

2009-08-28

Investigating Learning in an Intelligent Tutoring System through Randomized Controlled Experiments

Leena Razzaq
Worcester Polytechnic Institute

Follow this and additional works at: <https://digitalcommons.wpi.edu/etd-dissertations>

Repository Citation

Razzaq, L. (2009). *Investigating Learning in an Intelligent Tutoring System through Randomized Controlled Experiments*. Retrieved from <https://digitalcommons.wpi.edu/etd-dissertations/370>

This dissertation is brought to you for free and open access by [Digital WPI](#). It has been accepted for inclusion in Doctoral Dissertations (All Dissertations, All Years) by an authorized administrator of Digital WPI. For more information, please contact wpi-etd@wpi.edu.

Investigating Learning in an Intelligent Tutoring System through Randomized Controlled Experiments

By

Leena Razzaq

A Dissertation

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

in

Computer Science

by

August 2009

APPROVED:

Dr. Neil T. Heffernan
Advisor

Dr. David C. Brown
Committee Member

Dr. Robert W. Lindeman
Committee Member

Dr. Beverly Park Woolf
Univ. of Massachusetts - Amherst
External Committee Member

Dr. Michael A. Gennert
Head of Department

Abstract

In the United States, many students are doing poorly on new high-stakes standards-based tests that are required by the *No Child Left Behind Act* of 2002. Teachers are expected to cover more material to address all of the topics covered in standardized tests, and instructional time is more precious than ever. Educators want to know that the interventions that they are using in their classrooms are effective for students of varying abilities.

Many educational technologies rely on tutored problem solving, which requires students to work through problems step-by-step while the system provides hints and feedback, to improve student learning. Intelligent tutoring researchers, education scientists and cognitive scientists are interested in knowing whether tutored problem solving is effective and for whom. Intelligent tutoring systems have the ability to adapt to individual students but need to know what types of feedback to present to individual students for the best and most efficient learning results.

This dissertation presents an evaluation of the ASSISTment System, an intelligent tutoring system for the domain of middle school mathematics. In general, students were found to learn when engaging in tutored problem solving in the ASSISTment System. Students using the ASSISTment System also learned more when compared to paper-and-pencil problem-solving.

This dissertation puts together a series of randomized controlled studies to build a comprehensive theory about when different types of tutoring feedback are more appropriate in an intelligent tutoring system. Data from these studies were used to analyze whether interactive tutored problem solving in an intelligent tutoring system is

more effective than less interactive methods of allowing students to solve problems. This dissertation is novel in that it presents a theory that designers of intelligent tutoring systems could use to better adapt their software to the needs of students. One of the interesting results showed is that the effectiveness of tutored problem solving in an intelligent tutoring system is dependent on the math proficiency of the students. Students with low math proficiency learned more when they engaged in interactive tutoring sessions where they worked on one step at a time, and students with high math proficiency learned more when they were given the whole solution at once. More interactive methods of tutoring take more time versus less interactive methods. The data showed that it is worth the extra time it takes for students with low math proficiency.

The main contribution of this dissertation is the development of a comprehensive theory of when educational technologies should use tutored problem solving to help students learn compared to other feedback mechanisms such as hints on demand, worked out solutions, worked examples and educational web pages.

Acknowledgements

There are many people that I would like to acknowledge in the completion of this dissertation. Over the course of working on this dissertation, Neil Heffernan contributed a significant amount of time advising me and I want to thank him for his generous support and patience. His energy, enthusiasm and valuable advice were a great help. He makes it easy to believe that ASSISTments will take the world by storm.

I would also like to thank my dissertation committee, David Brown, Rob Lindeman and Beverly Woolf. Their expertise and experience have all contributed valuable advice and support.

The ASSISTment team has been a great group of people to work with. Although the ASSISTment lab is small and crowded, I have always felt lucky to work with such smart and talented people. Ming, Zach, Jozsef, Mike, Cristina and so many others have been friends as well as colleagues.

Thanks to all of the family and friends who supported me. Finally, a special note of thanks to Muttasem, Fatima and Zane for their love and support during my time at WPI.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
Table of Figures.....	vii
1. Introduction.....	3
1.1 Motivation.....	3
1.2 The System.....	5
1.3 Research Questions.....	5
1.4 Contributions.....	6
1.4.1 Evaluate student learning in the ASSISTment System.....	6
1.4.2 Investigate the effectiveness of tutored problem solving.....	7
1.4.3 Designed tutoring content.....	8
1.5 Dissertation Outline.....	10
2. Background on Tutoring.....	11
2.1 Human Tutoring.....	11
2.2 Computer-based tutoring.....	13
2.2.1 Computer-Aided Instruction.....	13
2.2.2 Intelligent Tutoring Systems.....	14
2.2.3 Dialog-based Intelligent Tutors.....	17
2.2.4 Interactive Tutoring.....	19
3. The ASSISTment System.....	22
3.1 Structure of an ASSISTment.....	23
3.2 Skill mapping.....	27
3.3 Teacher Reports.....	28
3.4 Problem Sequences.....	29
4. Is the ASSISTment System an effective tutor?.....	31
4.1 ASSISTment development and usage.....	31
4.2 Are students learning from the ASSISTment system?.....	35
4.3 ASSISTments with tutoring/help vs. paper-and-pencil homework.....	37
4.3.1 Participants.....	40
4.3.2 Content.....	41
4.3.3 Experiment Design.....	42
4.3.4 Procedure.....	43
4.3.5 Results.....	45
4.3.6 Discussion.....	46
4.4 ASSISTments with no tutoring vs. paper and pencil homework.....	46
4.4.1 Participants and Content.....	47
4.4.2 Experiment Design.....	47
4.4.3 Results.....	48
4.4.4 Conclusion.....	49
5. Tutored Problem Solving vs. Hints on Demand.....	51
5.1 Participants.....	55
5.2 Results.....	55
5.3 Hints on demand vs. proactive hints.....	58
5.4 Conclusion.....	62

6. Tutored problem solving vs. worked examples	64
6.1 Participants.....	66
6.2 Materials	67
6.3 Procedure	67
6.4 Results.....	72
6.5 Conclusions.....	74
7. Tutored problem solving vs. worked out solutions.....	77
7.1 Experiment Design.....	79
7.2 Results.....	81
7.3 Immediate vs. Delayed Feedback	84
7.3.1 Experiment Design.....	85
7.3.2 Results.....	87
7.4 Conclusion	87
8. Tutored problem solving vs. educational web pages	89
8.1 Can students learn from viewing a web page of educational content?.....	89
8.2 Is Tutored Problem Solving superior to viewing educational web pages?.....	91
8.3 Conclusion	93
9. Is there a tradeoff between the value of interactive tutoring and the extra time it takes?	94
9.1 Experiment Design.....	95
9.2 Results.....	98
9.3 Conclusions.....	100
10. Theory of the effectiveness of tutored problem solving.....	102
10.1 Is interactive tutored problem solving to help students solve math problems more effective than less interactive forms of help?	102
11. Conclusions.....	107
11.1 Implications.....	108
11.2 Directions for future work	108
11.2.1 Switching between strategies for individual students.....	108
11.2.2 Examining other student characteristics	109
11.2.3 Conceptual versus procedural knowledge	109
11.2.4 Ranking educational web content	109
References.....	111
Appendix A: Materials for ASSISTments vs. Paper and Pencil Study	
Appendix A.1 Pretests/Post-tests	
Appendix A.2: Homework Problem Sets	
Appendix A.3: Sample ASSISTment Problems	
Appendix A.4: Sample Hints	
Appendix B Materials for Experiment on Tutored Problem Solving vs. Worked Out Solutions	
Appendix B.1 Pretest	
Appendix B.2 Tutored Problem Solving Condition	
Appendix B.3 Hints on demand condition	
Appendix B.4 Worked Solutions Condition	
Appendix B.5 Post-test	

Table of Figures

Figure 1: The ASSISTment Builder and associated student screen	25
Figure 2: An item report tells teachers how students are doing on individual problems..	29
Figure 3: A Problem Sequence that is designed as a randomized controlled experiment	30
Figure 4: Item #19 from the 2003 Massachusetts state math test for 8th grade	33
Figure 5: 8th grade problem sets targeting skills set by the state of Massachusetts	34
Figure 6: A perimeter and area learning opportunity pair	37
Figure 7: Students were given two attempts to enter the correct answer.....	49
Figure 8: The tutored problem solving condition with three scaffolding questions is shown on the left. Four hints are shown on the right. Students ask for each hint by clicking on the hint button.	54
Figure 9: Results for average on post-test items by condition.....	56
Figure 10: Results on the transfer item for the first pre-test item by condition.....	57
Figure 11: Hints on demand, students ask for each hint by clicking on a hint button.....	60
Figure 12: Proactive hints: hints are presented automatically when a student submits an incorrect answer	61
Figure 13: The worked example condition requires students to read the example and then try to answer the question again.....	70
Figure 14: The tutored problem solving condition requires students to work through each step of the problem.....	71
Figure 15: An ASSISTment item showing the tutored problem solving and hints on demand conditions.	78
Figure 16: The delayed feedback explanation	80
Figure 17: There are interactions between condition and math proficiency approaching significance.	83
Figure 18: The immediate feedback condition gave students the solution immediately while the delayed feedback condition gave the student the solution at the end of the assignment.....	86
Figure 19: Students would be directed to a webpage on Venn Diagrams if they answered the problem incorrectly.	90
Figure 20: A symbolization problem shows the TPS approach on the left and the Solutions approach on the right.	97
Figure 21: Highly proficient students appeared to learn more by seeing Solutions, and students with low proficiency learned more by doing TPS.	100
Figure 22: Effect sizes of tutored problem solving over other strategies	106

1. Introduction

1.1 Motivation

In many states, students are doing poorly on new high-stakes standards-based tests that are required by the *No Child Left Behind Act* of 2002. For instance, Massachusetts has an assessment system, called Massachusetts Comprehensive Assessment System (MCAS), in which students take rigorous tests in English, Mathematics, History & Social Science and Science & Technology/Engineering in grades 3-12. Students must pass the Mathematics and English portions of the 10th grade MCAS test in order to get a high school diploma. In Massachusetts, which according to the National Assessment of Educational Progress (NAEP) was tied for first in 8th grade math performance, over 60% of 8th grade students did not reach the “Proficient” level in math in the Massachusetts state test in 2005-2006 according to the Massachusetts Department of Education. The problem is even more acute for minorities and those who come from low-income families. In the industrial city of Worcester, Massachusetts, 56% of 8th graders failed the Massachusetts state test in mathematics and 82% did not reach the Proficient level in 2005-2006.

In 2007, a full 10% of high school seniors were set to be denied a high school diploma due to having failed to pass the test on their fourth try. The failure rate for African Americans and Latinos was a disheartening 25% and 30%, respectively. Because students are more likely to fail the mathematics portion of the test, the state is focusing efforts on mathematics. The state of Massachusetts has singled out student performance

on the 8th grade math test as an area of highest need for improvement (MCAS Results - Massachusetts Comprehensive Assessment System).

The Worcester Public School District is representative of the many districts across the country that are struggling to address these problems. With new advances in educational technology and numerous studies showing varying degrees of its effectiveness, there are more choices for school administrators and teachers to make regarding which tools to use to help students learn and stay motivated. Complicating matters further, a study (Dynarski, Agodini, Heaviside, Novak, Carey, & Campuzano, 2007) commissioned by the U.S. Department of Education presented results that claimed that classrooms using selected math and reading educational software products did not differ significantly on standardized tests from classrooms that did not use the products. With the notoriety of this study, some have called into question the utility of education software.

The goal of this research is to use randomized controlled experiments to evaluate an intelligent tutoring system and to determine empirically-validated effective ways to interact with students via the computer. This will help inform the design of the tutoring strategies that can maximize learning results. This becomes more pertinent as teachers are expected to cover more material to address all of the topics covered in standardized tests, instructional time becomes more precious and teachers want to know that the interventions that they are using in their classrooms are effective for students of varying abilities.

1.2 The System

This research was carried out with the ASSISTment System (Razzaq et al., 2007), a web-based tutoring system that assesses students while it assists them in solving math and science problems. Assistance and assessment are integrated in the ASSISTment System which offers instruction to students while providing a detailed evaluation of their abilities to teachers. Each time students work on the website, the system “learns” more about the students’ abilities. The ASSISTment System is being built to identify the difficulties individual students – and the class as a whole – are having, and teachers will be able to use this detailed feedback to tailor their instruction to focus on those difficulties. Unlike other assessment systems, the ASSISTment system also provides students with intelligent tutoring assistance while assessment information is collected.

Over 3,000 students use the system as part of their regular math instruction. This work is mainly based on evaluations of middle school students using the ASSISTment System in Central Massachusetts.

1.3 Research Questions

This dissertation attempts to answer the following questions:

1. Is the ASSISTment System an effective tutor?
 - Do students learn when using the ASSISTment System in the classroom?
 - Can students learn more from using the ASSISTment System for homework compared to traditional paper and pencil homework?
2. Is the ASSISTment System’s method of interactive tutored problem solving to help students solve math problems more effective than other forms of help? Such as:

- Tutored problem solving vs. hints on demand
 - Tutored problem solving vs. worked examples
 - Tutored problem solving vs. worked out solutions
 - Tutored problem solving vs. educational web pages
3. Does a student's math proficiency determine what form of help to provide?
 4. Is there a tradeoff between the value of interactive tutoring and the extra time that it takes?

1.4 Contributions

This dissertation makes three main contributions to the field of student learning in intelligent tutoring systems, as described in the following sections. The first contribution concerns designing and developing tutoring content for the ASSISTment System. The second contribution focuses on evaluation of learning in the ASSISTment System. Finally, the third contribution investigates the effectiveness of tutored problem solving.

1.4.1 Evaluate student learning in the ASSISTment System

Since the *No Child Left Behind Act* of 2002, which stresses accountability, educators are concerned with showing that their students are making Adequate Yearly Progress. Adequate Yearly Progress, or AYP, is a measurement defined by the U.S. government that helps to determine how public schools are performing academically according to their students' results on standardized tests.

Intelligent tutoring systems have great potential to be effective learning tools that provide individualized instruction to students of different abilities and needs and help students to prepare for standardized tests. However, evaluations are needed to validate these systems. Are students learning when using the tutoring system? Do they learn more

from the tutoring system than they would from other more traditional approaches such as paper and pencil problem solving? What are the students' attitudes about the system?

I have shown that students do learn significantly when they are using the ASSISTment System to solve math problems. Students also learn significantly more when they are tutored by the ASSISTment System than when they engage in paper and pencil problem solving (Mendicino, Razzaq, & Heffernan, 2009). Student surveys showed that 72% Strongly Agreed or Agreed Somewhat that the ASSISTment system helped them prepare for the MCAS.

The implication of this contribution is that educators can believe that the ASSISTment System is an effective educational technology that can help their students to learn.

1.4.2 Investigate the effectiveness of tutored problem solving

Many educational technologies depend on tutored problem solving to help students learn. The ASSISTment System forces students to engage in tutored problem solving when they cannot solve a problem correctly. While this method can be effective, it is time-consuming and students have complained about being forced to go through each step. Developers of intelligent tutoring systems would like to know how to adapt tutoring to individual students by being able to decide when to use tutored problem solving and for which students.

This dissertation used randomized controlled experiments with hundreds of students to compare tutored problem solving to five other interventions: paper and pencil homework, hints on demand, worked examples, worked solutions and educational web pages. Based on the results of these experiments, this dissertation has shown that tutored

problem solving is an effective method of helping students to learn how to solve math problems. The results also show that tutored problem solving is most effective when the material is difficult for students. For students who have higher math proficiency, tutored problem solving is not the most efficient method. These students benefit more from being shown a worked out solution and from finishing more problems in the same amount of time. Students who have lower math proficiency benefit from spending more time per problem while doing tutored problem solving.

The implication of this contribution is that tutoring system developers can use these guidelines to design systems that adapt to individual students resulting in better and more efficient learning results, perhaps helping to close the achievement gap that affects economically disadvantaged and minority students.

1.4.3 Designed tutoring content

Although intelligent tutors have been shown to produce significant learning gains in students (Koedinger, Anderson, Hadley, & Mark, 1997), few intelligent tutoring systems have become commercially successful. The high cost of building intelligent tutors may contribute to their scarcity and a significant part of that cost concerns content creation. Murray (1999) asked why there are not more intelligent tutoring systems and proposed that a major part of the problem was that there were few useful tools to support intelligent tutoring systems creation. Murray, Blessing, and Ainsworth (2003) reviewed 28 authoring systems for learning technologies. Unfortunately, they found that there were very few authoring systems that were of "release quality", let alone commercially available.

Koedinger et al. (2004) introduced example-tracing tutors which mimic cognitive tutors but are limited to the scope of a single problem. The ASSISTment System uses a further simplified example-tracing tutor, called an ASSISTment, where only a linear progression through a problem is supported which makes content creation easier and more accessible. The ASSISTment System's approach is to allow users to create example-tracing tutors via the web to reduce the amount of time it takes to create content, thus reducing the cost.

The author has been part of the design team for the ASSISTment System since its start in 2003. She has helped to design the Builder which is used to build tutoring content in the ASSISTment System (Razzaq, et al., 2009).

The author has created close to 1,000 ASSISTments to provide tutoring in the ASSISTment System for problems taken from state math tests for grades 4 through 10. These ASSISTments can include scaffolding questions, hints, buggy messages, images, animations and tutoring strategies. These ASSISTments are used regularly by students in their regular math classes as well as for homework. Over 3,000 students use the ASSISTment System in central Massachusetts and Pittsburgh, Pennsylvania.

The author has organized hundreds of ASSISTments into problem sets based on the skills needed to solve the problems. Teachers can browse through the problem sets available for each grade level and assign materials to students based on what they are doing in class or based on what they would like to assess. The author has created problem sets to target the skills for 7th, 8th and 10th grade math as outlined by the Massachusetts Frameworks.

The implication of this contribution is that teachers can use this content to assign problems to their students to assess their progress on topics that are covered in the

Massachusetts state test. Students can use this tutoring content to be tutored on problems that they are having trouble solving. The large amount of tutoring content developed means that most students can use our system throughout the school year (generally used once a week in school) without running out of content.

1.5 Dissertation Outline

The rest of this dissertation is outlined as follows. Section 2 presents prior work on tutoring: both human and computer-based tutoring. Within computer-based tutoring, computer aided instruction, intelligent tutoring systems and dialog-based tutoring systems are discussed. Section 3 discusses the ASSISTment System which was used to carry out the experiments in this dissertation. Section 4 is focused on Research Question 1. Sections 5 through 8 focus on Research Questions 2 and 3. Table 1 shows a summary of the main experiments that will be discussed in these sections. Section 9 focuses on Research Question 4. Section 10 discusses the results of the experiments and develops the guidelines for using tutored problem solving. Section 11 concludes this dissertation.

Table 1: Summary of main experiments

Randomized controlled study	TPS	Paper & Pencil	Hints on Demand	Worked Solutions	Worked Examples	Web pages
Mendicino, Razzaq & Heffernan, 2009	x	x	x			
Razzaq & Heffernan 2006	x		x			
Razzaq, Heffernan & Lindeman, 2007	x		x	x		
Shrestha et al, 2009	x				x	
Unpublished study comparing TPS to web pages	x					x
Razzaq & Heffernan, 2009	x			x		

2. Background on Tutoring

This section describes prior work on human tutoring, computer aided tutoring and intelligent tutoring systems.

2.1 Human Tutoring

In education research, an effect size is used to indicate the number of standard deviation units between the result scores of treatment and control groups in a study. Thus, effect sizes can be used to describe the results of different studies on a uniform scale of effectiveness. For instance, an effect size of 0.5 standard deviations means that the average score in the experimental group is 0.5 standard deviations above the average score in the control group, or is higher than the scores of 49% of the control group. In education, results with effect sizes above 0.25 are considered large enough to be educationally meaningful (Cohen, 1969).

Studies indicate that human tutors provide the most effective form of instruction known (Bloom, 1984; Cohen, Kulik, & Kulik, 1982). They raise the mean performance about two standard deviations compared to students taught in classrooms. Intelligent tutoring systems offer excellent instruction, but not quite as good as human tutors. The best ones raise performance about one standard deviation above classroom instruction (Anderson, Corbett, Koedinger, & Pelletier, 1995). In other words, a human tutor can raise the student's grade by about two letter grades (e.g., from C to A) while a tutoring

system can raise it by about one letter grade (e.g., from C to B). The challenge is to create tutoring systems that are as effective as human tutors.

In studying what makes a tutoring session successful, VanLehn, Siler and Murray (1998) identified two principles for effective teaching. First, tutors should not offer strong hints or solve problems themselves when students make mistakes. Students miss the opportunity to learn how to solve a problem when they are given an answer and are not allowed to reason for themselves. Secondly, a tutor may need to emphasize the rules needed to solve a problem differently based on what is causing the student to have difficulty. For instance, a tutor may need to generalize a rule or explain why a rule applies. On other occasions an explanation of why an error was made does not help with the learning (VanLehn, Siler, Murray, & Baggett, 1998). Furthermore, VanLehn, Siler, Murray, Yamauchi and Baggett (2003) did further research on this topic and suggested that students were more likely to learn when they reached an impasse than if the tutor had applied the correct action for them.

Merrill, Reiser, Ranney and Trafton (1992) compared the effectiveness of human tutors and intelligent tutoring systems. Their study indicated that a major reason that human tutors are more effective is that they let the students do most of the work in overcoming impasses, while at the same time provided as much assistance as necessary. Merrill, D., Reiser, Merrill, S., and Landes (1995) argue that the main thing human tutors do is to keep students on track and prevent them from following “garden paths” of reasoning that are unproductive and unlikely to lead to learning. Merrill et al. (1995) pointed to the large number of remarks made by tutors that helped keep students on track while learning Lisp programming. They argue that this was especially important in this

domain because of the hierarchical nature of computer programming where if a student does one step wrong then everything after that step is likely to be wrong.

Modeling, coaching, and scaffolding (demonstrating, overseeing and supporting a student's learning) are described by Collins, Brown and Holum (1991) as the heart of cognitive apprenticeship (a model of instruction that works to make thinking visible), which they claim "help students acquire an integrated set of skills through processes of observation and guided practice." Tutors use *modeling* when they do a task in front of students to demonstrate how it is done, allowing the student to observe the processes needed. When students attempt problems themselves, tutors may *coach* them by giving some feedback designed to remind, hint or clarify. *Scaffolding* is when tutors do parts of a problem for students that they are struggling with and involves being able to assess a student's proficiency at that point. An important part of scaffolding is *fading*, which entails progressively removing the support of scaffolding as the student demonstrates proficiency (Collins, Brown, & Holum, 1991).

2.2 Computer-based tutoring

2.2.1 Computer-Aided Instruction

The first computer-based tutoring systems appeared over thirty years ago with the goal of approaching the effectiveness of human tutors. According to Corbett & Trask (2000) these systems, called computer-assisted instruction (CAI) offered one advantage of human tutors: individualized interactive learning support. While these systems were interactive and provided explicit instruction in the form of long web pages or lectures they offered no dialog.

Studies demonstrated the effectiveness of CAI in mathematics at the elementary level (Burns & Bozeman, 1981; Kulik, Kulik, & Bangert-Drowns, 1985) secondary level (Kulik, Bangert-Drowns, & Williams, 1983) and college level (Kulik & Kulik, 1986). In a meta-analysis of 254 studies involving CAI, Kulik and Kulik (1991) found that CAI improved student achievement by an average effect size (a measure of the magnitude of the treatment effect) of 0.3 over students receiving conventional instruction. In another meta-analysis, Kulik (1994) summarized 97 studies from the 1980's that compared classroom instruction to computer-based instruction and found an average effect size of 0.32 in favor of computer-based instruction. Kulik claimed that students learned more and learned faster in courses that involved computer-based instruction.

2.2.2 Intelligent Tutoring Systems

The next generation of computer-based tutoring systems moved beyond the simple presentation of pages of text or graphics. These new intelligent tutoring systems (ITS), called cognitive tutors, incorporated model-tracing technology, which is a cognitive model of student problem solving that captures students' multiple strategies and common misconceptions. The cognitive tutor uses model-tracing to understand students' input and to indicate when they have made a mistake. With model-tracing, cognitive tutors provide students individualized assistance that is just-in-time and sensitive to the students' particular approach to a problem (Anderson, Corbett, Koedinger, & Pelletier, 1995). They also provide hint messages that get more explicit as students continue asking for help until the tutor is telling the student exactly what to do. The feedback is immediate and is structured so as to lead students step-by-step toward expert-like performance. The tutor intervenes as soon as students deviate from the solution path but does not engage students

in dialog by asking new questions. Cognitive tutors also use knowledge-tracing technology that traces students' knowledge growth across problem solving activities and uses this information to select problems and to adjust the pacing to adapt to individual student needs.

Although cognitive tutors do not engage students in dialog, they nonetheless have had a significant impact on student learning in a variety of domains. For example, Koedinger et al. (1997) compared a cognitive tutor, "Cognitive Tutor Algebra" to traditional algebra instruction. The Cognitive Tutor was built to support the Pittsburgh Urban Mathematics Project (PUMP) algebra curriculum that is centrally focused on mathematical analysis of real world situations and the use of computational tools. The study evaluated the effect of the PUMP curriculum and Cognitive Tutor use and found that students in the experimental classes outperformed control classes by 100% on assessments of the targeted problem solving and multiple representations. These results also translated into a one standard deviation effect size. Other studies comparing Cognitive Tutor and traditional algebra instruction, have found improvements in the 50-100% range thus replicating the previous results (Koedinger, Corbett, Ritter, & Shapiro, 2000). Approximately 375,000 students in over 1000 schools currently use this cognitive tutor.

