12, 257 – 285.

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction, 16*(2), 165-169.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, *22*(2), 123-138.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and instruction*, *12*(3), 185-233.

Sweller, J., & Cooper, G. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction.* v2, n1, 59 – 89.

Sweller, J., & Owen, E. (1989). Should problem solving be used as a learning device in mathematics? *Journal for Research in Mathematics Education. 20*(3), 322–328

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296.

Tarmizi, R.A., and Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80(4), 424- 436.*

Ward, M., & Sweller, J. (1990) Structuring Effective Worked Examples. *Cognitive and Instruction, 7*(1), 1-39.

Zhu, X., & Simon, H. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction 4*(3), 137-166.

Author Information

David Allan Miller	**Matthew Schraeder**
https://orcid.org/0000-0001-9469-9374	https://orcid.org/0000-0002-7107-0546
West Virginia University	West Virginia University
1500 University Avenue	1500 University Avenue
Morgantown, West Virginia 26506	Morgantown, West Virginia 26506
United States	United States
Contact e-mail: *millerd@mail.wvu.com*	