Morgan and Ritter (2002) conducted a study comparing the Cognitive Tutor Algebra I course and a Traditional Algebra I Course, that used a different text, with students in their junior high school system. Dependent measures included the Education Testing Service (ETS) Algebra I end-of-course exam, course grades and a survey of attitudes towards mathematics. These measures certainly seem to have the benefit of not being defined by the experimenters themselves. When restricting the analysis to only

those teachers who taught both the Cognitive Tutor Algebra I course and the Traditional Algebra I Course, the researchers found statistically significant differences on all dependent measures in favor of the cognitive tutor. Morgan and Ritter state that the strongest components of teacher effects have to do with teacher education and professional development and only indirectly with their teaching practices. In their study the curriculum effect that they were examining had to do with teacher practices that would be expected to be relatively small. Therefore, they conclude that the effect size of 0.29 is impressive taken in this context.

Finally, VanLehn et al. (2005) evaluated Andes, an ITS developed to replace paper-and-pencil homework and to increase student learning in introductory college physics courses. Andes provides immediate feedback to student responses and also provides three kinds of help including: 1) pop up error messages when the error is probably due to lack of attention rather than lack of knowledge, 2) *What's Wrong Help* when the student is essentially asking what is wrong with that, and 3) *Next Step Help* if students are not sure what to do next. The *What's Wrong* and *Next Step Help* end with a bottom-out hint which tells students exactly what to do.

Andes was evaluated from 1999 to 2003 and in all years Andes students scored higher than control students with effect sizes ranging from 0.21 to 0.92. VanLehn et al. (2005) compared their results to the results of the Koedinger et al. (1997) study that they suggest is the benchmark study with respect to tutoring systems. The Koedinger et al. study evaluated the "Cognitive Tutor" intelligent tutoring system and a novel curriculum (PUMP), which Carnegie Learning distributes as the Algebra I Cognitive Tutor. Koedinger et al. used both experimenter-designed questions and standardized tests. While

analyzing the experimenter-designed tests, they found effect sizes of 1.2 and 0.7 and effect sizes of 0.3 while analyzing multiple-choice standardized tests. VanLehn et al. (2005) found very similar effect sizes (1.21 & 0.69) for their conceptual experimenter-written tests and similar effect sizes (0.29) for their multiple-choice standardized tests. Thus, both evaluations have similar tests and effect sizes. They both have impressive effect sizes for conceptual, experimenter-designed tests, and lower effect sizes on standardized, answer only tests.

The authors of the Andes study stated that their evaluation differed from the Koedinger et al. (1997) evaluation in a crucial way. The Andes evaluations manipulated only the way that students did their homework: on Andes vs. on paper. The evaluation of “Cognitive Tutor” was also an evaluation of the Pittsburgh Urban Mathematics Project curriculum (PUMP), which focused on analysis of real world situations and the use of computational tools such as spreadsheets and graphers. Therefore, how much gain was due to the tutoring system and how much was due to the new curriculum is not clear. Finally, VanLehn et al. (2005) stated that in their study, the curriculum was not reformed; therefore, the gains in their evaluation may be a better measure of the power of intelligent tutoring systems *per se*.

2.2.3 *Dialog-based Intelligent Tutors*

Both CAI and cognitive tutors have been shown to be more effective than traditional classroom instruction, yet neither has approached the effectiveness of human tutors. Perhaps they have not captured the features of human tutoring that account for its effectiveness. Researchers have recently developed ITS that incorporate dialog that is based on human tutors in specific domains. Preliminary results are promising.

The Tutoring Research Group at the University of Memphis has developed AutoTutor (Graesser, Person, & Harter, 2001), an ITS that helps students construct answers to computer literacy questions and qualitative physics problems by holding a conversation in natural language thus taking advantage of the interaction hypothesis (covered in Section 11). AutoTutor attempts to imitate a human tutor by reproducing the dialog patterns and strategies that were likely to be used by a human tutor. AutoTutor presents questions and problems from a curriculum script, attempts to comprehend learner contributions that are entered by keyboard, formulates dialog plans that are sensitive to the learner's contributions and delivers the dialog moves with an animated talking head that simulates facial expressions and speech to give the impression of a discussion between the tutor and student. AutoTutor has produced gains of 0.4 to 1.5 standard deviations depending on the learning performance measure, the comparison condition, the subject matter, and the version of AutoTutor (Graesser, Moreno, Marineau, Adcock, Olney, & Person, 2003).

Rosé et al. (2001) integrated Atlas and the Andes system to compare a model-tracing ITS with an ITS incorporating dialog. Atlas facilitates incorporating tutorial dialog while Andes is a model-tracing ITS for quantitative physics that provides immediate feedback by highlighting each step attempted in either red or green to indicate a right or wrong answer. Andes also provides hint sequences for students asking for help. The researchers were able to compare student learning between the original Andes and the integrated Atlas-Andes with dialog. Atlas-Andes students scored significantly higher on post-test measures with a difference of 0.9 standard deviations.

Heffernan (Heffernan & Koedinger, 2000) developed an ITS called Ms. Lindquist that uses dialog to help students write algebra expressions, a skill known as

symbolization. It models both student behavior and tutorial behavior by combining a cognitive model of student behavior with a tutorial model of tutoring strategies. The cognitive student model has a set of production rules that models symbolization skills and the tutorial model is based on the observation of an experienced human tutor. Heffernan (Heffernan & Croteau, 2004) found students using Ms. Lindquist completed only half as many problems as students in a control condition, but still showed learning gains over the control condition with an effect size of 0.5.

2.2.4 Interactive Tutoring

When tutoring a student, a tutor needs to decide whether to give a complete explanation on a topic or whether to draw out the explanation from the student through questions and interactive dialog. It is widely believed that interactive one-to-one dialog between students and tutors will result in more learning than reading text or listening to lectures. There are several arguments for stressing interactivity in tutoring. The tutor can identify misunderstandings or gaps in knowledge that the student may have and address them. Students must pay closer attention to the tutor when they are expected to participate and contribute to the tutoring session so they are less likely to daydream.

According to VanLehn et al. (2005), the *interaction hypothesis* is as follows “When one-on-one natural language tutoring, either by a human tutor or a computer tutor, is compared to a less interactive control condition that covers the same content, then the tutees will learn more than the nontutees.”

Several studies in the literature have found evidence in support of the interaction hypothesis with human tutors. In a comparison of Socratic and didactic tutoring strategies, Core, Moore and Zinn (2003) found that the more interactive (based on words

produced by students) Socratic tutorial dialogs had a greater correlation with learning. Also as expected according to the interaction hypothesis, Katz, Allbritton and Connelly (2003) found that students learned more when they participated in post practice dialogs with a tutor than students who did not. Chi, Siler, Jeong, Yamauchi and Hausmann (2001) found that students who engaged in a more interactive style of human tutoring were “able to transfer their knowledge better than the students in the didactic style of tutoring.” When Evens and Michaels (2006) compared expert human tutoring to reading a text book with the same material, they also found that the tutored students got significantly higher scores on a post-test.

Since the content covered in studies of these tutors can make a big difference to the results, studies of computer tutors have the advantage of being able to control the content better than studies with human tutors. Results that support the interaction hypothesis have also been found in studies of interactive ITS. Graesser, Moreno, Marineau, Adcock, Olney, & Person, (2003) did a comparison of natural language computer tutoring using AutoTutor and reading a textbook for a course on computer literacy. AutoTutor carried out dialog that was designed to imitate human tutor dialog with the result of the tutored students learning more than the control group.

It seems that a positive relationship between learning and tutor-student interaction exists, and we would expect students to learn more whenever they engage in interactive tutoring conditions than in less interactive conditions such as reading text. There is, however, evidence that this is not always the case. VanLehn, Graesser, Jackson, Jordan, Olney, & Rose (2005) reviewed several studies that hypothesize that the relationship between interactivity and learning exists, as well as a few studies that failed to find evidence for this relationship. VanLehn et al. (2005) found that when students found the

material to be difficult, tutoring was more effective than having the students read an explanation of how to solve a problem. However, this was not the case when the students found the material to be at their level: there was no significant difference between interactive tutoring and reading text. Craig, Driscoll and Gholson (2004) also reported mixed results in experiments comparing interactive tutoring with a computer to watching a video of a tutor.

In this dissertation, we look at tutoring strategies that differ in the level of interactivity between the tutor and the student. We consider tutoring to be high interactivity when the student is required to input an answer to each step of solving a problem in order to proceed to the next step. This is compared to tutoring that is low interactivity which we consider less interactive because they involve at most, clicking a button to request a hint and reading the hint message. Students do not have to respond to hints or input answers to each step. There is a spectrum of tutor-student interactivity involved in the strategies compared in this dissertation that involve the amount and frequency of student input, system-initiation vs. user initiation, and amount of information presented. Table 2 shows some high vs. low interactive strategies.

Table 2: Interactive vs. non-interactive instruction

High interactivity	Low interactivity
Tutored problem solving	Non-tutored problem solving
Self-explanations with feedback	Explanations/ Examples
Mixed user and system initiated	System-initiated

3. The ASSISTment System

The ASSISTment System is joint research conducted by Worcester Polytechnic Institute and Carnegie Mellon University and is funded by grants from the U.S. Department of Education, the National Science Foundation, and the Office of Naval Research. The ASSISTment System's goal is to provide assessment of students while providing tutoring assistance to students.

The ASSISTment System aims to assist students in learning the different skills needed for the MCAS test (or other state tests) while at the same time assessing student knowledge to provide teachers with fine-grained assessment of their students; it assists while it assesses. The system assists students in learning different skills through the use of *scaffolding questions*, *hint messages*, and messages that target incorrect answers known as *buggy messages* (Razzaq et al., 2007). Scaffolding questions, hint messages and buggy messages are described in the next section. Assessment of student performance is provided to teachers through real-time reports based on data from students using the system. Use of the web-based ASSISTment System is free and only requires registration on the website; no software need be installed. The system is primarily used by middle- and high-school teachers throughout Massachusetts who are preparing students for the MCAS tests. Currently, there are over 3,000 students and 50 teachers using the system as part of their regular math classes.

In addition to tutoring and assessing, the ASSISTment System attempts to support the full life cycle of content authoring with the tools available in the ASSISTment System. Teachers can create problems with tutoring, map each question to the skills

required to solve them (described in Section 3.2), bundle problems together in sequences that students work on, view reports on students' work and use tools to maintain and refine their content over time. Over 30 teachers use the system to create content.

3.1 Structure of an ASSISTment

Koedinger et al. (2004) introduced example-tracing tutors. Example-tracing tutors behave like cognitive tutors but are limited to the scope of a single problem. The ASSISTment System uses a further simplified example-tracing tutor, which is called an ASSISTment, where only a linear progression through a problem is supported, which makes content creation easier and more accessible to a general audience.

An ASSISTment consists of a single main problem, or what we call the “*original question*.” For any given problem, assistance to students is available either in the form of a hint sequence or scaffolding questions. Hints are messages that provide insights and suggestions for solving a specific problem, and each hint sequence ends with a *bottom-out* hint which gives the student the answer. Scaffolding questions are designed to address specific skills needed to solve the original question. Students must answer each scaffolding question in order to proceed to the next scaffolding question. When students finish all of the scaffolding questions, they may be presented with the original question again to finish the problem. Each scaffolding question also has a hint sequence to help the students answer the question if they need extra help. Additionally, messages called *buggy* messages are provided to students if certain anticipated incorrect answers are selected or entered. For problems without scaffolding, a student will remain in a problem until the problem is answered correctly and can ask for hints which are presented one at a time. If

scaffolding is available, the student will be programmatically advanced to the first scaffolding problems in the event of an incorrect answer on the original question.

Hints, scaffolds, and buggy messages together help create ASSISTments that are structurally simple but can address complex student behavior. The structure and the supporting interface used to build ASSISTments are simple enough so that users with little or no computer science and cognitive psychology background can use.

Figure 1 shows an ASSISTment being built on the left and what the student sees is shown on the right. Content authors can enter question text, hints and buggy messages by clicking on the appropriate field and typing; formatting tools are also provided for bolding, italicizing, etc. Images and animations can also be uploaded in any of these fields. In creating an environment that is easy for content creators to use, there is a tradeoff between ease of use and having a more flexible and complicated ASSISTment structure.

Figure 1 shows an example that involved understanding the skills of algebra, perimeter, and congruence. If the student had answered correctly, she would have moved on to a new item, but she incorrectly typed 5, to which the system responded, “Sorry, that is incorrect. Let’s move on and figure out why.” The system then engaged the student to give her some scaffolding questions that would help isolate for which of the skills she had an error, and to give her tutoring so that she could learn the correct actions. The tutor began by asking a “scaffolding” question that isolated the step involving *congruence*. Eventually she will get the scaffolding question correct (i.e., by answering AC), and then will be given a question that tries to determine if she understood *perimeter*.

The ASSISTment System assumes that students may know certain skills and rather than slowing them down by going through all of the scaffolding first, students are

allowed to try to answer questions without showing every step. This differs from Cognitive Tutors (Anderson, Corbett, Koedinger, & Pelletier, 1995) and Andes (VanLehn, et al., 2005) which both ask the students to fill in many different steps in a typical problem. The ASSISTment System's scaffolding pattern means that students get through items that they know faster and spend more time on items they need help on. It is not unusual for a single Cognitive Tutor Algebra word problem to take ten minutes to solve, while filling in a table of possibly dozens of sub-steps, including defining a variable, writing an equation, filling in known values, etc. To be sure, in circumstances where the student does not know these skills, this can be very useful. However, if the student already knows most of the steps this may not be pedagogically useful.

3.2 Skill mapping

The ASSISTment Builder also supports the mapping of skills, which are organized into sets (Razzaq, Heffernan, Feng, & Pardos, 2007). Skills are mapped to specific problems to indicate that a problem requires knowledge of that skill. The mapping between skills and problems allows the reporting system to track student knowledge over time using longitudinal data analysis techniques (Feng, Heffernan, & Koedinger, 2006).

In April of 2005, a subject-matter expert helped to identify skills and map all of the existing 8th grade MCAS items with these skills in a seven hour long "coding session". Content authors who are building 8th grade items can then tag their problems with one of the skills for 8th grade using the ASSISTment Builder. Tagging an item with a skill typically takes 2-3 minutes. The cost of building a skill model can be high initially, but the cost of tagging items is low.

There are more than twenty skill models available in the system with up to 300 skills in each. Content authors can map skills to problems and to scaffolding questions as they are building content. The Builder will automatically map problems to any skills that its scaffolding questions are marked with and vice versa.

3.3 Teacher Reports

The various reports in the ASSISTment System that are available on students' work are valuable tools for teachers. Teachers can see how their students are doing on individual problems or on complete assignments (an organized set of problems called a problem sequence). They can also see how their students are performing on each skill. These reports allow teachers to determine where students are having difficulties and they can adapt their instruction to the data found in the reports. For instance, Figure 2 shows an item report which shows teachers how students are doing on individual problems. Teachers can tell at a glance which students are asking for too many bottom-out hints (cells are colored in yellow). Teachers can also see what students have answered for each question, whether the answer was correct, what percent of the class got the answer correct and the percent correct for the whole problem set.

Module 4 Period 7

Legend: Hit bottom hint of the problem

Module G.4.8 Pythagorean Theorem (8 items)

Problem	Average	#131	#1648	#4665	#248	#1503	#74	#215	#1541	Total hints
	50%	25%	41%	66%	75%	48%	74%	55%	21%	
basan*	100%	+ 14	+ 60	+ C. 5	+ 120	+ a triangle with sides measuring 20, 21 and 29	+ 55	+ 60	+ 18 units	0
	62%	x 192	x 70	+ C. 5	+ 120	+ a triangle with sides measuring 20, 21 and 29	+ 55	+ 60	x 5 units	0
atooe*	16%	x 192	x 121 4 times	x Hint requested	+ 120 2 times	+ a triangle with sides measuring 20, 21 and 29	x 125 2 times	x 148 2 times	x 13 units	53

Figure 2: An item report tells teachers how students are doing on individual problems.

3.4 Problem Sequences

Problems can be arranged in problem sequences in the system. The sequence is composed of one or more sections, with each section containing problems or other sections. This recursive structure allows for a rich hierarchy of different types of sections and problems.

The section component, is an abstraction for a particular ordering of problems. Currently, section types include “Linear” (problems or sub-sections are presented in linear order), “Random” (problems or sub-sections are presented in a pseudo-random order), and “Choose Condition” (a single problem or sub-section is selected pseudo-randomly from a list, the others are ignored).

The ASSISTment System can be used to find the best ways to tutor students by running randomized controlled experiments. Figure 3 shows a problem sequence that has been designed to run an experiment that compares two conditions: giving students

scaffolding questions or allowing them to ask for hints. Three main sections are presented in linear order, a pre-test, experiment and post-test sections. Within the experiment section there are two sections (one for each condition) and students will randomly be presented with one of them.

Working with the ASSISTment system puts researchers in a unique position to be able to find out what works with intelligent tutoring systems, for whom, and under what circumstances. The ASSISTment project's collaboration with the Worcester Public Schools in Worcester, Massachusetts makes it possible to test hypotheses with real students in an urban school district. There are many ways to tutor children using educational technology and it is difficult to determine the best tutoring practice without randomized controlled experiments and rigorous analysis of data.

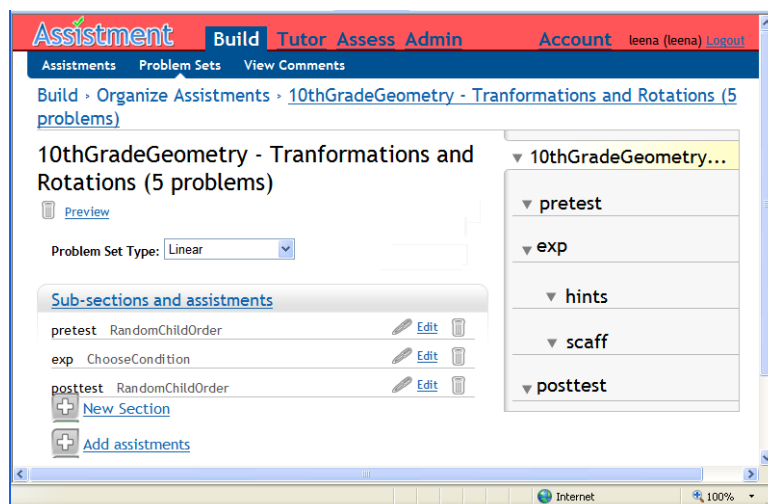


Figure 3: A Problem Sequence that is designed as a randomized controlled experiment

4. Is the ASSISTment System an effective tutor?

This section describes work to create and evaluate ASSISTments. Section 4.1 describes the process of creating ASSISTments. Section 4.2 presents evidence that students are learning in the ASSISTment system. Section 4.3 describes a study to determine whether students can learn more from using the ASSISTment System to do their homework than from doing traditional paper-and-pencil homework. Section 4.4 describes a study to compare traditional paper-and-pencil homework to using the ASSISTment System with immediate feedback but no tutoring or help.

4.1 ASSISTment development and usage

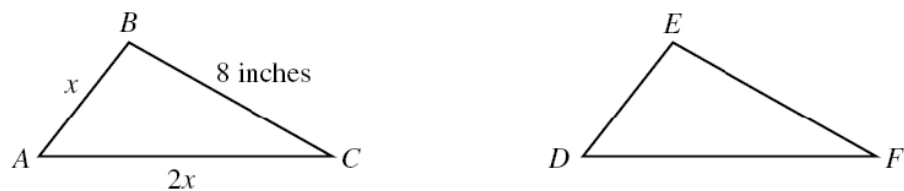
In December of 2003, Neil Heffernan and the author met with the Superintendent of the Worcester Public Schools in Massachusetts, and were subsequently introduced to the three math department heads of three out of the four Worcester middle schools. The goal was to get these educators involved in the design process of the ASSISTment System at an early stage. The main activity done with these teachers was meeting about one hour a week to do “knowledge elicitation” interviews, whereby the teachers helped design the pedagogical content of the ASSISTment System.

The procedure for knowledge elicitation interviews was as follows. A teacher was shown a MCAS test item and asked how she would tutor a student to solve the problem. What kinds of questions would she ask the student? What hints would she give? What kinds of errors did she expect and what would she say when a student made an expected error? These interviews were videotaped. The interviewer took the videotape and filled out an “ASSISTment design form” from the knowledge gleaned from the teacher. The

ASSISTment was then implemented using the design form. The first draft of the ASSISTment was shown to the teacher to get her opinion and she was asked to edit it. Review sessions with the teachers were also videotaped and the design form revised as needed. When the teacher was satisfied, the ASSISTment was released for use by students. For instance, a teacher was shown a MCAS item on which her students did poorly, such as item number 19 from the year 2003, which is shown in Figure 4. About 15 hours of knowledge elicitation interviews were used to help guide the design of approximately 20 ASSISTments. Figure 1 shows an ASSISTment that was built for item 19 of 2003.

As mentioned in Section 3, each ASSISTment typically consists of an original question and a list of scaffolding questions (in the case of 2003's item 19, 5 scaffolding questions). The first scaffolding question appears only if the student gets the item wrong. After an error, students are not allowed to try the item further, but instead must then answer a sequence of scaffolding questions (or "scaffolds") presented one at a time. Students work through the scaffolding questions, possibly with hints, until they eventually get the problem correct. If the student presses the hint button while on the first scaffold, the first hint is displayed, which would have been the definition of congruence in this example. If the student clicks the hint button again, the hint describes how to apply congruence to this problem. If the student asks for another hint, the answer is given. Once the student gets the first scaffolding question correct (by choosing AC), the second scaffolding question appears.

- 19 Triangles ABC and DEF shown below are congruent.



The perimeter of $\triangle ABC$ is 23 inches. What is the length of side \overline{DF} in $\triangle DEF$?

Figure 4: Item #19 from the 2003 Massachusetts state math test for 8th grade

Since 2003, the author has created close to 1,000 ASSISTments to provide tutoring for problems taken from state math tests for grades 4 through 10. These ASSISTments can include scaffolding questions, hints, buggy messages, images, animations and tutoring strategies. These ASSISTments are used regularly by students in their regular math classes as well as for homework. Over 3,000 students use the ASSISTment System in central Massachusetts and Pittsburgh, Pennsylvania.

The author has organized thousands of ASSISTments into problem sets based on the skills needed to solve the problems. Teachers can easily browse through the problem sets available and assign materials to students based on what they are doing in class or based on what they would like to assess. For instance, Figure 5 shows the problem sets that were created to target the skills for 8th grade math as outlined by the Massachusetts Frameworks. Similar problem sets were created for 7th grade and 10th grade math based on the Massachusetts Frameworks.

Defined Curriculums

Display curriculum for grade

2246 - Introduction To ASSISTment **Do this question first. It will walk you through how to use the system**

2 - Choose 20 Pretest **Here are 20 randomly-chosen MCAS items appropriate for 8th grade**

16 - 7th Grade Easy and Decomps **Here are some questions for students that need extra practice**

2630 - D.1.8 and D.4.8-Probabilities-Combinatorial (9 items)

2625 - D.2.8-Venn-Stem-n-Leaf-Circle-Graphs (9 items)

2626 - D.2.8-Reading Graphs(Bar and X-Y)(8 items)

2637 - D.2.8 N.1.8 and N.6.8-Ordering-Numbers-Numberlines (8 items)

2627 - D.3.8-Mean-Median-Mode (11 items)

2628 - D.4.8-Probability-Set One (9 items)

2629 - D.4.8-Probability-Set Two (9 items)

2631 - G.1.8 and G.2.8 Triangles-Sums-of-Interior-Angles (7 items)

2632 - G.3.8 and G.4.8 Supplementary Angles and Pythagorean Theorem (8 items)

2745 - G.6.8-G.7.8-G.8.8 Transformations-Rotations-3d-and-2d-Figures (13 items)

2634 - M.1.8 and M.2.8-converting-from-one-measure-to-another (12 items)

2635 - M.3.8-Area-Perimeter (12 items)

2636 - M.3.8-Area-and-Circumference-of-Circles (10 items)

2638 - N.1.8-Fractions-Decimals-Percents (6 items)

2639 - N.1.8 and N.10.8-Percents (14 items)

2645 - N.2.8 and N.5.8-Square-roots-and-Prime-Numbers (9 items)

2643 - N.3.8 M.4.8 and M.5.8-Ratios-Proportions-and-Rate-of-Change (13 items)

2644 - N.4.8 and N.7.8-Exponents-and-Scientific-Notation (7 items)

2646 - N.8.8-N.9.8-N.12.8-Order-of-Operations-and-Inverse-Operations (9 items)

5132 - N.10.8 Estimation - WE.2 (10 items)

2640 - N.10.8-Fraction-Multiplication (9 items)

2641 - N.10.8-Compute-with-and-without-Decimals (10 items)

2642 - N.11.8-Discounts-and-Compounding-Interest (11 items)

2647 - P.1.8-Pattern-Finding-Set-One (9 items)

2648 - P.1.8-Pattern-Finding-Set-Two (10 items)

2649 - P.1.8-Inducing-Functions (12 items)

2651 - P.2.8 and P.3.8-Substitution-and-Algebraic-Manipulation (13 items)

2652 - P.4.8-Symbolization-Articulation (11 items)

5198 - P.4.8-Writing symbolic expressions v.3 (16 items)

2653 - P.5.8 and P.7.8-Interpreting-Linear-Equations-and-Plotting-Points (11 items)

2654 - P.7.8-Setting-up-and-Solving-Equations-and-Inequalities (10 items)

2655 - P.7.8 P.8.8 and P.9.8 Solving-Equations-Modeling-and-Understanding-Covariation (10 items)

2650 - P.10.8-Understanding-Graph-Shapes-and-Linear-Growth (10 items)

Figure 5: 8th grade problem sets targeting skills set by the state of Massachusetts

The author has spent hundreds of hours observing the ASSISTment System use in classrooms. This time is used to work with teachers to try to improve content and to work with students to note any misunderstandings they sometimes bring to the content. For instance, if it is noted that several students are making similar errors that were not anticipated, the ASSISTment Builder can be logged into and a buggy message added that addresses the students' misconception.

Time in classrooms is also spent helping teachers to use their students' data to inform their instruction. For instance, teachers can determine which skills their students have mastered and which skills need to be reviewed in class from reports in the ASSISTment System. Teachers can also determine which misconceptions students have and which students are not taking their work seriously. For instance, while working with a teacher at Burncoat Middle School, the author was able to show, to the teacher's surprise, that the majority of her students did not understand Venn Diagrams. She decided to spend more time on Venn Diagrams in class since this was a topic that regularly appeared on the state test every year.

4.2 Are students learning from the ASSISTment system?

Razzaq et al. (2007) reported evidence that students were learning from using the ASSISTment system in their regular math classes. Students potentially saw 33 different problem pairs in random order. Each pair of ASSISTments included one based on an original MCAS item and a second "morph" intended to have different surface features, such as different numbers, but the same deep features or knowledge requirements, such as approximating square roots or finding the area of a figure given the perimeter (see Figure 6). Learning was assessed by comparing students' performance the first time they

were given one of a pair with their performance when they were given the second of a pair. If students tend to perform better on the second of the pair, it indicates that they may have learned from the instructional assistance provided by the first of the pair.

Both a student-level analysis and an item-level analysis were done. The hypothesis was that when students worked on pairs of items that required similar skills and got the first item incorrect, they should show learning by getting the second item correct. The pairs of items that were chosen for this analysis had been completed by at least 20 students.

For the student-level analysis there were 742 students who contributed to the analysis by comparing how they did on the first opportunity versus the second opportunity on a similar skill. A gain score per item was calculated for each student by subtracting the student's score (0 if they got the item wrong on their first attempt, and 1 if they got it correct) on their first opportunity from their score on the second opportunity. Then an average gain score for all of the sets of similar skills in which they participated was calculated. A student analysis was done on learning opportunity pairs seen on the same day by a student, and the t-test showed statistically significant learning ($p = 0.0244$). It should be noted that there may be a selection effect in this experiment in that better students are more likely to do more problems in a day and therefore are more likely to contribute to this analysis.

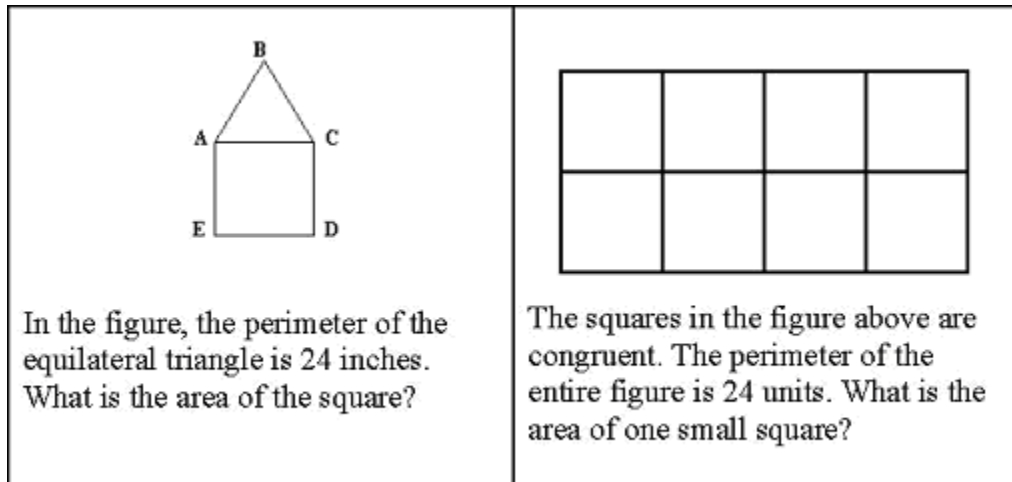


Figure 6: A perimeter and area learning opportunity pair

An item analysis was also done. There were 33 different sets of skills that met the criteria for this analysis. The five sets of skills that involved the most students were: Approximating Square Roots (6.8% gain), Pythagorean Theorem (3.03% gain), Supplementary Angles and Traversals of Parallel Lines (1.5% gain), Perimeter and Area (4.3% gain) and Probability (3.5% gain). A t-test was done and showed that the average gain scores per item were significantly different than zero ($p = 0.04$).

4.3 ASSISTments with tutoring/help vs. paper-and-pencil homework

Web-based systems that allow students to do their homework online such as Blackboard (www.blackboard.com), WebCT (www.webct.com), Homework Service (hw.utexas.edu/bur/overview.html) and WeBWorK (webwork.rochester.edu) are becoming more widely used. At the K-12 level, systems such as Study Island (studyisland.com) and PowerSchool (powerschool.com) are gaining popularity among teachers. Some states, such as Maine, Indiana, Michigan and Virginia, have begun to implement one-to-one computing programs, where each child gets his/her own laptop to

use during school and occasionally to take home as well (Bonifaz & Zucker, 2004). In fact, the Maine Learning Technology Initiative (2002-2004) supplied every Maine seventh and eighth grade student and their teachers with laptop computers, with 40% of the middle schools allowing students to take their laptops home. There are few research studies on the effects of one-to-one computing on teaching and learning; however, teachers report that students in one-to-one computing programs were more engaged, motivated and interact better with teachers (Bebell, 2005; Silvernail & Lane, 2004). As the digital divide narrows and more states become committed to one-to-one computing, the opportunities for students to do their homework online increase. The important question is, do students learn more by using computers to do their homework than by doing traditional paper-and-pencil homework?

Advantages of homework assistance systems are immediate feedback to students and automatic grading for instructors. Automatic grading is helpful by a) saving time for teachers who do not have time to grade all of their students' paper-and-pencil homework carefully by hand and b) by prompting students to take homework more seriously because they know it will be graded and the grade recorded. With these systems, students can get immediate feedback on their answers to problems and sometimes hints.

Although there are benefits to using these web-based homework assistance systems, there can be disadvantages, too. One disadvantage is that many of these systems require students to enter a single answer for each problem and do not take how the student arrived at the answer into consideration. Students may do less work on paper which can help them to be more organized and try to do more math in their heads instead. Teachers may spend less time looking at student work and figuring out exactly where

they are having difficulties. Because these systems often do not consider student work, cheating may be easier among students.

Research has shown positive results for using intelligent tutors instead of traditional paper and pencil homework. “MasteringPhysics” is a web-based physics homework tutor developed at MIT that uses mastery learning to help students reach mastery when solving physics homework problems. Students can request hints on problems and can receive feedback on common student errors. Some hints will ask a question which behaves similar to “scaffolding questions” in the ASSISTment system. Warnakulasooriya & Pritchard (2005) found that twice as many students were able to complete a set of problems in a given time with the help provided by MasteringPhysics when compared to students who worked on the problems without help (administered by MasteringPhysics but without hints or feedback).

Quantum Tutors (www.quantumsimulations.com/) is a web-based system that is commercially available for students to do homework in the sciences and math. Students can choose topics to work on, enter their own problems and choose from a list of questions they may have on particular problems. For instance, students working on percents can choose “I need to find the percentage one number is of another” and solve problems provided by the system or enter their own values to solve. They can also choose from a list of questions such as “Why would I want to convert a percent to a fraction?” In a press release, Quantum Tutors describes a week-long study done in 2005 (www.quantumsimulations.com/news15.html), where students using Quantum Tutors for homework in a high school chemistry course outperformed a control group that did paper-and-pencil homework on a post-test by just over a full letter grade. The difference between groups became larger as the problems increased in difficulty.

The Andes system is an intelligent tutoring system that can be used on the internet to provide support for problem-solving for physics homework. Andes requires students to complete whole derivations step-by-step and offers feedback after each step. Students can also ask for hints on each step to find out the nature of their error (What’s Wrong Help) or to get help on what the next step is (Next Step Help). Andes was used and evaluated in introductory physics classes from 1999 – 2003 at the U.S. Naval Academy. VanLehn et al. (2005) presented evidence that students who used Andes for homework got significantly higher exam scores than students in control groups who did pencil and paper homework. The Andes studies seem to be the most closely related to the following comparison of ASSISTments and paper-and-pencil homework (Mendicino, Razzaq, & Heffernan, 2009) which attempts to replicate Andes’ positive results.

4.3.1 Participants

The setting for this study was four fifth grade classrooms and students’ home computers. The school was located in a small town in a rural county. Approximately 350 students were enrolled in the school at the time of the study with at least 50% receiving free or reduced lunch. All four classes were typical elementary classes with a mix of below average, average, and above average students. There were a total of 92 students, 54 with internet access at home. The breakdown of the participants is shown in Table 3.

Table 3: Participants

	Class A	Class B	Class C	Class D	Total
Male students	13	12	9	12	46
Female students	10	13	13	10	46
Students with internet at home	12	15	17	10	54

4.3.2 Content

Two problem sets were used in both the web-based homework and the paper-and-pencil homework assignments each consisting of ten problems. One problem set consisted of number sense problems and the other was a mix of problems. The Number Sense problem set included problems for which students had to demonstrate understanding of numbers, ways of representing numbers, and relationships among numbers and number systems. The Mixed problem set included problems in the algebra, geometry, data analysis and probability domains. Students had to demonstrate understanding of patterns, relations, and functions, describe spatial relationships using coordinate geometry and other representational systems, develop and evaluate inferences and predictions that are based on models, and apply and demonstrate an understanding of basic concepts of probability (see Appendix A for the problems in both homework sets). Students in the classes had prior learning experience with the homework material during the course of the school year. However, since the experiment took place at the end of the school year, it was not recent experience and more of a review.

The worksheets that students completed for paper-and-pencil homework assignments were identical to the web-based homework assignment problems, with the same formats (i.e., multiple choice or short answer). This was possible because each class did the Number Sense or Mixed problem set for computer homework and the opposite problem set for paper-and-pencil homework. This will be explained in detail in the Experimental Design section below. The pre-test and post-test items were “morphs” of the Number Sense and Mixed problem sets and were designed to have different surface features such as names and numbers while keeping the same deep features or knowledge requirements such as analyzing patterns. Finally, the same hints used in the problem sets

on the web-based tutor were used while going over paper-and-pencil homework problems in class to ensure that each class had the same instruction.

4.3.3 Experiment Design

A counterbalanced within-subjects experimental design was used in this study in which students in two classrooms participated in the web-based homework condition first while students in the other two classrooms participated in the paper-and-pencil condition first. All students participated in both web-based and paper-and-pencil conditions and all students were given pre-tests and post-tests for each condition in which they participated.

In the web-based first group, one class was given a pre-test for the Number Sense problem set and the other class was given a pre-test for the Mixed problem set. Students were then given a homework assignment consisting of ten problems in their respective problem sets on the web-based system. After completing the web-based homework the students were given post-tests in class the next day. The groups were then reversed, with the Number Sense group given the Mixed pre-test and the Mixed group given the Number Sense pre-test. Both groups were then given a paper-and-pencil homework assignment that consisted of ten problems on a worksheet for their respective problem sets. They were given post-tests on the following day. The paper-and-pencil first group participated in the same overall experimental design (see Table 4 for the overall experimental design).

The design counterbalances the content (Number Sense versus Mixed) as well as counterbalances the order of condition (web-based versus paper and pencil) so that we can draw valid inferences that any gains students make will not be attributed to these outside factors.

4.3.4 Procedure

Two of the four classes were assigned to the computer-first group and two were assigned to the paper-and-pencil-first group. Within each group one class was assigned the number sense set and the other was assigned the mixed set. On day one of the experiment, students in all four classes were given the appropriate pre-test (two classes were administered the number sense pre-test, and two the mixed pre-test). So for both conditions, computer-based homework and paper-and-pencil homework, one class was completing the number sense assignment while the other class completed the mixed assignment. Students in the computer homework condition were randomly assigned by the computer either to a tutored problem solving condition or hints on demand condition.

After completing the pre-test, each class in the paper-and-pencil homework condition was given a worksheet containing ten problems to complete for homework. They were instructed to bring the worksheet to school the following day so the teacher could go over it and answer any questions they had. Students in each class in the computer homework condition completed their pre-tests then were taken to the media room where they were instructed in how to log into the ASSISTment System. They all were given a school identifier which was the same for all students, then given an individual screen name that consisted of their first name and last initial. After each student logged into the system they were shown how to select their teachers' names and how to select their homework for the evening. They were also instructed on how to select and work on a demonstration problem to familiarize themselves with the system and how problems were presented. The demonstration problems were not a part of their homework problems.

On day two, the students in the paper-and-pencil condition were given the answers to their worksheet problems and then given the opportunity to ask questions for review. When answering questions the teacher used the exact hints used in the computer hints condition to ensure uniformity across groups. The review was limited to ten minutes per group. Post-tests were then administered to students. The two classes in the computer condition were administered post-tests on day two. Days three and four followed the same procedures as days one and two, but the groups were reversed. That is, the two classes in the computer condition switched with the two classes in the paper-and-pencil condition and the two classes that did number sense problems switched with the two classes doing mix problems.

Table 4: Overall Experimental Design

Day	Computer First Group		Paper-and-Pencil First Group	
	Class A	Class B	Class C	Class D
Mon	Pretest Number Sense Intro to ASSISTments Web-based assignment	Pretest Mixed Intro to ASSISTments Web-based assignment	Pretest Mixed Paper-and-pencil assignment	Pretest Number Sense Paper-and-pencil assignment
Tues	Post-test Number Sense	Post-test Mixed	Review assignment Post-test Mixed	Review assignment Post-test Number Sense
Wed	Pretest Mixed Paper-and-pencil assignment	Pretest Number Sense Paper-and-pencil assignment	Pretest Number Sense Intro to ASSISTments Web-based assignment	Pretest Mixed Intro to ASSISTments Web-based assignment
Thurs	Assignment review Post-test Mixed	Assignment review Post-test Number Sense	Post-test Number Sense	Post-test Mixed

4.3.5 Results

Fifty-four students had internet available at home and could participate fully in the study. Of the 54 students, 28 students completed the web-based homework and the paper-and-pencil homework.

For the following analyses, t-tests were run on the computer gain scores from pre-test to post-test and on the paper-and-pencil gain scores from pre-test to post-test. There was learning in both conditions; however, when comparing the main effect of computer homework as a whole (tutored problem solving & hints) and paper-and-pencil homework (including only those students who completed both homework assignments), there was a statistically significant difference in favor of the computer homework condition. The results showed a p-value of 0.05 with an effect size of 0.61. The 95% confidence interval for this effect size of 0.61 is [0.15 – 1.21].

An Analysis of Variance (ANOVA) was run to see if there was differential learning between the two computer conditions, scaffolding and hints. The mean gain for students in the hints condition was 2.47 problems, while for students in the scaffolding condition it was 2.17 problems. The results were not statistically significant, p-value of .716. We observed that students who had scaffolding did tend to spend more time on their homework. The average time spent on the computer doing homework was twenty-four minutes and fifteen seconds. Students in the scaffolding condition averaged almost thirty minutes (29.54) and the students in the hints condition spent a little more than eighteen minutes (18.17).

4.3.6 Discussion

In this experiment, students had significantly higher learning gains when doing homework on the computer as compared to when they did homework in a traditional paper-and-pencil manner. A substantial effect size (0.61) was found in favor of students doing homework via computer for students that completed both computer-based homework and pencil-and-paper homework.

As more and more web-based homework assistance systems become available, teachers can take advantage of the convenience of having homework automatically graded and recorded and students can benefit more if they take homework more seriously when they know it will be recorded.

By using a system such as the ASSISTment system, students can learn more than they would by doing their homework with paper and pencil. Students get immediate feedback on their answers and help when they need it. One limitation of the ASSISTment system is that it is not able to grade open responses or essay-type questions and teachers are limited to multiple choice or short answer questions.

In this study, we were limited by the number of students who had internet at home. As the digital divide narrows and more K-12 students have access to computers and internet at home, more teachers can take advantage of the promise of web-based homework assistance systems. This study was also limited in that assignment of conditions was by class rather than by student.

4.4 ASSISTments with no tutoring vs. paper and pencil homework

The previous experiment compared ASSISTments with tutored problem solving and/or hints to paper and pencil problem solving. The advantage of the ASSISTment System

over paper and pencil homework can be attributed to the tutoring/hints that students saw, or it could be attributed to the immediate feedback on their answers. With paper and pencil homework, the students waited until the next day to find out whether their answers were right or wrong, while students using ASSISTments could find out immediately. This section describes a study that compares paper and pencil homework to ASSISTment homework with immediate feedback on answers but no tutoring or hints.

4.4.1 Participants and Content

Seventy-three high school students taking Algebra II participated in this study. Two homework worksheets were given: Quadratic equations and Exponents. Students were asked to use the ASSISTment System to enter their answers for one of the assignments and to do the other assignment with paper and pencil. When doing the assignment using the ASSISTment System, the teacher decided that students were allowed two attempts to enter the correct answer. If after two attempts, the student failed to enter the correct answer, the answer was given (see Figure 7).

4.4.2 Experiment Design

All of the students participated in both conditions so a repeated measures analysis was used. A counterbalanced design controlled for the order of condition and assignment was by student. Pretests and post-tests were given before and after each assignment and a gain score was calculated for each assignment by subtracting the pretest score from the post-test score. See Table 5 for the experiment design.

Table 5: Experiment Design

Day 1	Quadratic Equations Pretest
	Randomly assigned students to ASSISTment homework or paper and pencil homework on Quadratic Equations
Day 2	Quadratic Equations Post-test
Day 3	Exponents Pretest
	Assigned students to ASSISTment homework or paper and pencil homework on Exponents based on the condition they were assigned to on Day 1.
Day 4	Exponents Post-test

4.4.3 Results

Because of technical difficulties on the first homework day, only 29 students completed both homework assignments. The mean gain for the paper and pencil homework was 3.6 points (out of 10 possible) and the mean gain for the ASSISTment homework was 2.7 points out of 10. The difference between the gain scores was not significant ($p = 0.3$).

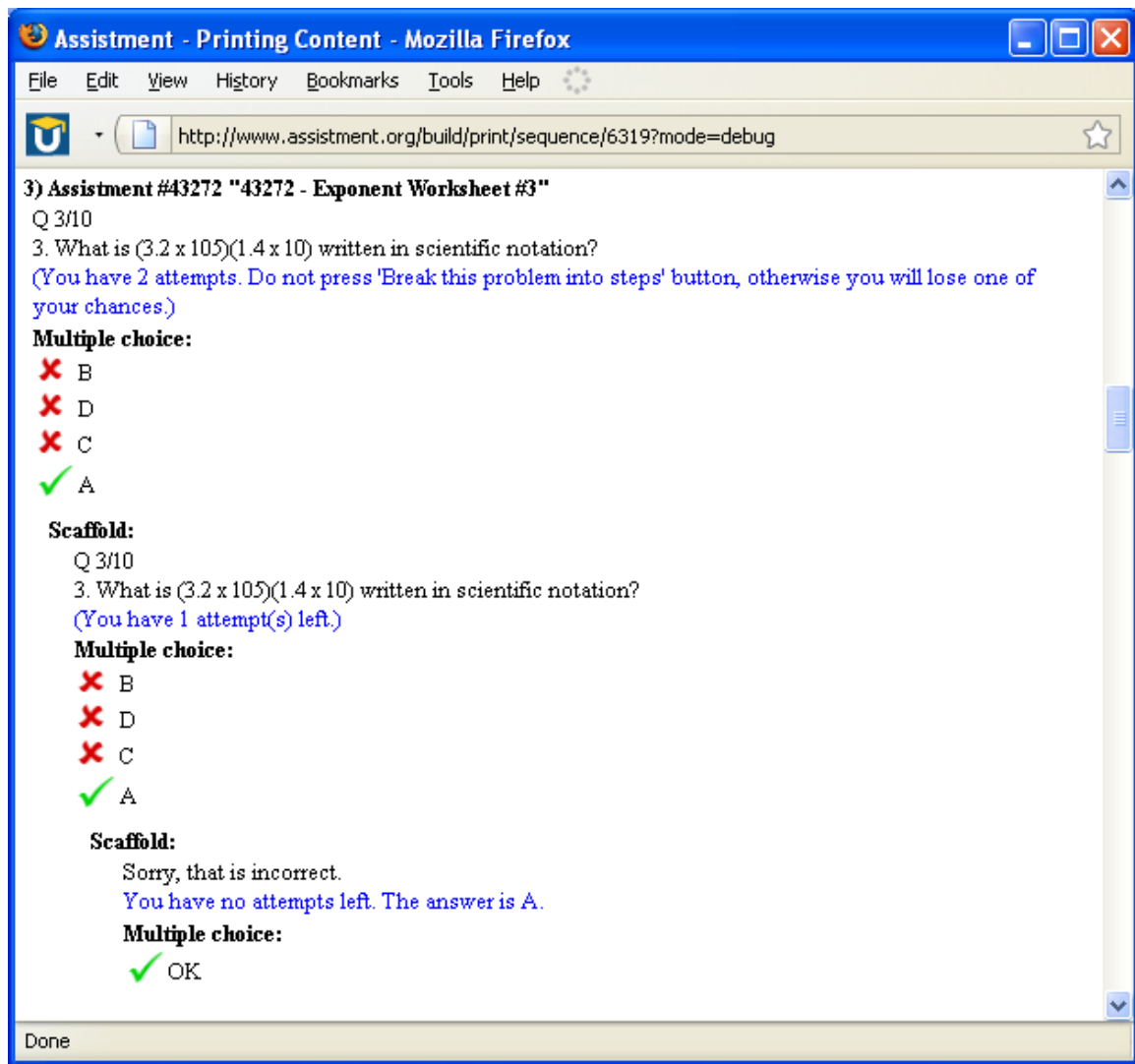


Figure 7: Students were given two attempts to enter the correct answer.

4.4.4 Conclusion

The results of this study may be an indication that the advantage of using the ASSISTment System for homework over paper and pencil homework may come, not from the immediate feedback on student answers, but rather from the tutored problem solving and hints that helps students to reach the answer. However, we are limited by the fact that the students in the two studies in this section are not the same students. We

would be able to draw more convincing conclusions by comparing these methods with the same students which is recommended as future work.

5. Tutored Problem Solving vs. Hints on Demand

Early evidence that the ASSISTment system was causing students to learn was reported by Razzaq et al. (2007), but we were uncertain if that was simply due to students getting more practice on math problems or more due to the "interactive tutored problem solving" that the system forced students to participate in if they got a problem wrong. A survey indicated that some students were frustrated by being forced to do the tutored problem solving. In addition, a lot of time was invested into authoring scaffolding questions used for tutored problem solving. An experiment described by Razzaq & Heffernan (2006) was conducted to see if students learned on a set of four problems if they were forced to do the scaffolding questions, which would *ASK* them to complete each step required to solve a problem, compared with being given hints on demand, which would *TELL* them the same information without expecting an answer to each step. In this study, the "tutored problem solving" condition represents a more interactive learning experience than the "hints on demand" condition.

A similar experiment carried out in 2004 tested to see whether students learned more from tutored problem solving or hints on demand in the ASSISTment system. In that experiment, 11 problems on probability were presented to 8th grade students in Worcester, Massachusetts. We will refer to this as the Probability Experiment. Some students received the tutored problem solving condition while others received the hint condition. In the tutored problem solving condition, the computer broke each problem down into 2-4 steps (or scaffolds) if a student got the original question wrong. In the hints condition, students could ask for hints if they needed help. The number of problems was

controlled for. When students completed all 11 problems, they worked on a post-test to test if they had learned how to solve the problems.

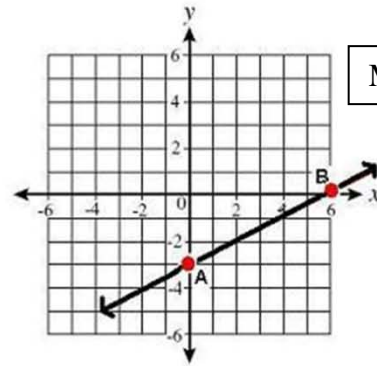
The results of the statistical analysis showed a large gain for those students that did the tutored problem solving, but it was discovered that there was a selection-bias. There were about 20% fewer students in the tutored problem solving condition that finished the experiment, and those students that finished were probably the better students. One reason for this bias could be due to the fact that students in the hint condition could finish problems faster than students in the scaffold condition.

A new experiment was designed (Razzaq & Heffernan, 2006) that focused on problems that involved slope and intercept, which according to data from within the ASSISTment system, students found difficult. We will refer to this experiment as the Slope Experiment. Four problems were chosen for the experiment and four more were chosen for the post-test to determine whether the students had learned how to do slope problems. Two of the post-test problems were also presented at the beginning of the experiment to serve as pretest problems. Students who got both pretest items correct did not participate in the experiment as they probably had already mastered the material. Students who got a pre-test item wrong were not told the answer nor given any tutoring on the item. They were shown a message that told them that they would come back to this problem at the end of class.

To make sure that the students had the opportunity to complete the post-test, they were timed. The students were given 20 minutes to work on an assignment containing the two pretest problems and four experiment problems. They were then asked to complete another assignment containing the four post-test problems and could work on them until the end of class. Unlike the Probability Experiment, students had to complete all of the

problems before proceeding to any other assignment. This procedure also ensured that students would work on the post-test regardless of which condition they were in.

Figure 8 shows a slope problem used in the experiment. The problem on the left, in the tutored problem solving condition, shows that a student has answered incorrectly and is immediately presented with a scaffolding question. The problem on the right, in the hints condition, shows that a student got the problem wrong and received the message, outlined in red, of “That is incorrect”. The hint shown outlined in green appears when the student requests a hint by clicking on the Hint button. The content of the hints in the hints condition were similar to the scaffolding questions so that the content was kept constant. The difference is that the students in the tutored problem solving condition were forced to give answers to the individual steps in the problem. The hypothesis was that if there was a difference between tutored problem solving and hints in this experiment it would be due to forcing students to work actively to solve each step of a problem, i.e., learning by doing.



Main question

Tutored Problem Solving

Hints on demand

What is the slope of the line graphed above?

Hmm, no.
Let me break this down for you.

Slope is a number that measures the steepness of a straight line. We measure slope by picking 2 points and dividing the change along the y-axis by the change along the x-axis.

What is the change along the y-axis for this line from point A to point B?

What is the change in x for this line from point A to point B?

Good. Remember, the slope is the change in y divided by the change in x. What is the slope for this line?

The change in y from point A to point B is 3.
The change in x from point A to point B is 6.
The slope can be found by dividing the change in y by the change in x.
The slope is 3/6.

Sorry, that is incorrect.

Slope is defined as the change in y divided by the change in x from point A to point B. What is the change in y from point A to point B?

The change in y is 3.
What is the change in x from point A to point B?

The change in x is 6.
What is the slope?

The slope is 3/6

Figure 8: The tutored problem solving condition with three scaffolding questions is shown on the left. Four hints are shown on the right. Students ask for each hint by clicking on the hint button.

5.1 Participants

There were a total of 178 students from three middle schools in Worcester who participated in the Slope Experiment. Of the participants, 25 students were excluded for getting both pretest items correct, 11 were in the tutored problem solving condition and 14 were in the hints condition. Another 5 students were excluded because they did not complete the post-test: 2 in the tutored problem solving condition and 3 in the hints condition. After these exclusions, there were 77 students in the tutored problem solving condition and 71 students in the hints condition.

5.2 Results

An ANOVA was run to test whether the two conditions differed by pretest. The result was not statistically significant so it was concluded that the groups were fairly balanced in incoming knowledge. However, of the two pretest problems given, one of them was much harder than the other; 18% of the students got the first pretest problem correct as opposed to 45% who got the second pretest problem correct. The first pretest problem concerned finding the y-intercept from an equation (What is the y-intercept in this equation: $y = 3/4x - 2$?). The second pretest problem presented the student with 3 points and asked them to choose the graph that contained the points.

Two different analyses of the data were done. The first method, Analysis #1, analyzes the post-test scores, while the second method, Analysis #2, only uses a single problem, but has the advantage of being able to use performance on the pretest by calculating a gain score. Analysis #1 compared the two groups' average post-test scores

but ignored pretest scores, while Analysis #2 looked at differing performance on the harder of the two pretest problems that was repeated in the post-test.

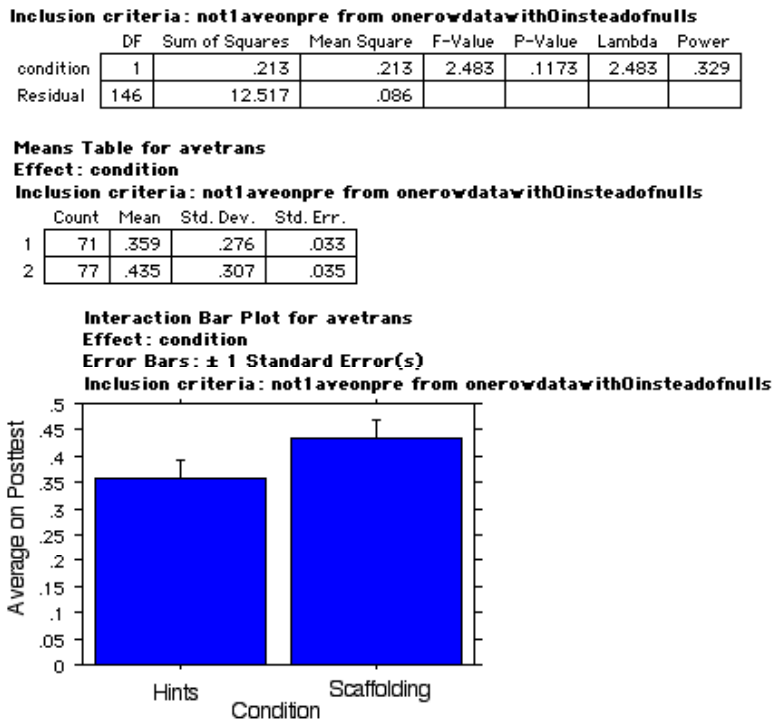


Figure 9: Results for average on post-test items by condition.

For Analysis #1, the results showed that students learned more from tutored problem solving with a p-value of 0.117 with an effect size of 0.3 (see Figure 9). We also calculated the 95% confidence interval for this effect size of 0.3 and got [-0.03, 0.6]. Because zero is included in this interval, we do not have 95% confidence that the effect size is real. We wanted to get a sense of the significance of this effect size so we calculated the 90% confidence interval and found the range to be [0.01, 0.56]. This implied that the effect size was greater than 0.01 with 90% confidence. We interpret this

as somewhat weak evidence in support of the hypothesis that students learned more in the scaffolding condition.

We also looked at scores on the transfer items that students had seen as pretest items. For the first pre-test item, which concerned finding the y-intercept from an equation, the ANOVA showed a statistically significant p-value of 0.005 with an effect size of 0.85 (see Figure 10). The 95% confidence interval of the effect size of 0.85 is [0.5, 1.2], meaning that we are 95% confident that the effect size is somewhere between 0.5 and 1.2, implying that the effect size seems to be at least greater than 0.5, which is a very respectable effect size.

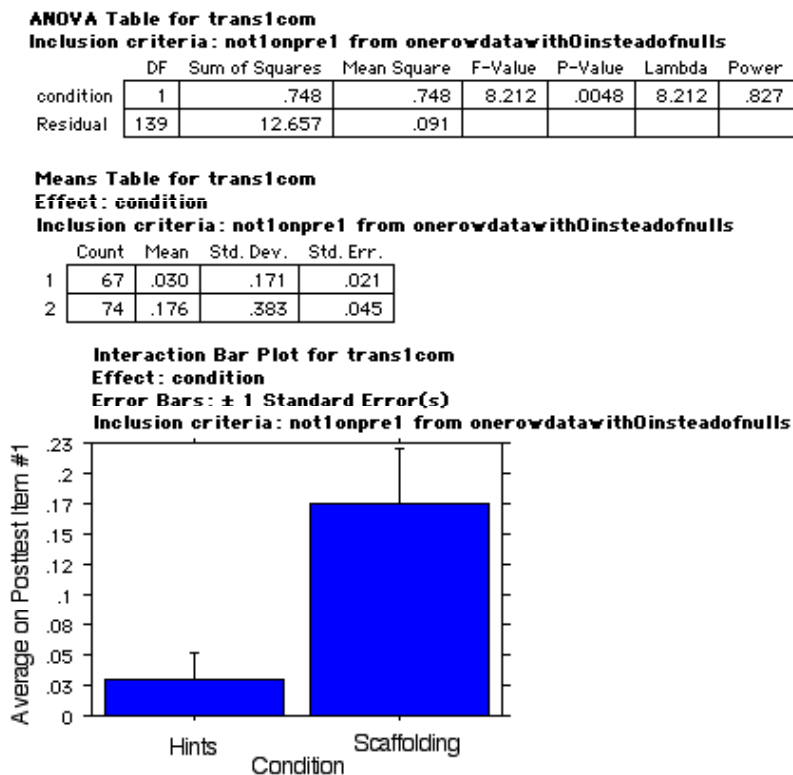


Figure 10: Results on the transfer item for the first pre-test item by condition

5.3 Hints on demand vs. proactive hints

Perhaps the reason that hints on demand was not as effective as tutored problem solving was the fact that hints on demand depended on student initiative: students were expected to ask for a hint when they wanted one whereas the tutored problem solving condition guided students through each step and provided the same information in the hints proactively. Murray and VanLehn (2006) found that proactive help was more effective for some students. “Proactive help when a student would otherwise flounder can save time, prevent confusion, provide valuable information at a time when the student is prepared and motivated to learn it, and avoid the negative affective consequences of frustration and failure.” In the ASSISTment system, students only see hints if they ask for them and they are less likely to ask for hints on multiple choice questions when they can guess more easily.

There are advantages for allowing the student to have control over when to ask for help (Anderson, 1993) because of the difficulty of deciding what kind of error the student has made. For instance, a tutor would respond differently to an error caused by a slip (the student knows the skill but slipped up) or by a misconception or by missing background knowledge. Students may also be able to figure out the correct answer on their own.

To test this hypothesis, a small study was conducted to compare proactive hints to hints on demand in the ASSISTment System. This study was concerned with “context-sensitive” hints or hints that are pertinent to the task at hand and help the student to learn a skill by doing.

There were two conditions in this study, hints on demand (see Figure 11) and proactive hints (see Figure 12), and two topics (symbolization and slope/intercept). Students participated in both conditions in a counterbalanced design that controlled for the order of conditions, the order of topics and the order of problems, which were randomly assigned by student (see Table 6).

A plumber not only charges \$30 per hour, but he also charges \$10 just to come to your house. Write an expression for how much you have to pay if this plumber works for h hours.

[Comment on this question](#)

Let's try substituting a number for h and computing the answer first.
If the plumber worked for 3 hours at your house how much would you owe him?

Well, the **amount owed to the plumber** is equal to the **amount charged per hour** times the **number of hours worked** plus the **amount charged for the house call**.

The amount charged per hour is **30 dollars** and the number of hours worked is **3**.

The amount charged for the house call is **10 dollars**.

So, the plumber will charge **90** dollars for the work done plus **10** dollars for the house call, or $90 + 10 = 100$ dollars

[Comment on this hint](#)

Now, if the plumber worked for 3 hours the expression that correctly finds the **amount owed to the plumber** is

the **number of hours** times the **amount charged per hour** plus the **amount charged for the house call**.

or $30 \cdot 3 + 10$

[Comment on this hint](#)

Now, remember we don't actually know how many hours the plumber will work, so we need to use h instead of 3.

We need to add the **amount charged for the work** to the **amount charged for the house call**.

The amount charged for the work is **$30h$**

The amount charged for the house call is **10**

The amount owed is **$30h + 10$**

[Comment on this hint](#)

Type your answer below (mathematical expression):

Submit Answer

Figure 11: Hints on demand, students ask for each hint by clicking on a hint button

A plumber not only charges \$30 per hour, but he also charges \$10 just to come to your house. Write an expression for how much you have to pay if this plumber works for h hours.

[Comment on this question](#)

Break this problem into steps

Type your answer below (mathematical expression):

Submit Answer

Let's move on and figure out this problem.

Here is a hint:

Let's try substituting a number for h and computing the answer first. If the plumber worked for 3 hours at your house how much would you owe him?

Well, the amount owed to the plumber is equal to the amount charged per hour times the number of hours worked plus the amount charged for the house call. The amount charged per hour is 30 dollars and the number of hours worked is 3. The amount charged for the house call is 10 dollars.

So, the plumber will charge 90 dollars for the work done plus 10 dollars for the house call, or $90 + 10 = 100$ dollars

Try writing an expression for how much you owe the plumber if he worked for h hours.

[Comment on this question](#)

Break this problem into steps

Type your answer below (mathematical expression):

Submit Answer

✘ Sorry, that is incorrect. Let's move on and figure out why!

Here is another hint:

Now, if the plumber worked for 3 hours the expression that correctly finds the amount owed to the plumber is the number of hours times the amount charged per hour plus the amount charged for the house call. or $30 \cdot 3 + 10$

What is the expression if the plumber worked for h hours?

[Comment on this question](#)

Break this problem into steps

Type your answer below (mathematical expression):

Submit Answer

Figure 12: Proactive hints: hints are presented automatically when a student submits an incorrect answer

Sixty-one middle school students participated in the study. During one class period, students solved problems in the two topics. A pretest and post-test were given before and after each topic.

Table 6: Students were randomly assigned to one of four conditions.

	Condition 1	Condition 2	Condition 3	Condition 4
First	Symbolization Hints on Demand	Symbolization Proactive Hints	Slope/Intercept Hints on Demand	Slope/Intercept Proactive Hints
Second	Slope/Intercept Proactive Hints	Slope/Intercept Hints on Demand	Symbolization Proactive Hints	Symbolization Hints on Demand

The results showed that students learned significantly ($p = 0.035$) more from having control over when to ask for a hint (mean gain score = 0.13) compared to having the computer control when to give a hint (mean gain score = 0.04).

5.4 Conclusion

This chapter described two experiments. The results of the first experiment showed that there is more learning with tutored problem solving than with hints on demand, although the difference was not always significant between the two conditions. The first pretest problem on finding the y-intercept from an equation proved to be a difficult problem for all of the students and tutored problem solving helped significantly. Perhaps the tutored problem solving had a greater positive effect on learning for the first pretest item because it was much more difficult for the students than the second pretest item. This result leads

into the search for an aptitude treatment interaction concerning the effectiveness of tutored problem solving.

The second experiment in this chapter showed that it is possible that the reason students did not learn as much from the hints on demand condition may not be because of a problem with student initiative, but perhaps because tutored problem solving required students to solve each step correctly before proceeding, while hints on demand did not require students to work on any of the steps leading to the answer.

6. Tutored problem solving vs. worked examples

Students are often taught new material in mathematics by first being introduced to the principles needed to understand the new material, then worked examples that show how to use the principles to solve related problems and finally, practice problems for the students to work on. Traditionally, teachers often present only a few examples and assign a large number of practice problems. Likewise, learning technologies for mathematics often focus heavily on tutoring step-by-step problem solving with positive learning results [i.e., Cognitive Tutors (Anderson J. R., Corbett, Koedinger, & Pelletier, 1995), Andes (VanLehn, Graesser, Jackson, Jordan, Olney, & Rose, 2005) and the ASSISTment System (Razzaq, et al., 2007)] rather than presenting information about principles or presenting many worked examples.

Cognitive scientists have been interested in the role of worked examples in reducing cognitive load and helping students to learn, and there have been numerous studies on the effectiveness of worked examples (Chi M. T., Siler, Jeong, Yamauchi, & Hausmann, 2001; Graesser, Moreno, Marineau, Adcock, Olney, & Person, 2003). Sweller and Cooper (1985) presented evidence that supported their hypothesis that worked examples helped novices to acquire “schemas” which they defined as “mental constructs that allow patterns or configurations to be recognized as belonging to a previously learned category and which specify what moves are appropriate for that category.” It appears that novices who have not learned the required schemas have to depend on superficial search strategies in solving problems (Larkin, McDermott, Simon, & Simon, 1980) while experts can choose the next appropriate step based on their ability to correctly categorize the problem.

Sweller and Cooper's work suggested that problem-solving practice did not help students to acquire schemas as efficiently as the use of worked examples perhaps because of the change of focus from "goal-directed problem-solving" to "problem-state configurations." Kalyuga, Ayres, Chandler & Sweller (2001) presented results that point to a benefit of using worked examples with novice students and then using problem-solving practice later as students show more understanding.

Tutored problem solving helps students solve a problem by providing feedback and help on each step of a problem and is more interactive than reading worked examples. Several studies in the literature have found evidence of the benefit of greater interaction. Comparing Socratic and didactic tutoring strategies, Core, Moore & Zinn (2003) found that the more interactive (based on words produced by students) Socratic tutorial dialogs correlated more with learning. Chi, Siler, Jeong, Yamauchi and Hausmann (2001) found that students who engaged in a more interactive style of human tutoring were "able to transfer their knowledge better than the students in the didactic style of tutoring." Evens and Michaels (2006) compared expert human tutoring to reading a text book with the same material and found that the tutored students got significantly higher scores on a post-test. Results that support greater interaction have also been found in studies of intelligent tutoring systems (Graesser, Moreno, Marineau, Adcock, Olney, & Person, 2003; VanLehn, Graesser, Jackson, Jordan, Olney, & Rose, 2005). Additionally, there is evidence that tutored problem solving is more effective for less proficient students than less interactive methods of tutoring (Razzaq, Heffernan, & Lindeman, 2007).

Other researchers have been interested in comparing tutored problem solving to worked examples. Kim et al. (2009) found an advantage for worked examples for

conceptual problems over tutored problem solving while tutored problem solving was better for procedural problems. Schwonke et al. (2007) found that students learned more from gradually fading worked examples to tutored problem solving than from tutored problem solving alone. In Schwonke et al.'s work, the fading of worked examples was the same for all students and did not depend on their demonstration of understanding.

Salden, Alevan, Renkl and Schwonke (2008) experimented with an adaptive fading scheme where worked examples were gradually faded when students showed understanding based on their self-explanations. Salden et al. found evidence that adaptively fading worked examples was more effective than fixed fading.

The study described in (Shrestha, Maharjan, Wei, Razzaq, Heffernan, & Heffernan, 2009) investigated whether students in a classroom setting would benefit more from interactive tutored problem solving than from worked examples given as a feedback mechanism. The study attempted to determine whether results will differ depending on the students' math proficiency. It was expected that less proficient students would benefit more from tutored problem solving than more proficient students. The ASSISTment System was used to test this hypothesis.

6.1 Participants

This experiment was conducted with 8th grade students in three local middle schools located in central Massachusetts. One of the schools was suburban, while the other two were urban. Over 80% of the students who participated were from a school which according to its state test scores is in the bottom 5% in the state and has been labeled by the *No Child Left Behind Act* as not making adequate yearly progress. The experiment took place in the months of April and May of 2008 at the computer labs of the respective

schools. The students who participated in this experiment were exposed to both conditions: tutored problem solving, and worked examples. They were given problem sets to work on, and their actions were logged and later analyzed.

6.2 Materials

For the experiment we created nine problem sets, each consisting of four to five ASSISTments. All of the main questions of the ASSISTments were taken from 6th Grade MCAS tests for Mathematics (2001 – 2007) focusing on the *Patterns, Relations and Algebra* section, which concentrates on different mathematics skills: populating a table from a relation, finding a missing value in a table, using fact families, determining equations for relations, substituting values into variables, interpreting relations from number patterns, and finding values from a graph.

6.3 Procedure

Each problem set in this study was a collection of ASSISTments grouped into three sections: pre-test, experiment, and post-test. For the experiment, students were considered to have completed a problem set only if they finished every part of it. Time on task was not controlled. The gain score from pre-test to post-test was used to determine whether students had learned anything from the conditions.

When students started a problem set, they were first given a pre-test problem. The pre-test was an ASSISTment with a single question, and did not include any form of help or hint. In order to make sure that the students understood what was happening, they were informed that the question was a pre-test and that they would not receive feedback on

whether their answer was right or wrong. They were also informed that the question would be repeated at the end of the problem set.

In the pretest, students were allowed only one attempt to answer the question, so the first answer they provided was considered as the final answer for the pre-test and it could not be changed. After the one question pretest, students were presented with the first question from a randomly chosen condition. The computer randomly assigned either the tutored problem solving or the worked example condition to the students. This part consisted of two or three ASSISTments all in the same condition. Within the two conditions, students did the same number of questions, so the content of the questions was held constant between the conditions. Finally, when the students finished all of the ASSISTments in the experiment section, they were given the post-test which was the pre-test question repeated. Also similar to the pre-test, the first response of the student was recorded and used for analysis. However, unlike the pre-test, we did inform the students regarding the correctness of their answer. Learning was assessed by comparing the results of the pre-test and the post-test.

In the worked example condition, when a student gave an incorrect answer or pressed on the “Break this problem into steps” button, a problem that was similar to the main question was shown solved step by step. As such, the students would have a pattern to follow in order to solve the problem. The worked example condition was shown in Figure 13. The student was asked to read through the worked example and choose “I have read the example and now I am ready to try again” when he/she was done. The student was then asked to do the original problem again.

In the tutored problem solving condition, students who got a problem wrong were asked to answer a set of questions that break down the main problem into steps

(presented one at a time), shown in Figure 14. If the student provided a wrong answer or pressed on the “Break this problem into steps” button, he/she would be directed to the first scaffolding question, which helped the student to understand the first step to solve the original problem. Students could ask for hints on each step if they needed more help. If the student pressed on the “Show me a hint” button, hints would be shown one by one until the student reached a “bottom-out hint” which was typically the answer to the scaffolding question. After answering the question correctly, the student was directed to the next scaffolding question. The number of scaffolding questions depended on the complexity of the original question. At the end, the student was expected to understand how to do the original problem step by step.

During the experiment, teachers introduced the problem sets as a regular assignment. As such, students were not aware of the randomized controlled experiment. They were neither briefed about the problem set structure nor the number of ASSISTments in a problem set. Thus, students might not have been aware that they were taking a pre-test until they submitted an answer, as they were told that the question they answered was a pre-test only after answer submission.

We do not distinguish the experiment section from the post-test with any specific instruction or notice like we do in the pre-test. The only way a student can know that they were in the post-test is if they realized that the pre-test question has been repeated. It should be noted that some students were not exposed to either of the conditions since conditions are introduced only when a student makes a mistake in the first response. If students answered all of the ASSISTments in a problem set correctly in their first attempt then they would not have been exposed to any of the conditions and their performance on that problem set were not included in the study.

Karen purchased a new camera for \$60. She also purchased 5 rolls of film. The total cost of the camera and the rolls of film was \$90. Karen's purchase is represented by the equation below. In the equation, f stands for the cost of each roll of film.

$$5f + 60 = 90$$

What was the cost of each roll of film that Karen purchased?

Break this problem into steps

Type your answer below (mathematical expression):

Submit Answer

Let's move on and figure out this problem.

Let's look at the solution for a problem similar to the one above:

Simond purchased a new play-station for \$100. He also purchased 4 game CDs. The total cost of the game and the CDs was \$200. Simond's purchase is represented by the equation below. In the equation, f stands for the cost of each CD.

$$4f + 100 = 200$$

What was the cost of each CD that Simond purchased?

Solution:

Let us first try to find out the amount Simond spends to buy CDs. Simond spends \$100 out of \$200 on the play-station and the rest on CDs. So to find the amount he spends on CDs we need to subtract 100 from 200.

$$200 - 100 = 100.$$

So he spends \$100 on cds.

Now, we know that Simond spends \$100 on CDs and that he buys a total of 4 CDs. Let us try to find the cost of each CD. Simond bought 4 CDs for \$100. So if f represents the cost of each CD, we add f 4 times to get \$100.

$$f + f + f + f = 100$$

We need to find a number that when added 4 times to itself gives us 100. This number can be found by dividing 100 by 4.

$$100 \div 4 = 25.$$

So, \$25 must be the cost of each CD.

[Comment on this question](#)

I have read the example and now I am ready to try again.

Submit Answer

✔ Correct!

Karen purchased a new camera for \$60. She also purchased 5 rolls of film. The total cost of the camera and the rolls of film was \$90. Karen's purchase is represented by the equation below. In the equation, f stands for the cost of each roll of film.

$$5f + 60 = 90$$

What was the cost of each roll of film that Karen purchased?

The cost of each roll of film is \$6.
Type in 6. [Comment on this hint](#)

Type your answer below (mathematical expression):

Submit Answer

✔ Correct!

You are done with this assignment!

Figure 13: The worked example condition requires students to read the example and then try to answer the question again.

Karen purchased a new camera for \$60. She also purchased 5 rolls of film. The total cost of the camera and the rolls of film was \$90. Karen's purchase is represented by the equation below. In the equation, f stands for the cost of each roll of film.

$$5f + 60 = 90$$

What was the cost of each roll of film that Karen purchased?

[Comment on this question](#)

Break this problem into steps

Type your answer below (mathematical expression):

Submit Answer

Let's move on and figure out this problem.

Karen spends a total of \$90 on films and cameras, and spends the rest to buy 5 films. Out of the \$90, she spends \$60 on camera

What is the amount of money Karen spends to buy films?

[Comment on this question](#)

Karen spends \$60 out of \$90 on the camera and the rest on films. So to find the amount she spends on films we need to subtract 60 from 90.

[Comment on this hint](#)

$90 - 60 = 30$. So she spends \$30 on films. Type in 30.

[Comment on this hint](#)

Type your answer below (mathematical expression):

Submit Answer

✓ Correct!

We know that Karen spends \$30 on films and that she buys a total of 5 films.

What is the cost of each film?

[Comment on this question](#)

Karen bought 5 rolls for \$30. So if f represents the cost of each roll we add f 5 times to get \$30.

$$f + f + f + f + f = 30$$

[Comment on this hint](#)

We need to find a number that when added 5 times to itself gives us 30. This number can be found by dividing 30 by 5.

[Comment on this hint](#)

$$30 \div 5 = 6$$

[Comment on this hint](#)

$6 + 6 + 6 + 6 + 6 = 30$. So 6 must be the cost of each roll. Type in 6.

[Comment on this hint](#)

Type your answer below (mathematical expression):

Submit Answer

✓ Correct!

You are done with this assignment!

Figure 14: The tutored problem solving condition requires students to work through each step of the problem.

6.4 Results

Our experiment used a repeated measures design where students participated in a different number of experiments, and each time the student started an experiment, he/she was randomly assigned to one of the two conditions. For the analysis, we only considered the students who had completed at least one problem set in both of the conditions and ignored all other students who were exposed to only one condition. Problem sets that were not completed were ignored. In addition, we also ignored students who correctly answered both the pre-test and the post-test questions, as we assumed the student had mastered that material. Since repeated measure design suffers from ordering effects, we relied on the random assignment of conditions as a control for that effect.

Out of a total of 186 participants, 166 students completed at least one problem set and we had data from a total of 866 attempts at completing a problem set. We then ignored data where both pre-test and post-test answers were correct. We also ignored data from students who completed only one of the two conditions. We then had a total of 68 students who participated in both tutoring conditions. So this means each of the 68 students completed at least one problem set where they were given tutored problem solving and at least one problem set where they were given worked examples.

For each student, the average learning gain from tutored problem solving and the average learning gain from worked examples were calculated. Learning gain for a problem set was defined to be the post-test score minus the pre-test score. Average learning gain for the tutored problem solving condition was defined to be the average of the learning gains for the entire problem sets that the student did when they were

assigned to the tutored problem solving condition. Similarly, the average learning gain for worked examples was the average of the gains for all of the problem sets that the students did when they were assigned to the worked examples condition. There was no need to check if both groups were balanced at pre-test since our experiment was a repeated measure design and each student participated in both conditions.

There was a significant effect for condition with tutored problem solving receiving higher gain scores than worked examples (35% average gain vs. 13% average gain), $t(67) = 2.38, p = 0.02$. These results are shown in Table 1.

Table 7: Mean gain scores for both conditions.

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Scaffold	.3498	68	.53799	.06524
	Worked	.1306	68	.52168	.06326

To determine whether there was an aptitude-treatment interaction we calculated a student math proficiency score using an Item Response Theory (IRT) model which takes into account difficulty of ASSISTments and how students performed on ASSISTments throughout the school year.

We did not have IRT scores for five students, so this analysis was done on data from 63 students. We did a median split on the IRT scores to categorize students as either high or low proficiency. We did not find a significant difference based on math proficiency ($F(1, 61) = .158, p = 0.69$). Both high proficiency and low proficiency students learned more from the tutored problem solving condition than from the worked

example condition. High proficiency students had a mean gain of 47% with tutored problem solving and 22% with worked examples ($t(29) = 1.599, p = 0.12$). Low proficiency students had a mean gain of 20% with tutored problem solving and 3% with worked examples ($t(32) = 1.404, p = 0.17$).

Because the tutored problem solving is more interactive, it does consume more time. Tutored problem solving ($M = 244$ seconds) took significantly more time on average than worked examples ($M = 166$ seconds), ($t(66) = 2.93, p = 0.002$).

6.5 Conclusions

Our study compared the effectiveness of tutored problem solving versus worked examples when used as feedback. Students participated in the study in a classroom environment and the problems were presented as classroom assignments. Our results indicate that tutored problem solving is significantly better than worked examples in terms of the average gain of students in each condition. Furthermore, we did not find an aptitude-treatment interaction.

Our study differed from previous studies in that we compared worked examples to tutored problem solving rather than untutored problem solving. It also differs in that we presented worked examples as feedback after students unsuccessfully attempted to solve a problem rather than presenting them before they attempted problem solving.

We speculate that many studies that have found positive results for worked examples were done in lab settings, where an adult lab attendant provided the extra focusing attention that a classroom environment does not provide. Perhaps in the classroom setting, the more interactive tutored problem solving condition was superior

due to the fact that the higher interactivity level required from tutored problem solving better engages students' focus. This theory suggests that students with greater focus might yield results that would be more in line with the current literature.

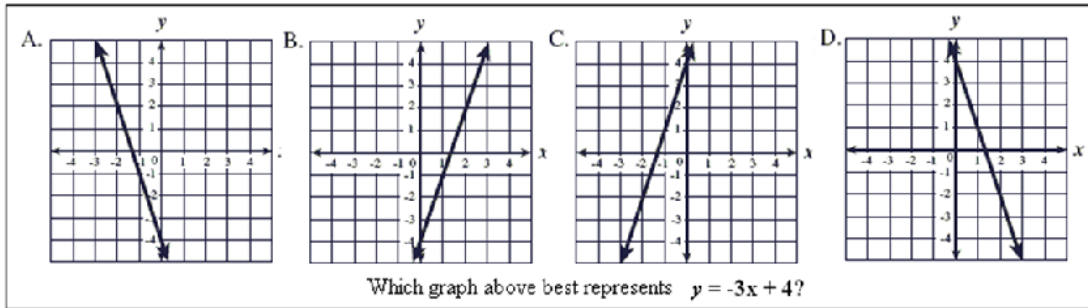
Salden et al. (2008) thought of their results as an instance of the *Assistance Dilemma* coined by Koedinger and Alevan (2007) which studies the dilemma of when to give assistance to students versus when to withhold information in an attempt to get students to generate information on their own. The Assistance Dilemma would consider worked examples to be “high assistance” while tutored problem-solving to be “low assistance”. However, this does not seem to consider that these may be seen differently by different students. For instance, Chi, Bassok, Lewis, Reimann, & Glaser (1989) found a difference in the way that students used worked examples based on their proficiency in problem-solving: “... we find that the Good students use the examples in a very different way from the Poor students. In general, Good students, during problem solving, use the examples for a specific reference, whereas Poor students reread them as if to search for a solution.”

In the next section, (Razzaq, Heffernan, & Lindeman, 2007) found that students who received worked-out *solutions* to problems rather than tutored problem-solving learned more only if they were above average students. Below average students did better with tutored problem-solving. This is important because it raises the question about whether worked examples are always a better thing to do before problem solving for all students. We think our theory can explain the current results in this area. In particular, we speculate that the students in the recent Salden et al. study (2008) might have been just the right type of well focused students that could benefit from reading worked examples. However, to help the less focused student then tutored problem solving is superior.

This conclusion is reasonable in a few respects. Firstly, these two conditions have different degrees of interactivity. In the worked examples condition, a student is shown a completely solved example problem which is similar to the main problem. The student is only one click away from answering the original problem again. In contrast, the tutored problem solving condition asks several subsequent questions pertaining to the main problem, all of which have to be completed before returning to the main problem again. For most students, it is reasonable to assume that answering questions frequently keeps them more focused than just reading from a screen.

7. Tutored problem solving vs. worked out solutions

This study described by Razzaq, Heffernan & Lindeman (2007) pays attention to the interaction between proficiency of the student and condition. We look at three different conditions, in the ASSISTment system. The first two conditions are the same as in Razzaq & Heffernan (2006) (*tutored problem solving* and *hints on demand*). The third condition is a *delayed feedback* condition where students get no feedback from the tutor until they finish all of the problems in the experiment, whereupon they receive worked out solutions to all of the problems.



Tutored Problem Solving

Will the graph that best represents this equation have a positive slope or a negative slope?
 Negative Positive

Right. The slope of the line will be negative. Which graph(s) show a line with a negative slope?
 A. B. and C.
 A. and B. C.
 A. and D. C. and D.
 B. D.

Good. Now we know that the correct answer will be graph A or graph D. According to the equation, $y = -3x + 4$, what is the y-intercept of the line?

A.

D.

Now which graph shows a line with a y-intercept of 4?
 A.
 D.

A.

D.

The y-intercept of the line is the value of y at the point where the line crosses the y-axis. Which graph shows a y-intercept of 4?
 The answer is D. Please choose D.

Hints on demand condition

No, that is not correct. Please try again.

The equation of a line can be written as follows:
 $y = mx + b$
 where m is the slope and b is the y-intercept.
 According to the equation $y = -3x + 4$, what is the slope of this line? Is the slope of the line positive or negative?

The slope of the line is -3, so it is negative. Which graph(s) show a line with a negative slope?

The line is sloping upward (left to right), the slope is positive.

The line is sloping downward (left to right), the slope is negative.

Graphs A and D have negative slopes.

According to the equation, what is the y-intercept of the line?
 The y-intercept is 4. Which graph shows a line with a y-intercept of 4?

A.

D.

The y-intercept is the y value of the point where the line crosses the y-axis. Which graph shows a line with a y-intercept of 4?
 The answer is Graph D.

Figure 15: An ASSISTment item showing the tutored problem solving and hints on demand conditions.

The purpose of this experiment was to determine which level of interaction worked best for students learning math: *tutored problem solving*, *hints on demand* or *delayed feedback*, and how their math proficiency influenced the effectiveness of the feedback provided.

7.1 Experiment Design

Problems in this experiment addressed the topic of interpreting linear equations. Figure 15 shows an item used in the experiment. The item shows the different feedback that students can receive once they have answered a question incorrectly. We call this top-level question the original question.

A student in the *tutored problem solving* condition is immediately presented with the first scaffolding question. Students must answer a scaffolding question correctly to proceed and receive the next scaffolding question (or finish the problem). Students can ask for hints on the scaffolding questions, but not on the original question. They cannot go back and answer the original question, but rather are forced to work through the problem. Figure 15 shows the scaffolding questions for one of the problems in the experiment.


Students in the hints condition receive a message, outlined in red, of “No, that is not correct. Please try again.” The hints, outlined in green, appear when the student requests them by pressing the Hint button. Students do not see the hints unless they ask for them. Figure 15 shows a sequence of seven hints to solving the problem outlined in green. The bottom-out hint gives the student the answer to the problem.

Students in the *delayed feedback* condition did not receive any feedback on the problems that they did until they had finished all of the problems. At that time, the students were presented with the answers and explanations of how to solve the problems. Figure 16 shows the explanation that students in the *delayed feedback* condition received for the item shown in Figure 15.


Based on the results of Razzaq and Heffernan (2006), we hypothesized that less-proficient students would need more interaction and benefit more from the tutored problem solving than more-proficient students.

The correct answer is Graph D. Read the following explanation to see how to find the answer.

The equation of a line can be written as follows:
 $y = mx + b$
 where m is the slope
 and b is the y-intercept.
 According to the equation $y = -3x + 4$, what is the slope of this line?
 Is the slope of the line positive or negative?
 The slope of the line is -3 , so it is negative.
 Which graph(s) show a line with a negative slope?



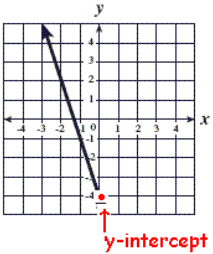
The line is sloping upward (left to right), the slope is positive.



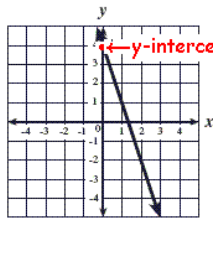
The line is sloping downward (left to right), the slope is negative.

Graphs A and D have negative slopes.
 According to the equation, what is the y-intercept of the line?
 The y-intercept is 4. Which graph shows a line with a y-intercept of 4?

A.



D.



The y-intercept is the y value of the point where the line crosses the y-axis.
 Which graph shows a line with a y-intercept of 4?
 The answer is Graph D.

Figure 16: The delayed feedback explanation

For this experiment, the number of problems was held constant, but students took as much time as they needed to finish all of the problems. Students were presented with two pretest problems, four experiment problems and four post-test problems that addressed the topic of interpreting linear equations. There were three versions of the experiment problems, one for each condition. The two pretest problems were repeated in the post-test.

The ASSISTment system randomly assigned students to the *tutored problem solving*, *hints on demand* or *delayed feedback* conditions with equal probability. There were 366 eighth grade students from the Worcester Public Schools in Worcester, Massachusetts who participated in the experiment: 131 students were in honors level classes and 235 were in regular math classes. There were 119 students in the *tutored problem solving* condition, 124 students in the *hints on demand* condition and 123 students in the *delayed feedback* condition. The students worked on the problems during their regular math classes.

7.2 Results

We excluded students who got all of the pretest problems correct from the analysis because we assumed that they knew the material. Fifty-one students got perfect scores on the pretest and were excluded. We first checked to make sure that the groups were not significantly different at pretest by doing an ANOVA on pretest averages by condition. There was no significant difference between groups at pretest ($p = 0.556$). Students learned overall from pretest to post-test ($p = 0.005$).

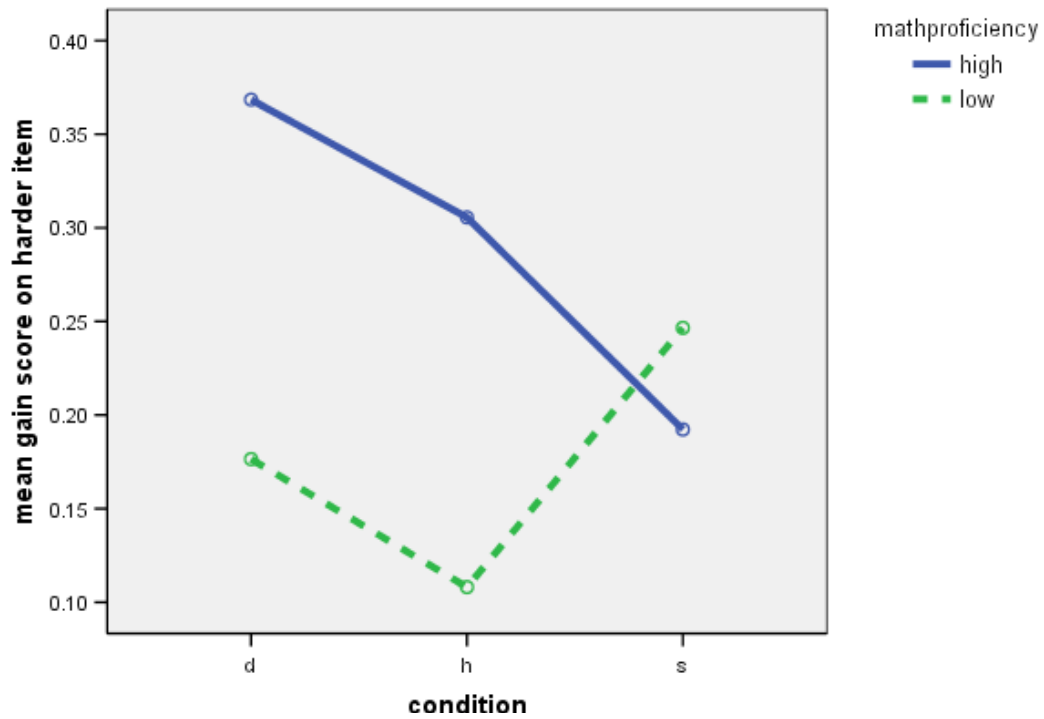
When we look at performance on the post-test by condition, the difference is not significant; however there is an interesting trend when we separate students by math

proficiency. There is a significant interaction ($p = 0.045$) between condition and math proficiency on the post-test average. The regular students seem to benefit more from the *tutored problem solving* condition, while honors students seem to benefit more from the *delayed feedback* condition.

We decided to take a closer look at the item that proved most difficult for students. The problem concerned finding the y-intercept from an equation and was presented to students in the pretest and again in the post-test. We did a one-way ANOVA using math proficiency as a covariate. An interaction between condition and math proficiency ($p = 0.078$) shows that honors students performed best when they received *delayed feedback* and the regular students performed best when they received *tutored problem solving*. Less-proficient students learned significantly more with Tutored Problem Solving than Worked-out Solutions ($p < 0.05$). More-proficient students learned significantly more with Worked-out Solutions than Tutored Problem Solving ($p < 0.075$). (See Figure 17)

We interpret the low p-values on the interaction term to mean that there are different rates of learning on the single items based upon the interaction between the level of math proficiency and condition. Students who come in with less knowledge benefit more from the *tutored problem solving* than students who come in with more knowledge. Students who come in with more knowledge benefit from the *delayed feedback* more than the other groups.

Interaction between condition and proficiency



Descriptive Statistics

Dependent Variable: gainscore_y_intercept

mathproficiency	condition	Mean	Std. Deviation	N
high	d	.3684	.48885	38
	h	.3056	.46718	36
	s	.1923	.40192	26
	Total	.3000	.46057	100
low	d	.1765	.38405	68
	h	.1081	.31264	74
	s	.2466	.43400	73
	Total	.1767	.38234	215

Tests of Between-Subjects Effects

Dependent Variable: gainscore_y_intercept

Source		Type III Sum of Squares	df	F	Sig.
Intercept	Hypothesis	14.522	1	124.897	.007
	Error	.237	2.039		
mathproficiency	Hypothesis	.835	1	1.968	.295
	Error	.853	2.010		
condition	Hypothesis	.232	2	.272	.786
	Error	.852	2		
mathproficiency * condition	Hypothesis	.852	2	2.577	.078
	Error	51.099	309		

Figure 17: There are interactions between condition and math proficiency approaching significance.

7.3 Immediate vs. Delayed Feedback

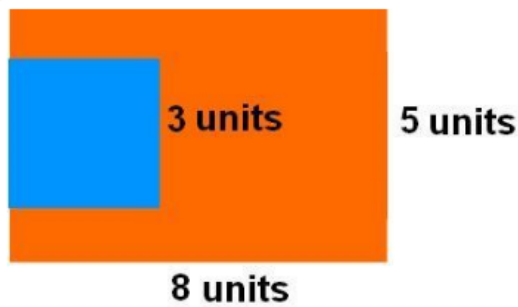
What was it about the solution condition that benefited some of the students? Was it the worked out solutions themselves that were presented at the end of the assignment or *when* they were presented?

Both immediate and delayed feedback have been shown to be helpful to students (Mathan & Koedinger, 2003). Razzaq and Heffernan (2004) reported that a human tutor provided immediate feedback to student errors, on most occasions, keeping students on the correct solution path. There was one out of 26 problems where the human tutor gave delayed feedback to a student error. In this instance, the tutor allowed the student to continue where she had made an error and then let her check her work and find the error seemingly promoting evaluative skills. This happened with the student who was taking a more advanced algebra class and was slightly more advanced than the other students in the study. However, for the other 25 problems, the tutor provided immediate feedback to promote the development of generative skills. McArthur et al. (1990) reported something similar in their examination of tutoring techniques in algebra where they collected one and a half hours of videotaped one-on-one tutoring sessions. “In fact, for every student error we recorded, there was a remedial response. At least the tutors we observed were apparently not willing to let students explore on their own and perhaps discover their own errors...teachers may believe that such explorations too frequently lead to unprofitable confusion for the student. (pp. 209)”

An experiment was run to compare showing solutions to students immediately versus showing the solutions after the assignment was completed. The assignment was focused on finding the area of irregular figures (see Figure 18).

7.3.1 Experiment Design

The first condition presented the immediate feedback of a worked solution when students made an error. The second condition presented delayed feedback of worked solutions to students after they completed a set of five problems. Students worked on two pretest problems before the assignment and were randomly assigned to a condition. After finishing the five problems in the immediate or delayed feedback condition, students completed two post-test problems which were the same problems that they saw in the pretest. The dependent measure was gain from pre- to post-test.



The figure above shows a square inscribed in a rectangle.
 What is the area of the orange colored part of the figure?

[Comment on this question](#)

Break this problem into steps

Type your answer below:

Submit Answer

Here is a solution to the problem. Read it carefully and then choose Ok to go to the next problem.

To find the area of the orange part of the figure, you need to do the following.

1. Find the area of the rectangle.
2. Find the area of the square.
3. Subtract the area of the square from the area of the rectangle.



So let's do the first step. What is the area (in square units) of the rectangle?

To find the area of a rectangle, we multiply the length by the width.

So the area of the rectangle is $8 * 5 = 40$ square units.

Now we need to find the area of the square.

To find the area of a square, we also multiply the length by the width although for a square these lengths will always be equal.

So the area of the square is $3 * 3 = 9$ square units.

Finally, step 3 is to subtract the two areas we just found to get the area of the orange colored part of the figure.

$40 - 9 = 31$ square units

Figure 18: The immediate feedback condition gave students the solution immediately while the delayed feedback condition gave the student the solution at the end of the assignment.

7.3.2 Results

Thirty 8th grade students participated in the study with nine students in the delayed feedback condition and 21 students in the immediate feedback condition. The average gain for students in the immediate feedback condition was 0.056 and the average gain for students in the delayed feedback condition was 0.167. There was not a significant difference between the two conditions ($p = 0.23$), so we cannot conclude from this experiment whether or not delaying feedback to students plays a role in learning. The results may be due to the small sample size or the short duration of the experiment.

7.4 Conclusion

The results of Razzaq, Heffernan and Lindeman (2007) were surprising. We did find evidence to support the interaction hypothesis for regular students. The regular students performed best in the *tutored problem solving* condition, which is the most interactive condition. We did not expect students in the *delayed feedback* condition to learn more than in other groups, however, the honors students did better in this condition than in the *tutored problem solving* or *hints on demand* conditions.

One possible explanation is that less-proficient students benefit from more interaction and coaching through each step to solve a problem while more-proficient students benefit from seeing problems worked out and seeing the big picture. Another possible explanation, put forth by one of the eighth grade teachers, is that honors students are often more competitive and like to know how they do on their work. The *delayed feedback* group had to wait until the end of the assignment to see whether they got the questions right or wrong. Perhaps the honors students ended up reading through the

explanations more carefully than they would have read the scaffolding questions or hints because they were forced to wait for their results.

We believe the results of this experiment provide further evidence of the interaction hypothesis for less-proficient students. However, the interaction between condition and math proficiency presents a good case for tailoring tutor interaction to types of students to maximize their learning.

8. Tutored problem solving vs. educational web pages

The World Wide Web has thousands of pages of educational content, some of them are very well written. These web pages often include definitions, examples, images and animations. Can students learn from educational web pages? How do these web pages compare to tutored problem solving that is specific to a problem?

8.1 Can students learn from viewing a web page of educational content?

The purpose of this study was to determine if students could learn from existing public web pages to help them solve problems in the ASSISTment System. Was it the case that web pages would be ignored? What percentage of students would show any benefit from having visited a web page? This study was conducted with eighth graders for two topics that are typically covered in middle school: Pythagorean Theorem and Venn Diagrams. Google's search engine was used to find web pages about the two topics on November 24, 2008 and December 15, 2008 and two "good" pages for each topic were chosen.

We evaluated 13 pages about Pythagorean Theorem, using our own judgment before finding two that we wanted to use. For instance, we decided not to use the first result, found on Wikipedia.com (http://en.wikipedia.org/wiki/Pythagorean_theorem), because it appeared to be too advanced for eighth graders. We also excluded the state of New York's Regents Exam Prep web page (regentsprep.org/regents/Math/fpyth/Pythag.htm) because it had the answer to one of the questions in the problem set. We chose PBS's page on the Pythagorean Theorem, (www.pbs.org/wgbh/nova/proof/puzzle/theorem.html), because it was age appropriate and highly ranked by Google. Math Forum's page on the Pythagorean Theorem

(mathforum.org/dr.math/faq/faq.pythagorean.html) was excluded because it contained a link to the PBS page. Similarly, we chose two web pages on the topic of Venn Diagrams.

Students worked through six problems in each topic. Two of the problems, the second and fifth problems, asked students to click on a link to display a web page if they had answered the question incorrectly instead of tutoring them through each step. After studying the webpage, students were asked to go back to the problem and try answering again. A counter-balanced design was used to control for the order of the problems, the order of the topics and the order of the web pages.

The Venn diagram below shows the number of seventh-grade students at Berkshire Middle School enrolled in Band, Art, Spanish, or any combination of the three elective classes.

Region	Number of Students
Band only	40
Art only	32
Spanish only	30
Band & Art	12
Band & Spanish	10
Art & Spanish	14
All three (Band, Art, Spanish)	8

What is the total number of seventh-grade students at Berkshire Middle School who are enrolled in Art or Spanish, but not in Band?

Answer: 14

Submit Answer

✗ Sorry, that is incorrect. Let's move on and figure out why!

Click on the following link and read the webpage. When you are ready, return to this page and answer the question below.

[Click here](#)

I found this webpage to be:

Select one:

- Very Useful
- Useful
- Somewhat Useful
- Not Useful
- I had technical difficulties trying to open the webpage

Submit Answer

Lesson Page
Venn Diagrams
(Sets)
Math A

A Venn diagram is a drawing, in which circular areas represent groups of items sharing common properties. The drawing consists of two or more circles, each representing a specific group. This process of visualizing logical relationships was devised by John Venn (1834-1923).

Each Venn diagram begins with a rectangle representing the universal set. Then each set in the problem is represented by a circle. Any values that belong to more than one set will be placed in the sections where the circles overlap.

The Venn diagram at the left shows two sets **A** and **B**. Values that belong to both set **A** and set **B** are

Figure 19: Students would be directed to a webpage on Venn Diagrams if they answered the problem incorrectly.

There were 130 students who participated in the study. The measure of learning was whether a student who got a problem wrong could get the problem correct right after they visited the web page. It was found that 60% of the students who got it wrong the first time, after visiting the first web page on Venn Diagrams, could then do the problem. Of the students who visited the first web page on Pythagorean Theorem, 37% of the students who got the problem wrong the first time could then correctly solve the problem. Overall, the results were encouraging in that so many students showed some benefit. However, it is not known if some of the students would have gotten the item correct on their 2nd attempt without a web page for instance. Nevertheless, overall, it confirmed that a student's success can be changed by their exposure to a web page for a few minutes.

Two web pages were chosen for each topic, and students saw them in a random order. Gain scores were not significantly different when we compared one web page to another.

8.2 Is Tutored Problem Solving superior to viewing educational web pages?

The purpose of this randomized controlled experiment was to compare tutored problem solving to viewing an educational web page. The study described in the previous section showed that students could learn from viewing a web page about the topic of a problem to be solved. If viewing a web page can help students to learn as much or more than tutored problem solving, time and resources spent on content development could be significantly reduced.

Seventy-one middle school students participated in the study. During one class period, students solved problems in two topics (Venn Diagrams and Pythagorean Theorem) and participated in both conditions in a repeated measures design. The experiment controlled for the order of topics and the order of conditions which were randomly assigned by student (see Table 8). A pretest and post-test was given before and after each topic.

Table 8: Students were randomly assigned to one of four conditions.

	Condition 1	Condition 2	Condition 3	Condition 4
First	Venn Diagram TPS	Pythagorean Theorem Web Page	Venn Diagram Web Page	Pythagorean Theorem TPS
Second	Pythagorean Theorem Web Page	Venn Diagram TPS	Pythagorean Theorem TPS	Venn Diagram Web Page

The Pythagorean Theorem topic was more difficult (mean pretest score = 45%) compared to the Venn Diagram topic (mean pretest score = 58%). The results showed that there was learning overall ($p < 0.001$) with a mean gain of 33% for the Pythagorean Theorem topic and 19% gain for the Venn Diagram topic. The difference between tutored problem solving and web pages was not statistically reliable ($p = 0.27$), however, students learned more with tutored problem solving than viewing the web pages with 0.22 effect size. Looking at the more difficult topic, Pythagorean Theorem, showed a 0.4 effect size for tutored problem solving over the web page.

8.3 Conclusion

These studies showed that students could learn from viewing educational web pages, which can encourage tutoring system developers to make more use of free educational content on the web. Tutored problem solving appeared to help students more, though not reliably more, especially when the topic was more difficult. Could this result be because the tutored problem solving condition targeted specific problems and how to solve them while the web pages were more focused on the broader principles involved in the topic? We do not know, but believe that there is potential for future work here.

9. Is there a tradeoff between the value of interactive tutoring and the extra time it takes?

Tutored problem solving takes longer than hints on demand or worked solutions. Do students benefit more from this interactive tutoring or from getting practice on more problems which they could do in the same amount of time?

This is an important question since teachers are being pressured to cover more material in less time. Teachers would like to get more “bang for their buck” since instructional time is more precious. If students can cover more material in the same amount of time with comparable learning results then that approach may be preferable.

The following experiment, reported in (Razzaq & Heffernan, 2009), controls for time rather than the number of problems and should tell us more about when it is beneficial for students to spend more time on problems by going through scaffolding, and when it would be better to get more problems done for more practice.

In the previous study (Razzaq, Heffernan, & Lindeman, 2007), the more time-consuming interactive tutored problem solving was indeed more helpful to less proficient students when compared to simply showing them a solution to the problem. On the other hand, it was not as helpful to more proficient students who benefited more from seeing solutions. We hypothesize that in the classroom setting, tutored problem solving was superior for less-proficient students due to the fact that the higher interactivity level required from tutored problem solving better engages students’ attention. This theory would suggest that students who were better able to learn from reading a solution had greater focus. In addition, the more-proficient students may have more prior knowledge that prepares them to learn from reading text (Schwartz & Bransford, 1998).

The previous study (Razzaq, Heffernan, & Lindeman, 2007) controlled for the number of problems done. Tutored problem solving was found to take significantly more time than seeing a solution. The purpose of this study was to determine whether tutored problem solving was worth the extra time it took (particularly for less-proficient students) or if students would benefit from practice on more problems given the same amount of time.

The hypothesis is that less-proficient students in a classroom setting will benefit more from interactive tutored problem solving than from reading solutions and doing more problems, while more-proficient students will benefit more from reading solutions and doing more problems than less-proficient students. Unlike the study in Razzaq, Heffernan & Lindeman (2007) there was no delay in presenting the worked solutions to students.

9.1 Experiment Design

A counterbalanced design was used where each student participated in two conditions: Tutored Problem Solving and Solutions. We designed two problem sets: 1) slope, intercept and linear equations and 2) symbolization. Students had been introduced to these topics in their regular math class before this study took place. Figure 20 shows a problem that appeared in the symbolization problem set, with the Tutored Problem Solving approach shown on the left and the Solutions approach shown on the right.

To control for order effects each group received treatments in a different order. Four classes of 8th grade students participated in the study, which took place over two days in the school computer lab. Only students who completed both problem sets were

included in the analysis. Students were asked to work on their own without help from their classmates. On the first day, students worked for 20 minutes on one of the problem sets using one of the strategies. On the second day, students worked for 20 minutes on the second problem set using the second strategy. After 20 minutes of working on the problem set, all of the students were asked to stop working. Then they were given the post-test and asked to finish all of the problems on the post-test. Students could work on the post-test until the end of class time, approximately 20 minutes. The pre-test problems were the same as the post-test problems, although students received no feedback on the pre-test whether they answered them correctly or not. Table 9 shows the experiment design.

Table 9: Experiment design.

Session (1 class period per session)	Symbolization First Group		Slope and Intercept First Group	
	Class A	Class B	Class C	Class D
Day 1	Pretest Symbolization (Solutions) Post-test	Pretest Symbolization (TPS) Post-test	Pretest Slope & Intercept (TPS) Post-test	Pretest Slope & Intercept (Solutions) Post-test
Day 2	Pretest Slope & Intercept (TPS) Post-test	Pretest Slope & Intercept (Solutions) Post-test	Pretest Symbolization (Solutions) Post-test	Pretest Symbolization (TPS) Post-test

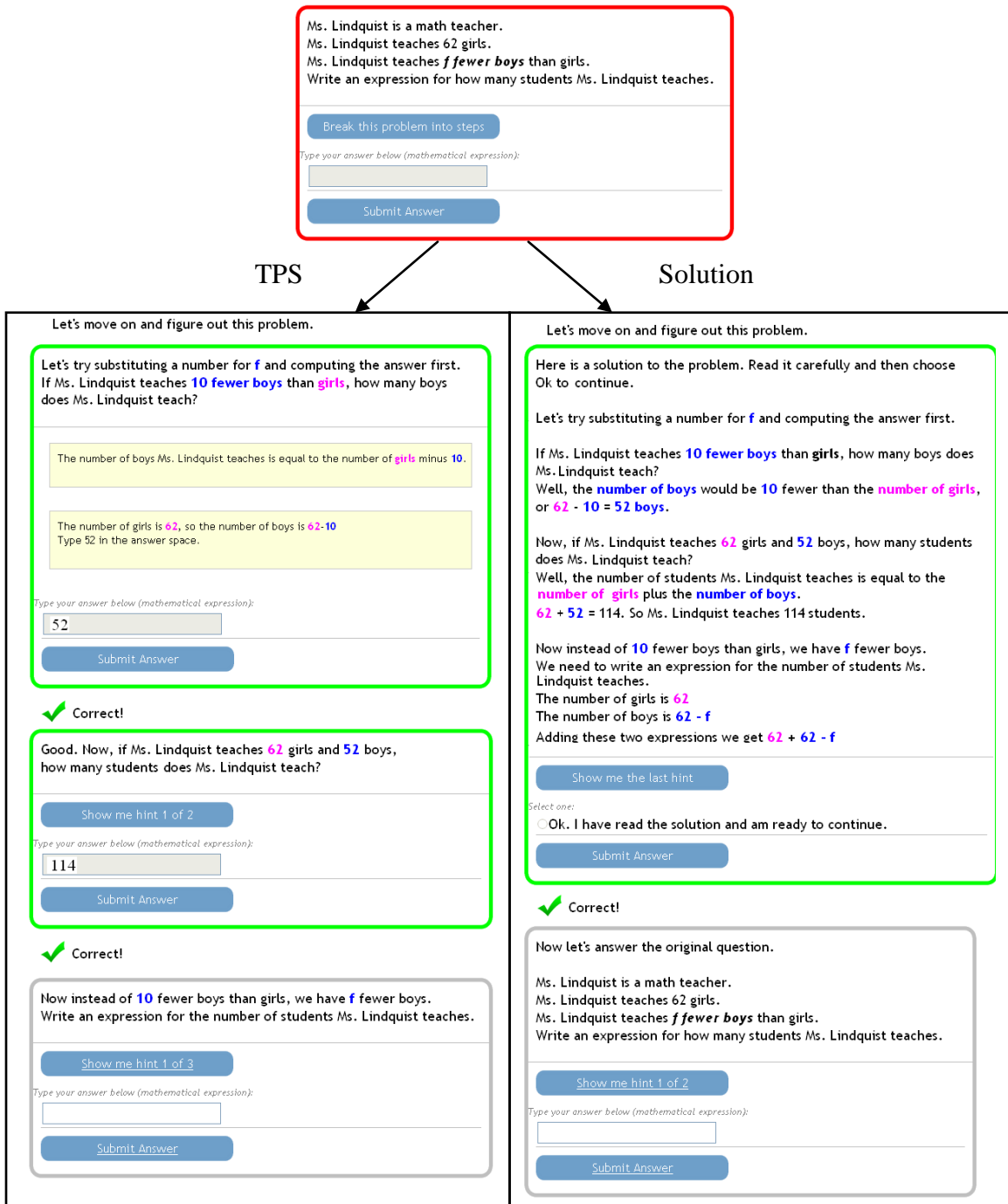


Figure 20: A symbolization problem shows the TPS approach on the left and the Solutions approach on the right.

9.2 Results

A gain score was calculated for each student by subtracting their pre-test scores from their post-test scores. The slope and intercept problem set contained three pre- and post-test problems and the symbolization problem set contained four pre- and post-test problems. For this reason, we calculated a z-score for each student's gain score on each problem set (to compare gain scores from distributions with different means). Thus, the transformed scores have a mean of zero and a standard deviation of one.

Overall, we found that there was significant learning in both problem sets ($p < 0.01$). The mean gain for the slope and intercept problem set was 15% and the mean gain for the symbolization problem set was 20%.

Students did significantly more problems with Solutions than with Tutored Problem Solving in both problem sets. For instance, in the symbolization problem set, students using Solutions did an average of 16.57 problems and students using TPS did an average of 11.59 problems ($t(82) = 16.66, p < 0.001$).

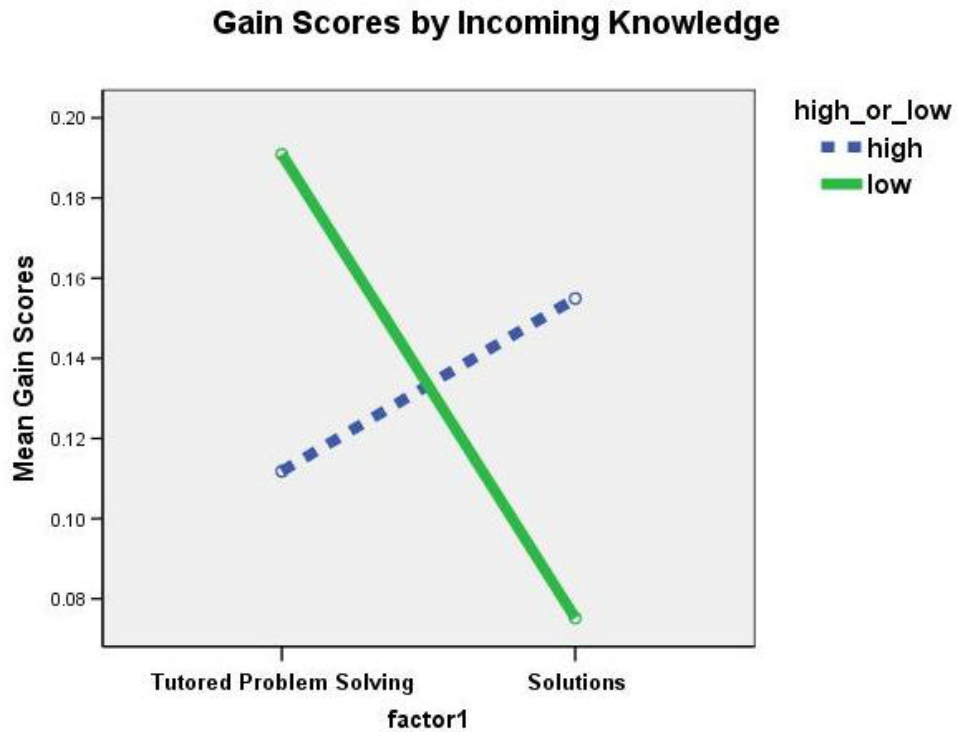
We used students' performance on a practice MCAS math test for 8th grade to categorize them as high proficiency or low proficiency. The practice MCAS test was given to the students as preparation for the MCAS test that they will take at the end of the school year. The average score on the practice MCAS for the students who participated in this study was 56% correct and the median was 57% correct. Therefore we placed students who scored greater than 56% on the practice test in the high math proficiency group and students who scored 56% or less in the low math proficiency group.

Our hypothesis was that highly proficient students would benefit more from Solutions and practice on more problems and that students with low proficiency would

benefit more from Tutored Problem Solving even though it was more time-consuming. Since every student participated in both conditions, we treated the problem sets as a repeated measure. The interaction between proficiency and condition ($F(1, 80) = 2.823, p = 0.097$) shows that highly proficient students learned more when they were shown Solutions than from doing the Tutored Problem Solving. Less proficient students learned more from Tutored Problem Solving than from seeing Solutions. Figure 21 shows the results of this analysis.

For more-proficient students, the difference between the two conditions was not significant ($p = 0.5$), however condition made a bigger difference for the less-proficient students. For these students, the mean gain for Tutored Problem Solving was 0.19 and the mean gain for the Worked Solutions was 0.07. This difference was approaching significance ($p = 0.12$).

In an attempt to explain why less-proficient students were not learning as much from getting more information as the more-proficient students, we decided to look more closely at how much time students spent reading through the solutions. Although we would expect less-proficient students to need to spend more time reading solutions, we found that they spent less than half as much time reading through solutions (mean = 12 seconds) than more-proficient students did (mean = 31 seconds). The difference between the time spent reading solutions was significant ($F(1, 30) = 14.801, p = 0.001$). We do not know if this was because of differences in reading ability or difficulty in focusing on the longer pieces of text.



	high_or_low	Mean	Std. Deviation	N
Tutored problem solving gain score	high	.1118	.25336	44
	low	.1909	.39730	34
	Total	.1463	.32412	78
Worked solutions gain score	high	.1549	.29913	44
	low	.0752	.25656	34
	Total	.1202	.28242	78

Figure 21: Highly proficient students appeared to learn more by seeing Solutions, and students with low proficiency learned more by doing TPS.

9.3 Conclusions

In (Razzaq, Heffernan, & Lindeman, 2007), we found evidence that choosing between giving or withholding information from students may depend on a student’s knowledge level. However, in that study we controlled for the number of problems and not for time spent. This study attempted to answer the question of when Tutored Problem Solving is worth the extra time that it takes and who benefits most from it.

This study showed that tutored problem solving was worth the extra time that it takes for students with low math proficiency. These students spent less time reading the Solutions although they arguably needed to spend more time. However, as expected, students with high math proficiency did not benefit from spending more time on tutored problem solving. They learned more from reading complete worked out solutions and solving more problems in the same amount of time. These students spent more time reading through Solutions. The difference between these groups of students could be due to motivation or reading ability. Further studies would be needed to determine which factor was more important.

10. Theory of the effectiveness of tutored problem solving

This research investigated ways of interacting with students with an intelligent tutoring system. The question was whether interactive tutored problem solving in an ITS where students must answer questions about each step of a problem is more effective than other methods of feedback. Randomized controlled experiments were used to answer this question. Experiments compared tutored problem solving to paper and pencil homework, hints on demand, worked examples, worked solutions and educational web pages.

10.1 Is interactive tutored problem solving to help students solve math problems more effective than less interactive forms of help?

The term “assistance dilemma” was coined by Koedinger and Alevan (2007). The assistance dilemma seeks to answer the question of how tutoring systems should balance *giving* and *withholding* information to optimize learning. Giving information can benefit students in that it is less time-consuming and students will make fewer errors (Klahr & Nigam, 2004). However, students may find it hard to stay focused and engaged. Withholding information can help students to stay focused and engaged while helping them to generate the information on their own (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). However, it is more time-consuming and students can make more errors from which it is difficult to recover. “The crux of the assistance dilemma is prescribing decision criteria (e.g., conditions and cutoff parameters) for when it is best to switch between information giving (more assistance) and information withholding (less assistance). This dilemma may be the fundamental open problem in learning and instructional science. (Koedinger & Alevan, 2007, p. 261)”

In this work, Tutored Problem Solving represents withholding information in an attempt to encourage students to construct knowledge themselves. Students must respond to questions and solve each step in order to proceed. They can get help and feedback on each step to help them solve the problem. On the other hand, presenting Worked Solutions represents giving information where students are given all of the information needed to solve the problem (including the answer). Students do not have to produce any response to the solution although they are asked to read and understand how to do the problem before moving on.

Table 10: Summary of randomized controlled experiments

Randomized controlled study	TPS	Paper & Pencil	Hints on Demand	Worked Solutions	Worked Examples	Web pages	Summary (**p<0.05, *p<0.1)
Mendicino, Razzaq & Heffernan, 2009	x	x	x				Large advantage to using the ASSISTment System with TPS and hints on demand over paper and pencil problem solving with an effect size of 0.6**.
Razzaq & Heffernan 2006	x		x				TPS showed an advantage over hints on demand with 0.3 effect size. For the more difficult problem, TPS outperformed hints with an effect size of 0.85**
Razzaq, Heffernan & Lindeman, 2007	x		x	x			For students with higher knowledge, worked solutions outperformed TPS with a 0.4* effect size. For students with lower knowledge, TPS outperformed worked solutions with a 0.27** effect size.
Shrestha et al, 2009	x				x		TPS outperformed worked examples with a 0.4** effect size. No subject-treatment interaction was detected.
Unpublished study comparing TPS to web pages	x					x	Overall, TPS outperformed viewing web pages with 0.2 effect size. For the more difficult topic, TPS outperformed viewing web pages with a 0.4 effect size
Razzaq & Heffernan, 2009	x			x			When we control for time on task, worked solutions outperformed TPS for high knowledge students with a 0.2 effect size. Low knowledge students learned more from TPS with a 0.48 effect size.

Table 10 summarizes the studies done in this dissertation to evaluate tutored problem solving and compare it to other forms of tutoring. Results of the studies done in this dissertation showed that tutored problem solving is indeed more helpful to *some* students but it is not always more helpful to others. This was especially apparent when we compared tutored problem solving to worked solutions, where this difference depended on the student’s math proficiency. We found that a student’s math proficiency

determined whether we should withhold information by presenting Tutored Problem Solving or give information by presenting a solution to the problem. This was true whether we controlled for time or for the number of problems. Students with high proficiency benefited from getting all of the information needed to solve a problem at once and students with low proficiency benefited more from getting information only on the step they were working on. We also found that students differed in how much time they spent reading Solutions. More-proficient students spent more than twice the amount of time reading Solutions than less-proficient students did. We do not know if this is due to a difference in focus, motivation or reading ability, but we believe that this difference may explain why less-proficient students did not learn as much from reading the Solutions.

We did not find this effect when we compared tutored problem solving to Worked Examples although students were given a complete solution of a similar problem rather than the same problem. It is possible that our results can be explained by cognitive load theory: perhaps the tutored problem solving reduces cognitive load even more than worked examples as students are walked through problems step by step and sub-goals are set for them. There may be a tradeoff in that students may lose the big picture by working on pieces of a problem at a time and are not asked to induce principles. However, sub-goal learning has been found to help guide problem solving by helping learners focus on the steps (Catrambone, 1998).

According to Sweller and Cooper (1985), “The use of worked example problems may redirect attention away from the problem goal and toward problem-state configurations and their associated moves.” Perhaps using worked examples as feedback

increased cognitive load as students tried to read the example and solve the problem at the same time.

Presenting educational web pages also did not show any advantage over tutored problem solving for high knowledge students. While Worked Examples showed a complete solution of a similar problem, the web pages usually focused on the principles needed to solve problems and gave examples that were not necessarily similar to the problem the student was trying to solve. Cognitive load theory may also explain why tutored problem solving was more effective than educational web pages.

Based on the results of the studies done in this dissertation, it appears that tutored problem solving in a tutoring system is an effective method of teaching low knowledge students to solve problems that they get wrong. The effectiveness of tutored problem solving is more significant when the student has low math proficiency or when the problem is more difficult. We can help students to learn more efficiently if they have high math proficiency by forgoing tutored problem solving and presenting more information. The information should pertain directly to the problem as in a worked solution as opposed to information about principles or similar examples. Figure 22 summarizes these results.

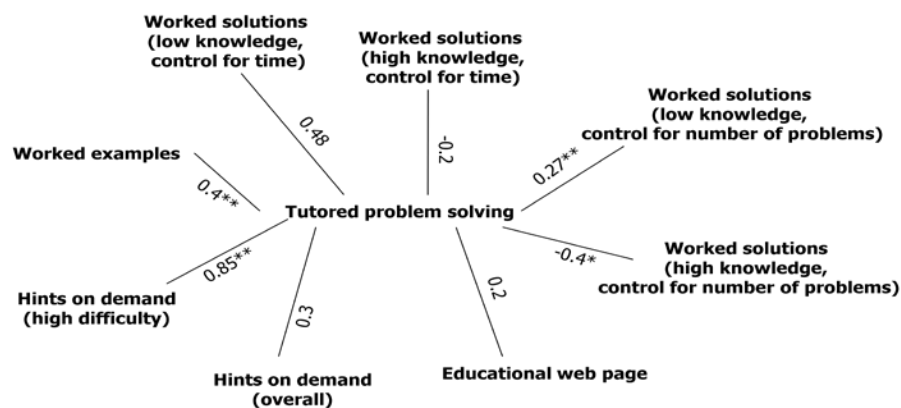


Figure 22: Effect sizes of tutored problem solving over other strategies (**p<0.05, *p<0.1)

11. Conclusions

This dissertation focused on student learning from an intelligent tutoring system. The ASSISTment System is a web-based tutoring system that assesses as it assists students while solving math problems. The system has over 3,000 student users who use it in class as part of their math instruction.

It is important to determine whether our method of tutoring students with tutored problem solving is effective. Numerous studies were designed and run to determine that students are learning when they use the ASSISTment System. There is a large advantage to using the ASSISTment System over paper and pencil problem solving with an effect size of 0.6. A majority of students believe that the ASSISTment System helps them to prepare for the Massachusetts state test.

When comparing tutored problem solving to five other methods of helping students to learn to solve problems that they got wrong, we found that tutored problem solving was more effective than paper and pencil homework, hints on demand, worked examples, and educational web pages. The effectiveness of tutored problem solving generally increases when the material is difficult or the student has low math proficiency. Tutored problem solving was more effective than worked solutions for students with low math proficiency, but not for students with high math proficiency. For students with high math proficiency, it is more effective and efficient to give them the whole solution at once.

11.1 Implications

We believe this research will aid the intelligent tutoring community in addressing the assistance dilemma. Of course, we do not claim that this study will definitively answer the assistance dilemma, but we believe it may take us a step closer to understanding the problem, helping us to optimize learning in an intelligent tutoring system by presenting the most effective and efficient approach to students determined by their knowledge level and the problem's difficulty. Students who have high proficiency would not have to waste time going through long problems step-by-step, causing them to become frustrated or bored. Students who have low proficiency may need to spend the extra time and obtain the help with focusing that tutored problem solving provides.

Results of this work could help inform the design of more effective tutoring strategies that adapt to the student to maximize learning results. This becomes more relevant and more pertinent as teachers are expected to cover more material to address all of the topics covered in standardized tests, instructional time becomes more precious and teachers want to know that the interventions that they are using in their classrooms are effective for students of varying abilities.

11.2 Directions for future work

11.2.1 Switching between strategies for individual students

In this dissertation, we have shown that students of low math proficiency benefited more from tutored problem solving than seeing worked solutions. As low knowledge students gain knowledge on a topic, can they gain more from worked solutions? Should high

knowledge students be given tutored problem solving if they begin to struggle? A dynamic approach would keep track of students' proficiency in each skill and switch to the appropriate strategy.

11.2.2 Examining other student characteristics

Thus far, much of my research in this area has been concentrated on adapting tutoring strategies to students with different degrees of background knowledge or performance levels. However, there are many other criteria that we can use to adapt instruction in learning technologies, such as determination/focus, reading ability, gender, affect, learning styles, and learning disabilities to name a few. Determining how these factors influence tutored problem solving in a tutoring system can help design a rich decision model for choosing the most effective tutoring strategies for individual students. This will help in the design of educational technologies that are able to deliver individualized, dynamic, effective learning content in real time.

11.2.3 Conceptual versus procedural knowledge

Kim et al. (2009) and Schwonke et al. (2007) found a difference between how effective tutored problem solving was compared to worked examples when the problem concerned procedural versus conceptual knowledge. It would be interesting to see if there is a similar effect for presenting worked solutions.

11.2.4 Ranking educational web content

Developers of educational technology often spend hundreds of hours developing content for their systems for a single topic. The World Wide Web has thousands of pages of educational content, some of them quite good. How can we utilize this wealth of effective

instructional content on the web, saving time and resources while helping students to learn? First, we need to be able to separate the wheat from the chaff. We have shown that students can learn from these web pages, but some educational web pages are better than others. It would be useful to be able to evaluate and rank the effectiveness of educational content on the web. This work will be useful to content developers as well as educators.

References

- Anderson, J. R. (1993). *Rules of the Mind*. Hillsdale: Erlbaum.
- Anderson, J. R., & Pelletier, R. (1991). A development system for model-tracing tutors. *Proceedings of the International Conference of the Learning Sciences*, (pp. 1-8).
- Anderson, J. R., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences* , 4 (2), 167-207.
- Anderson, J. R., Corbett, A., Koedinger, K., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences* , 4(2), 167-207.
- Arroyo, I., Beck, J., Woolf, B., Beal, C., & Schultz, K. (2000). Macroadaptating Animalwatch to gender and cognitive differences with respect to hint interactivity and symbolism. In G. Gauthier, K. VanLehn, & C. Frasson (Ed.), *Proceedings of the Conference on Intelligent Tutoring Systems*, (pp. 574-583).
- Bebell, D. (2005). *Technology promoting student excellence: An investigation of the 1st year of 1:1 computing in New Hampshire middle schools*. Technology and Assessment Study Collaborative.
- Bonifaz, A., & Zucker, A. A. (2004). *Lessons learned about providing laptops for all students*. Newton: Education Development Center.
- Burns, P. K., & Bozeman, W. C. (1981). Computer-assisted instruction and mathematics achievement: is there a relationship? *Educational Technology* , 21, 32-39.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General* , 127 (4), 355-376.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science* , 13, 145 – 182.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science* , 13, 145-182.
- Chi, M. T., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from tutoring. *Cognitive Science* , 25, 471-533.
- Chi, M. T., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from tutoring. *Cognitive Science* , 25, 471-533.

Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator* , 6 (11), 38-46.

Corbett, A., & Trask, H. (2000). Instructional interventions in computer-based tutoring: Differential impact on learning time and accuracy. *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 97-104). New York: ACM.

Core, M. G., Moore, J. D., & Zinn, C. (2003). The role of initiative in tutorial dialogue. *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics* , 67-74.

Core, M. G., Moore, J. D., & Zinn, C. (2003). The role of initiative in tutorial dialogue. *Proceedings of the tenth conference of the European chapter of the Association for Computational Linguistics* , 1, 67-74.

Craig, S. D., Driscoll, D. M., & Gholson, B. (2004). Constructing knowledge from dialog in an intelligent tutoring system: Interactive learning, vicarious learning and pedagogical agents . *Journal of Educational Multimedia and Hypermedia* (13), 163–183.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., & Campuzano, L. (2007). *Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.

Evens, M. W., & Michael, J. A. (2006). *One-on-one tutoring by humans and machines*. Mahway, NJ: Lawrence Erlbaum Associates.

Feng, M., Heffernan, N. T., & Koedinger, K. R. (2006). Predicting state test scores better with intelligent tutoring systems: developing metrics to measure assistance required. In Ikeda, Ashley, & Chan (Ed.), *Proceedings of the Eighth International Conference on Intelligent Tutoring Systems* (pp. 31-40). Berlin: Springer Verlag.

Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N. (2003). AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, & J. Kay (Ed.), *Proceedings of Artificial Intelligence in Education* (pp. 47-54). Amsterdam: IOS Press.

Graesser, A. C., Moreno, K., Marineau, J., Adcock, A., Olney, A., & Person, N. (2003). AutoTutor improves deep learning of computer literacy: Is it the dialog or the talking head? In U. Hoppe, F. Verdejo, & J. Kay (Eds.). *Proceedings of Artificial Intelligence in Education* (pp. 47-54). Amsterdam: IOS Press.

Graesser, A. C., Person, N., & Harter, D. (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education* , 12 (3), 257-279.

Heffernan, N. T., & Croteau, E. (2004). Web-based evaluations showing differential learning for tutorial strategies employed by the Ms. Lindquist tutor. In J. C. Lester, R. M. Vicari, & F. Paraguacu (Ed.), *Proceedings of the 7th Intelligent Tutoring Systems Conference* (pp. 491-500). Springer.

Heffernan, N. T., & Koedinger, K. R. (2000). Intelligent tutoring systems are missing the tutor: Building a more strategic dialog-based tutor. In C. P. Rose, & R. Freedman (Ed.), *Building Dialogue Systems for Tutorial Applications, Papers of the 2000 AAAI Fall Symposium*, (pp. 14-19). Menlo Park.

Heffernan, N. T., Koedinger, K. R., & Razzaq, L. (2008). Expanding the model-tracing architecture: A 3rd generation intelligent tutor for Algebra symbolization. *The International Journal of Artificial Intelligence in Education* , 18 (2), 153-178.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology* , 93 (3), 579–588.

Katz, S., Allbritton, D., & Connelly, J. (2003). Going beyond the problem given: How human tutors use post-solution discussions to support transfer. *The International Journal of Artificial Intelligence in Education* (13), 79-116.

Kim, R., Weitz, R., Heffernan, N., & Krach, N. (2009). Tutored Problem Solving vs. “Pure”: Worked Examples. In N. A. Taatgen, & H. van Rijn (Ed.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science* , 15 (10), 661–667.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review* , 19, 239-264.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with Cognitive Tutors. *Educational Psychology Review* , 19, 239-264.

Koedinger, K. R., Aleven, V., Heffernan, N. T., McLaren, B., & Hockenberry, M. (2004). Opening the door to non-programmers: Authoring intelligent tutor behavior by demonstration. *Proceedings of the 7th Intelligent Tutoring Systems Conference*, (pp. 162-173).

Koedinger, K., Anderson, J., Hadley, W., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education* , 8, 30-43.

Koedinger, K., Corbett, A., Ritter, S., & Shapiro, L. (2000). *Carnegie Learning's Cognitive Tutor Summary Research Results*. Retrieved November 12, 2007, from

carnegielearning.com:

http://www.carnegielearning.com/research/research_reports/CMU_research_results.pdf

Kulik, C. C., & Kulik, J. A. (1986). Effectiveness of computer-based education in colleges. *AEDS Journal* , 19, 81-108.

Kulik, C. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior* , 7 (1-2), 75-94.

Kulik, J. A., Bangert-Drowns, R., & Williams, G. (1983). Effects of Computer-Based Teaching on Secondary School Students. *Journal of Educational Psychology* , 75 (1), 19-26.

Kulik, J. A., Kulik, C. C., & Bangert-Drowns, R. L. (1985). Effectiveness of computer-based education in elementary schools. *Computers in Human Behavior* , 1 (1), 59-74.

Kulik, J. (1994). Meta-analytic studies of findings on computer-based instruction. In E. L. Baker, & H. F. O'Neil (Eds.), *Technology assessment in education and training* (pp. 9-33). Hillsdale: Lawrence Erlbaum.

Larkin, J., McDermott, J., Simon, D., & Simon, H. (1980). Expert and novice performance in solving physics problems. *Science* , 208, 1335-1342.

Mathan, S., & Koedinger, K. R. (2003). Recasting the Feedback Debate: Benefits of Tutoring Error Detection and Correction Skills. In U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Proceedings of the Conference on Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies* (pp. 39-46). Amsterdam: IOS Press.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction* (7), 197-244.

McLaren, B. M., Lim, S., & Koedinger, K. R. (2008). When and How Often Should Worked Examples be Given to Students? New Results and a Summary of the Current State of Research. In B. C. Love, K. McRae, & V. M. Sloutsky (Ed.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2176-2181). Austin, TX: Cognitive Science Society.

Mendicino, M., Razzaq, L., & Heffernan, N. (2009). Comparison of traditional homework with computer supported homework. *Journal of Research on Technology in Education* , 41 (3), 331-359.

Merrill, D., Reiser, B., Merrill, S., & Landes, S. (1995). Tutoring: guided learning by doing. *Cognition and Instruction* , 13 (3), 315-372.

Merrill, D., Reiser, B., Ranney, M., & Trafton, J. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *Journal of the Learning Sciences* , 2 (3), 277-305.

- Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor Algebra I on student knowledge and attitude*. Carnegie Learning, Inc.
- Murray, C., & VanLehn, K. (2006). A Comparison of Decision-Theoretic, Fixed-Policy and Random Tutorial Action Selection. In Ikeda, Ashley & Chan (Eds.), *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (pp. 116-123). Berlin: Springer Verlag.
- Murray, R. C., & VanLehn, K. (2005). Effects of Dissuading Unnecessary Help Requests While Providing Proactive Help. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Ed.), *Proceedings of the 12th International Conference on Artificial Intelligence in Education*, (pp. 887-889).
- Murray, T. (1999). Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education*, 10, 98-129.
- Murray, T., Blessing, S., & Ainsworth, S. (2003). *Authoring Tools for Advanced Technology Learning Environments*. Netherlands: Kluwer.
- Nuzzo-Jones, G., Walonoski, J. A., Heffernan, N. T., & Livak, T. (2005). The eXtensible tutor architecture: A new foundation for ITS. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Ed.), *Proceedings of the 12th Conference on Artificial Intelligence in Education*, (pp. 902-904).
- Ohlsson, S. (1986). Some principles for intelligent tutoring. *Instructional Science* (17), 281-307.
- Razzaq, Heffernan, N., Feng, M., & Pardos, Z. (2007). Developing Fine-Grained Transfer Models in the ASSISTment System. *Journal of Technology, Instruction, Cognition, and Learning*, 5 (3), 289-304.
- Razzaq, L., & Heffernan, N. T. (2006). Scaffolding vs. hints in the ASSISTment System. In Ikeda, Ashley, & Chan (Ed.), *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (pp. 635-644). Berlin: Springer.
- Razzaq, L., & Heffernan, N. (2009). To tutor or not to tutor: That is the question. In Dimitrova, Mizoguchi, d. Boulay, & Graesser (Ed.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education*, (pp. 457-464).
- Razzaq, L., & Heffernan, N. (2004). Tutorial dialog in an equation solving intelligent tutoring system. In Lester, Vicari, & Parguacu (Ed.), *Proceedings of the 7th Intelligent Tutoring Systems Conference*, (pp. 851-853).
- Razzaq, L., Heffernan, N. T., & Lindeman, R. W. (2007). What level of tutor interaction is best? In R. Luckin, K. R. Koedinger, & J. Greer (Ed.), *Proceedings of the 13th Conference on Artificial Intelligence in Education* (pp. 222-229). Amsterdam: IOS Press.

Razzaq, L., Heffernan, N. T., & Lindeman, R. W. (2007). What level of tutor interaction is best? In R. Luckin, K. R. Koedinger, & J. Greer (Ed.), *Proceedings of the 13th Conference on Artificial Intelligence in Education* (pp. 222-229). Amsterdam: IOS Press.

Razzaq, L., Heffernan, N. T., Koedinger, K. R., Feng, M., Nuzzo-Jones, G., Junker, B., et al. (2007). Blending Assessment and Instructional Assistance. In N. Nedjah, L. d. Macedo Mourelle, M. Neto Borges, & N. Nunesde de Almeida (Eds.), *Intelligent Educational Machines* (pp. 23-49). Berlin/Heidelberg: Springer.

Razzaq, L., Heffernan, N. T., Koedinger, K. R., Feng, M., Nuzzo-Jones, G., Junker, B., et al. (2007). Blending Assessment and Instructional Assistance. In N. Nedjah, L. d. Macedo Mourelle, M. Neto Borges, & N. Nunesde de Almeida, *Intelligent Educational Machines* (pp. 23-49). Berlin/Heidelberg: Springer.

Razzaq, L., Mendicino, M., & Heffernan, N. (2008). Comparing classroom problem-solving with no feedback to web-based homework assistance. In Woolf, Aimeur, Nkambou, & Lajoie (Ed.), *Proceeding of the 9th International Conference on Intelligent Tutoring Systems* (pp. 426-437). Berlin: Springer-Verlag.

Razzaq, L., Parvarczki, J., Almeida, S. F., Vartak, M., Feng, M., Heffernan, N. T., et al. (2009). The ASSISTment builder: Supporting the Life-cycle of ITS Content Creation. *IEEE Transactions on Learning Technologies* , 2 (2), 157-166.

Rose, C., Jordan, P., Ringenberg, M., Siler, S., VanLehn, K., & Weinstein, A. (2001). Interactive conceptual tutoring in Atlas-Andes. In J. D. Moore, C. L. Redfield, & W. L. Johnson (Ed.), *Proceedings of the International Conference on Artificial Intelligence in Education*, (pp. 256-266).

Salden, R., Aleven, V., Renkl, A., & Schwonke, R. (2008). Worked examples and tutored problem solving: redundant or synergistic forms of support? In C. Schunn (Ed.), *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 659-664). New York: Lawrence Erlbaum.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction* , 16 (4), 475-522.

Schwonke, R., Wittwer, J., Aleven, V., Salden, R., Krieg, C., & Renkl, A. (2007). Can tutored problem solving benefit from faded worked-out examples? In S. Vosniadou, D. Kayser, & A. Protopapas (Ed.), *Proceedings of European Cognitive Science Conference 2007* (pp. 59-64). New York: Lawrence Erlbaum.

Shrestha, P., Maharjan, A., Wei, X., Razzaq, L., Heffernan, N. T., & Heffernan, C. (2009). Are Worked Examples an Effective Feedback Mechanism During Problem Solving? In N. A. Taatgen, & H. van Rijn (Ed.), *Proceedings of the 31th Annual Conference of the Cognitive Society*.

- Silvernail, D. L., & Lane, D. M. (2004). *The impact of Maine's one-to-one laptop program on middle school teachers and students*. Maine Education Policy Research Institute. Gorham: University of Southern Maine.
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction* , 2, 59-89.
- Timms, M. (2007). Using Item Response Theory (IRT) to select hints in an ITS. In Luckin, & Koedinger (Ed.), *Proceedings of the 13th Conference on Artificial Intelligence in Education*, (pp. 213-221).
- Turner, T. E., Macasek, M. A., Nuzzo-Jones, G., Heffernan, N. T., & Koedinger, K. (2005). The ASSISTment Builder: A rapid development tool for ITS. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Ed.), *Proceedings of the 12th Artificial Intelligence In Education*, (pp. 929-931).
- VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rose, C. P. (2005). When is reading just as effective as one-on-one interactive human tutoring? *In Proceedings of the 27th Annual Meeting of the Cognitive Science Society* (pp. 2259-2264). Mahwah: Erlbaum.
- VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., et al. (2005). The Andes physics tutoring system: Lessons Learned. *International Journal of Artificial Intelligence and Education* , 15 (3), 1-47.
- VanLehn, K., Siler, S., Murray, C., & Baggett, W. B. (1998). What makes a tutorial event effective? In M. A. Gernsbacher, & S. J. Derry (Ed.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, (pp. 1084–1089).
- VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction* , 21 (3), 209-249.
- Warnakulasooriya, R., & Pritchard, D. E. (2005). Learning and problem-solving transfer between physics problems using web-based homework tutor. *Conference Proceedings: EdMedia -World Conference on Educational Multimedia, Hypermedia & Telecommunications*, (pp. 2976-2983).

Appendix A: Materials for ASSISTments vs. Paper and Pencil Study

Appendix A.1 Pretests/Post-tests

Name _____

Teacher _____

1.

Turf Coverage	
Pounds	Square Yards of Coverage
6	200
10	300
14	400
18	500

The table above shows the number of pounds of turf needed to cover a given area.

Based on the pattern in the table, how many pounds of turf are needed to cover 800 square yards?

2.

$$2X + 2 = 14$$

What value of X makes the equation shown above true?

- A. X = 6
- B. X = 8
- C. X = 10
- D. X = 12

3. The radius of a circle is 18 inches. What is the diameter of the circle?

4. Mrs. Chipps wrote five numbers on the white board in her room.

After class, one of the numbers was erased. The four numbers left are shown below.

20 32 44 12 ?

If the median of the five numbers that Mrs. Chipps wrote on the board was 20, which of the following could be true?

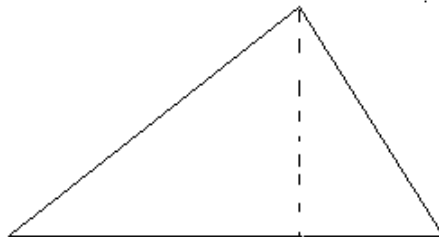
- A. The number that was erased was greater than 44
- B. The mode of the five numbers Mrs. Chipps wrote on the board was 31
- C. The number that was erased was less than or equal to 20
- D. The mean of the five numbers Mrs. Chipps wrote on the board was 56

5. Mr. Lamb drew an equilateral triangle.

Which of the following statements is true about the triangle?

- A. At least two angles are obtuse
- B. At least one angle measures 90 degrees
- C. All of the angles are less than 90 degrees
- D. All of the angles have different measurements

6



What is the area of the triangle shown above?

- A. 220 cm^2
- B. 156 cm^2
- C. 125 cm^2
- D. 225 cm^2

7.

Input	1	2	3	4	5	6	7
Output	6	10	14	18	22	26	30

Shelby created the input-output table shown above.

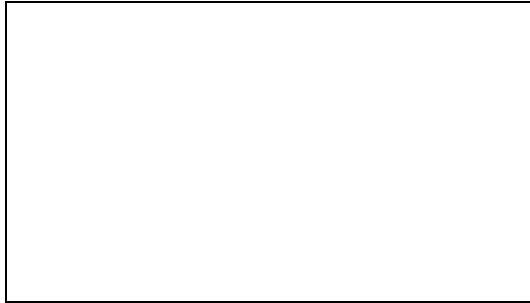
Which of the following rules is true for all values of Shelby's input-output table?

- A. Input + 5 = output
- B. Input times 5 = output
- C. (Input times 4) + 2
- D. (Input times 4) + 3

8. Joe is 22 years older than Bob. If Joe is 43 years old now, how old is Bob?

- A. 12 years old
- B. 18 years old
- C. 65 years old
- D. 21 years old

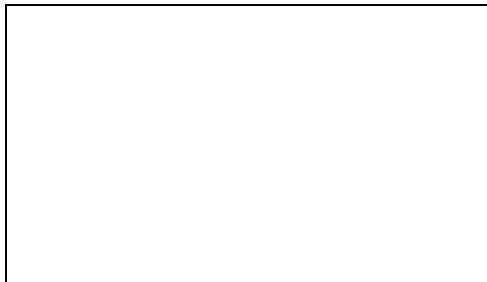
9.



On the coordinate grid above, Which point is located at (7,5)?

- A.
- B.
- C.
- D.

10.



What are the coordinates of point C?

Number Sense Pre Test

Name _____

Teacher _____

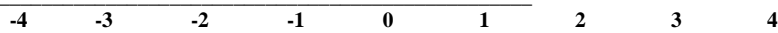
1. Which of the following is closest to the product of $397.8 * 10.3$?

- A. 3,000
 - B. 30,000
 - C. 400
 - D. 4,000
-

2. Which of the following shows the numbers in order from least to greatest?

- A. 0.452, 0.51, 0.432
 - B. 0.452, 0.432, 0.51
 - C. 0.432, 0.51, 0.452
 - D. 0.432, 0.452, 0.51
-

3.



Jeffrey is plotting points on the number line above.
Between which two numbers should Jeffery plot $-3 \frac{1}{2}$?

4. What is $15/60$ as a percent?

5. Write $10/25$ as a decimal.

6. What is the value of the following expression?

$$7 * 6 + 3$$

7. Destinee has a total of 12 fish in her aquarium.

Exactly 9 of the fish are gold fish.

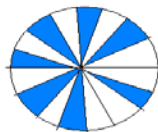
What percent of the fish in the aquarium are gold fish?

- A. 70%
 - B. 55%
 - C. 75%
 - D. 25%
-

8. David made the circle shown below using gray and white triangles.

What fractional part of the whole design is made up of gray triangles?

Write your answer as a fraction.



9.

Kendra is going on vacation. She packed the following clothes in her suitcase. How many different outfit combinations will Kendra have to choose from during vacation?

Suitcase for vacation

Tops	Pants
T-Shirt Sweatshirt	Khaki pants Sweat pants Jeans

10. What is the distance between point C and point D on the number shown below?



Appendix A.2: Homework Problem Sets

Homework/Mix Problems

Name: _____ Teacher: _____ Date: _____

1)

Fertilizer Coverage

Pounds	Square Yards of Coverage
4	100
8	200
12	300
16	400

The table above shows the number of pounds of fertilizer needed to cover a given area. Based on the pattern in the table, how many pounds of fertilizer are needed to cover 600 square yards?

2.)

$$2x + 2 = 10$$

What value of x makes the equation shown above true?



A. $x = 4$



B. $x = 6$



C. $x = 8$



D. $x = 12$

3.)

The radius of a circle is 14 inches. What is the diameter of the circle?

4.)

Mr. Young wrote five numbers on the board in his classroom.

After class, one of the numbers was erased. Four of the five numbers are shown below.

18 25 30 17 _?_

If the median of the five numbers that Mr. Young wrote on the board was 18, which of the following could be true?



A. The number that was erased was greater than 30.



B. The mode of the five numbers Mr. Young wrote on the board was 24.



C. The mean of the five numbers Mr. Young wrote on the board was 22.6.



D. The number that was erased was less than or equal to 18.

5.)

Mr. Donato drew an equilateral triangle.

Which of the following statements is true about the triangle?



A. At least one angle is obtuse.



B. All of the angles are acute.

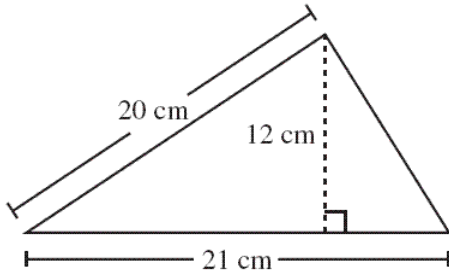


C. At least one angle measures 90 degrees.



D. All of the angles have different measurements.

6.)



What is the area of the triangle shown above?

- A. 126 cm²
- B. 210 cm²
- C. 252 cm²
- D. 420 cm²

7.)

Input	1	2	3	4	5	6	7
Output	4	6	8	10	12	14	16

Bridget created the input-output table shown above.

Which of the following rules is true for all values in Bridget's input-output table?

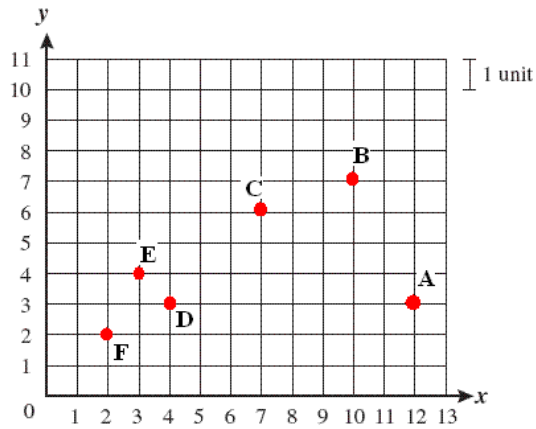
- A. Input + 3 = Output
- B. Input * 3 = Output
- C. (Input * 2) + 1 = Output
- D. (Input * 2) + 2 = Output

8.)

Sam is 37 years older than Dennis. If Sam is 55 years old now, how old is Dennis?

- A. 12 years old
- B. 18 years old
- C. 28 years old
- D. 92 years old

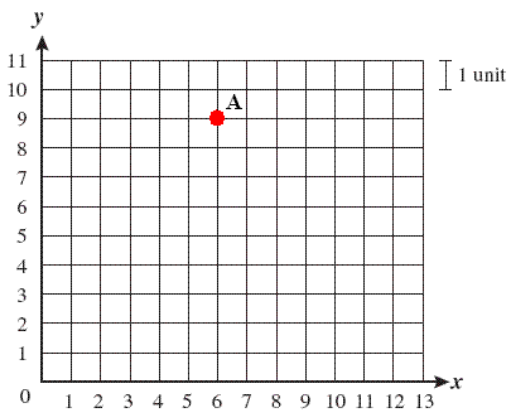
9.)



On the coordinate grid above, which point is plotted at (4, 3)?

- A.
- B.
- C.
- D.
- E.
- F.

10.)



What are the coordinates of Point A?

- (6, 10)
- (5, 9)
- (9, 6)
- (6, 9)

Homework/Number Sense Problems

Name: _____ Teacher: _____ Date: _____

1.)

Which of the following is closest to the product $298.7 * 10.1$?



A. 300



B. 2,000



C. 3,000



D. 20,000

2.)

Which of the following shows the numbers in order from least to greatest?



A. 0.765, 0.82, 0.791



B. 0.765, 0.791, 0.82

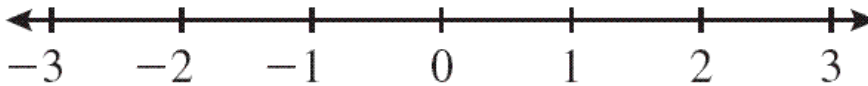


C. 0.791, 0.82, 0.765



D. 0.791, 0.765, 0.82

3.)



Marta is plotting points on the number line above.

Between which two numbers should Marta plot $-2\frac{1}{2}$?



A. 1 and 2



B. 2 and 3



C. -2 and -1



D. -3 and -2

4.)

Write $12/30$ as a percent.

5.)

Write $15/25$ as a decimal.

6.)

What is the value of the following expression?

$$3 + 6 * 4$$

7.)

Judith has a total of 8 fish in her aquarium.

Exactly 6 of the fish are guppies.

What percent of the fish in the aquarium are guppies?



A. 48%



B. 60%

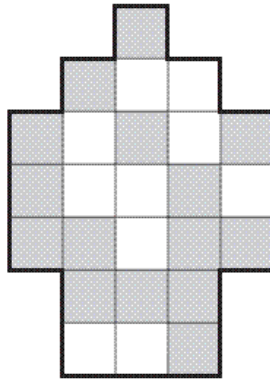




C. 68%



D. 75%

8.)



Key	
	represents 1 gray tile
	represents 1 white tile

Shing made the design shown above using gray square tiles and white square tiles.
 What fractional part of the whole design is made up of **gray** tiles? Write your answer as a fraction.

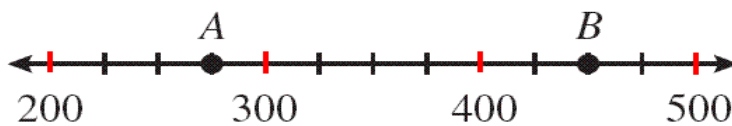
9.)

Salad Ingredients

Lettuce	Vegetable	Dressing
Iceberg	Carrot	Ranch
Romaine	Celery	Italian
	Broccoli	Caesar
	Cauliflower	Vinaigrette
		French

Rae is making a salad. The choices for the ingredients are shown in the chart above.
 What is the total number of different salads she can make using one lettuce, one vegetable, and one dressing?

10.)



What is the distance between point A and point B on the number line shown above?

Appendix A.3: Sample ASSISTment Problems

1.) "2005_5_gr6" (Problem ID: 12330) [MA - 2005 - SPRING - 5]

Which of the following is closest to the product $298.7 * 10.1$?

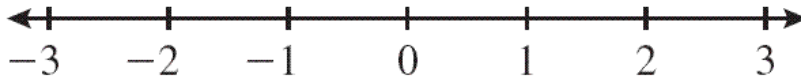
- A. 300
- B. 2,000
- C. 3,000
- D. 20,000

2.) "2005_23_gr6" (Problem ID: 12395) [MA - 2005 - SPRING - 23]

Which of the following shows the numbers in order from least to greatest?

- A. 0.765, 0.82, 0.791
- B. 0.765, 0.791, 0.82
- C. 0.791, 0.82, 0.765
- D. 0.791, 0.765, 0.82

3.) "2005_20_gr6" (Problem ID: 23332) [MA - 2005 - SPRING - 20]



Marta is plotting points on the number line above.

Between which two numbers should Marta plot $-2\frac{1}{2}$?

- A. 1 and 2
- B. 2 and 3
- C. -2 and -1
- D. -3 and -2

2.) "2005_6_gr6" (Problem ID: 12341) [MA - 2005 - SPRING - 6]

$$2x + 2 = 10$$

What value of x makes the equation shown above true?

- A. $x = 4$
- B. $x = 6$
- C. $x = 8$
- D. $x = 12$

3.) "2005_12_gr6" (Problem ID: 12361) [MA - 2005 - SPRING - 12]

The radius of a circle is 14 inches. What is the diameter of the circle?

4.) "2005_15_gr6" (Problem ID: 12366)

Mr. Young wrote five numbers on the board in his classroom.

After class, one of the numbers was erased. Four of the five numbers are shown below.

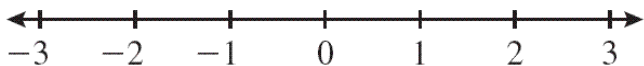
18 25 30 17 _?_

If the median of the five numbers that Mr. Young wrote on the board was 18, which of the following could be true?

- A. The number that was erased was greater than 30.
- B. The mode of the five numbers Mr. Young wrote on the board was 24.
- C. The mean of the five numbers Mr. Young wrote on the board was 22.6.
- D. The number that was erased was less than or equal to 18.

Appendix A.4: Sample Hints

3.)



Marta is plotting points on the number line above.

Between which two numbers should Marta plot $-2\frac{1}{2}$?

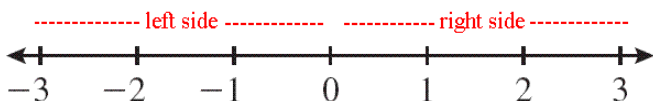
Answers:

A. 1 and 2

B. 2 and 3

C. -2 and -1

D. -3 and -2



$-2\frac{1}{2}$ should be $2\frac{1}{2}$ units away from zero.

What side of zero should $-2\frac{1}{2}$ be on?

Answers:

the left side

the right side

Hint 1:

The numbers to the right of zero on the number line are positive and are greater than zero.

The numbers to the left of zero on the number line are negative and are less than zero.

Hint 2:

Marta should plot $-2\frac{1}{2}$ on the left side of zero because it's negative.

Between which two numbers should Marta plot $-2\frac{1}{2}$?

Answers: (Interface Type: RADIO_BUTTON)

A. 1 and 2

B. 2 and 3

C. -2 and -1

D. -3 and -2

Hint 1:

Marta should plot $-2\frac{1}{2}$ on the left side of zero, between two negative numbers.

Hint 2:

$-2\frac{1}{2}$ should be $2\frac{1}{2}$ units away from zero. Is that to the left of -2 or to the right of -2?

Hint 3:

$-2\frac{1}{2}$ should be plotted between -3 and -2. Choose D.

4.) Write $\frac{12}{30}$ as a percent.

Answers:

40%

40

First, it will help to reduce the fraction to tenths. What is $\frac{12}{30}$ reduced?

Answers:

$\frac{3}{10}$

$\frac{4}{10}$

$\frac{4}{5}$

$\frac{6}{12}$

Hint 1:

What is a common factor between 12 and 30, that will give you 10 in the denominator?

Hint 2:

The common factor is 3. Divide 12 and 30 by 3.

Hint 3:

$$12/30 = 4/10$$

Hint 4:

Choose 4/10

Good. What is 4/10 in percent?

Answers:

0.4%

0.04%

4%

40%

Hint 1:

There are several ways to change a fraction to a percent.

One way is to divide the numerator by the denominator and move the decimal point 2 places to the right.

Hint 2:

What is 4 divided by 10?

Hint 3:

40 %

4/10 = 0.4. Now move the decimal point 2 places to the right.

Hint 4:

0.4 = 40%. Choose 40%

5. Write 15/25 as a decimal.

Answers: (Interface Type: ALGEBRA_FIELD)

0.6

First, it will help to reduce the fraction. What is 15/25 reduced?

Answers:

3/5

Hint 1:

What is the greatest common factor between 15 and 25?

Hint 2:

The greatest common factor is 5. Divide 15 and 25 by 5.

Hint 3:

$$15/25 = 3/5$$

Hint 4:

Type in 3/5

Good. What is 3/5 in decimal?

Answers:

0.6

Hint 1:

One way to change a fraction to a decimal is to divide the numerator by the denominator, but it is easier to convert this fraction to tenths first.

Hint 2:

$3/5 * 2/2 = 6/10$. To divide by 10, move the decimal point one place to the left. (Since there is no decimal point, think of 6 as 6.0 and move the decimal point one place to the left.)

Hint 3:

$$6/10 = 0.6 \text{ or } 0.60$$

Hint 4:

Type in 0.6

Appendix B Materials for Experiment on Tutored Problem Solving vs. Worked Out Solutions

Appendix B.1 Pretest

1) Assistent #9796 "9796 - pretest_y_intercept"

What is the y-intercept of the graph represented by the equation below?

$$y = \frac{4}{5}x - 2$$

Scaffold:

This was a **PRETEST** question, so we will not tell you if you got it right or wrong yet.

We will come back to this problem later and give you feedback on your answer.

Please click Ok to proceed.

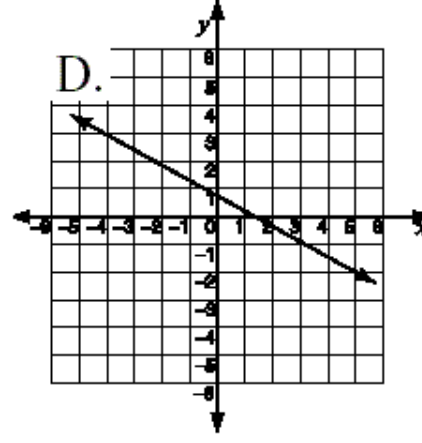
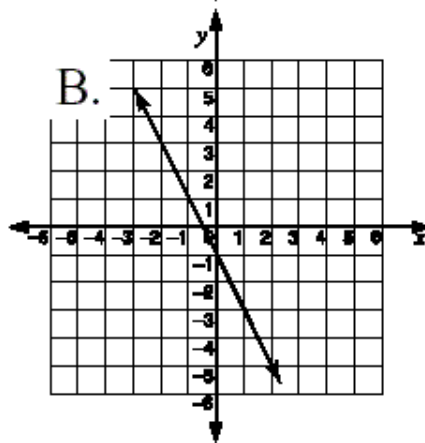
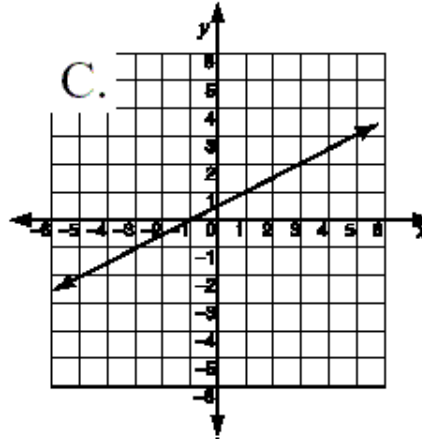
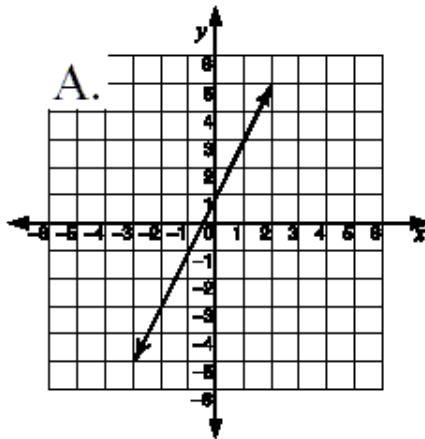
Multiple choice:

Ok

2) Assistent #9798 "9798 - pretest_3_2001"

Which graph above contains the points in the table below?

x	y
-2	-3
-1	-1
1	3



Scaffold:

This was a **PRETEST** question, so we will not tell you if you got it right or wrong yet.

We will come back to this problem later and give you feedback on your answer.

Please click Ok to proceed.

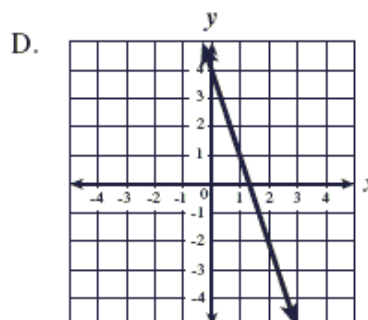
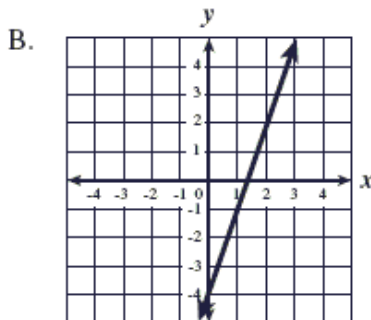
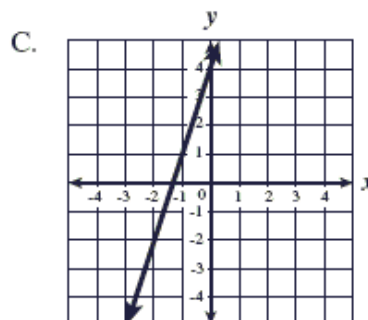
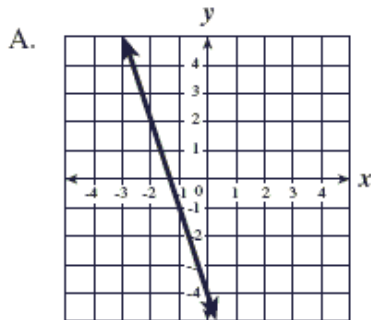
Multiple choice:

Ok

Appendix B.2 Tutored Problem Solving Condition

3) Assistentment #2724 "2724 - Item14_2004_graph_equation"

Which graph below best represents $y = -3x + 4$?



Multiple choice:

A.

B.

C.

D.

Scaffold:

Will the graph that best represents this equation have a positive slope or a negative slope?

Multiple choice:

Negative

Positive

The slope is -3, so it is negative.

Hints:

- The equation of a line can be written as follows:

$$y = mx + b$$

where **m** is the **slope**

and **b** is the **y-intercept**.

According to the equation $y = -3x + 4$, what is the slope of this line?

- The slope of the line would be -3
- Is the slope positive or negative?
- The answer is Negative.

Scaffold:

Right. The slope of the line will be negative.

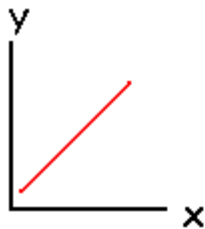
Which graph(s) show a line with a negative slope?

Multiple choice:

- A.
- A. and B.
- A. and D.
- B.
- B. and C.
- C.
- C. and D.
- D.

Hints:

- If the line is sloping upward from left to right (the direction that you read), the slope is positive.
- If the line is sloping downward from left to right, the slope is negative.



The line is sloping upward (left to right), the slope is positive.



The line is sloping downward (left to right), the slope is negative.

- Which graph(s) show lines that slope downward from left to right?
- The answer is Graphs A. and D.

Scaffold:

Good. Now we know that the correct answer will be graph A or graph D.

According to the equation, $y = -3x + 4$, what is the y-intercept of the line ?

Algebra:

✓ 4

Hints:

- Recall that the equation of a line can be written as follows:

$$y = mx + b$$

where **m** is the **slope**

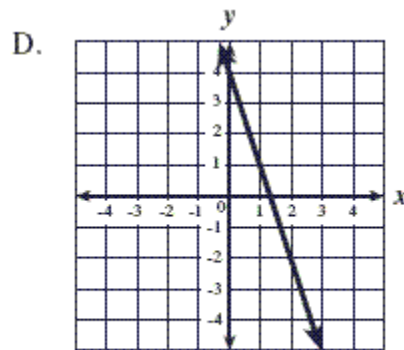
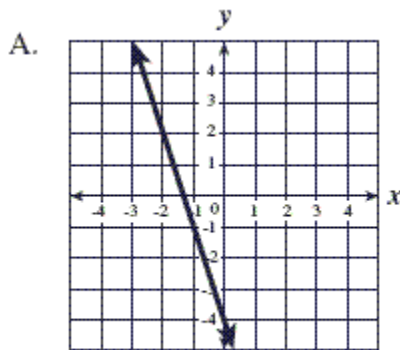
and **b** is the **y-intercept**.

According to the equation $y = -3x + 4$, what is the y-intercept of this line?

- The y-intercept of this line is 4.

Scaffold:

Now which graph shows a line with a y-intercept of 4?



Multiple choice:

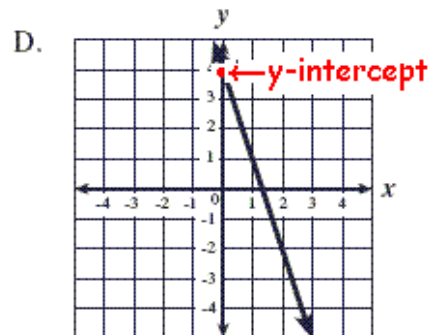
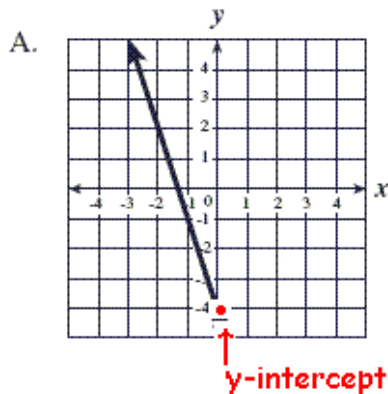
✗ A.

The y-intercept in this graph is -4.

✓ D.

Hints:

- The y-intercept of the line is the value of y at the point where the line crosses the y-axis.



- Which graph shows a y-intercept of 4?

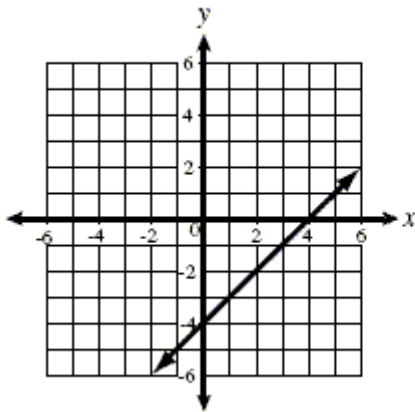
- The answer is D. Please choose D.

4) Assisment #2719 "2719 - Item3_2001_points_on_graph_morph"

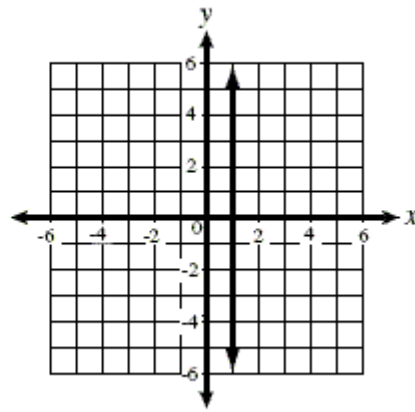
Which graph contains the points given in the table below?

x	y
-4	1
0	1
4	1

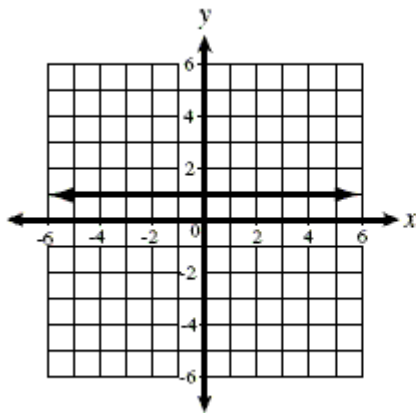
A.



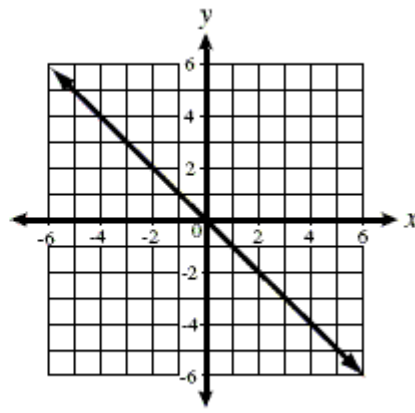
C.



B.



D.

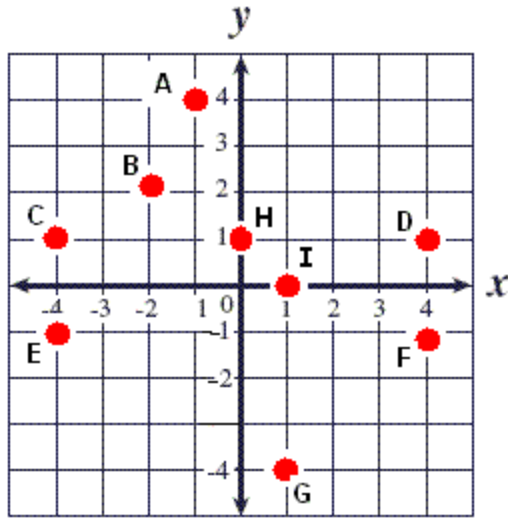


Multiple choice:

- A.
- B.
- C.
- D.

Scaffold:

Let's try plotting the first point in the table. Which point in the graph below has coordinates (-4, 1)?

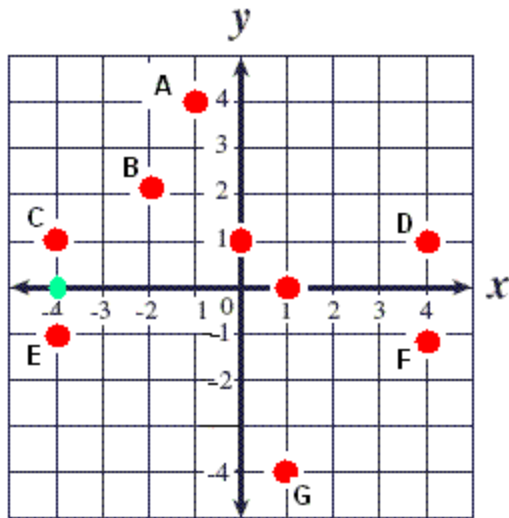


Multiple choice:

- A.
- B.
- C.
- D.
- E.
- F.
- G.

Hints:

- The x-coordinate is -4. Find -4 on the x-axis.
- The y-coordinate is 1. Find 1 on the y-axis.
- Find the point where -4 on the x-axis and 1 on the y-axis meet. Which point is it?

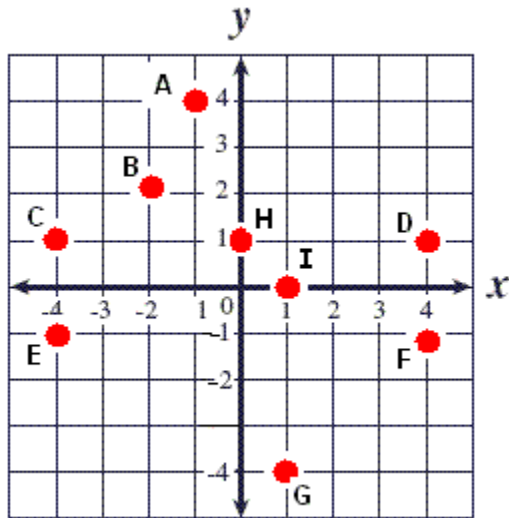


- The correct point is C

Scaffold:

Good. Now let's plot the next point in the table.

Which point in the graph below has coordinates (0, 1)?



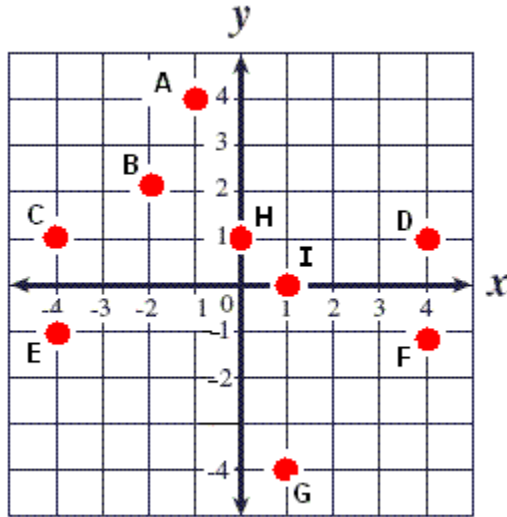
Multiple choice:

- A.
- B.
- C.
- D.
- E.
- F.
- G.

- ✓ H.
- ✗ I.

Hints:

- The x-coordinate is 0. Find 0 on the x-axis.
- The y-coordinate is 1. Find 1 on the y-axis.
- Find the point where 0 on the x-axis and 1 on the y-axis meet. Which point is it?

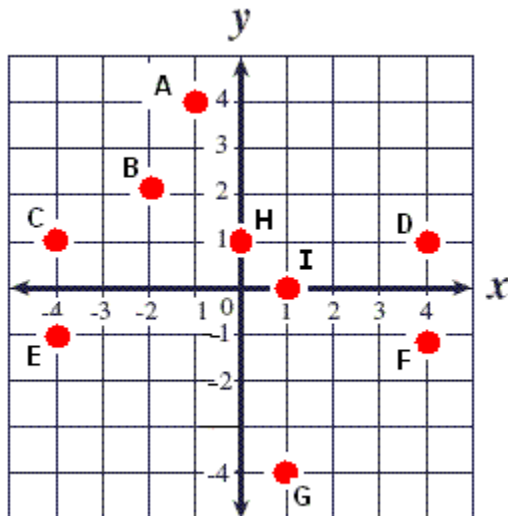


- The answer is H. Please choose H.

Scaffold:

Good. Now let's plot the last point.

Which point in the graph below has coordinates (4, 1)?

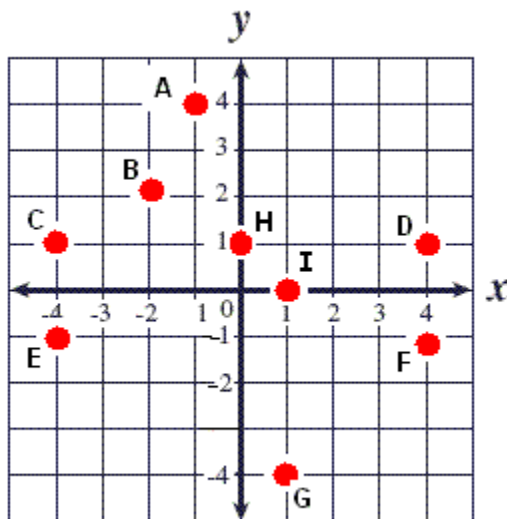


Multiple choice:

- ✗ A.
- ✗ B.
- ✗ C.
- ✓ D.
- ✗ E.
- ✗ F.
- ✗ G.
- ✗ H.
- ✗ I.

Hints:

- The x-coordinate is 4. Find 4 on the x-axis.
- The y-coordinate is 1. Find 1 on the y-axis.
- Find the point where 4 on the x-axis and 1 on the y-axis meet. Which point is it?

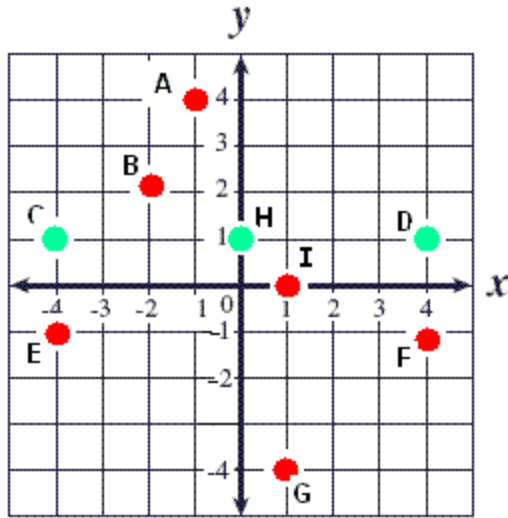


- The answer is D. Please choose D.

Scaffold:

Now we know that the correct graph will show a line that has the points shown (in green) below.

Which graph contains these points?

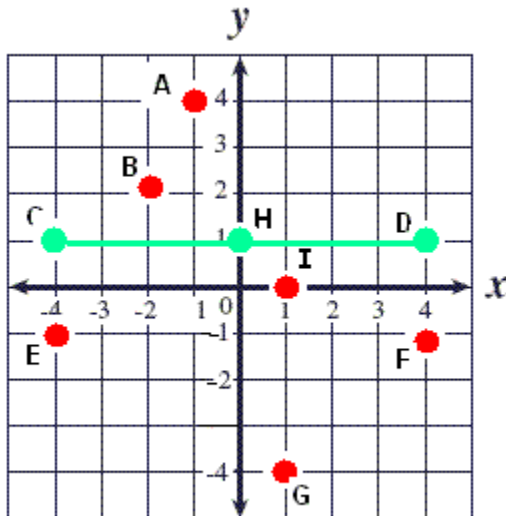


Multiple choice:

- A.
- B.
- C.
- D.

Hints:

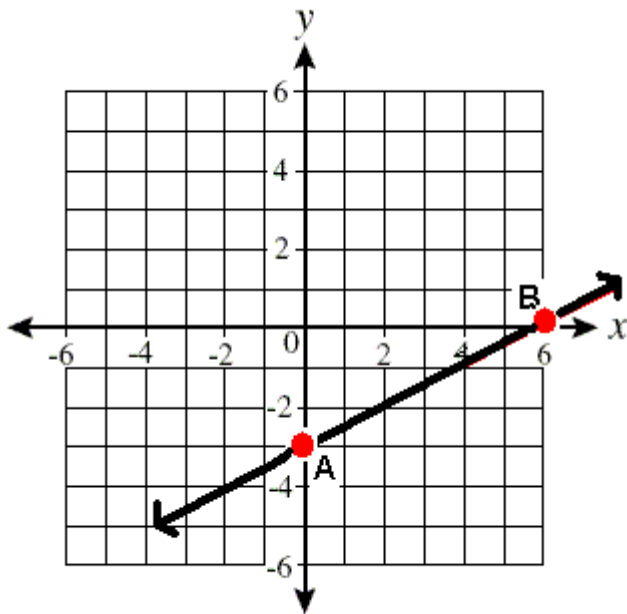
- Draw a line through the points. Which graph has the same line?



- Graph B has a line that goes through these points.
- The answer is B. Please choose B.

5) Assistentment #2716 "2716 - Item18_2003_find_slope_morph"

What is the slope of the line graphed below?



Algebra:

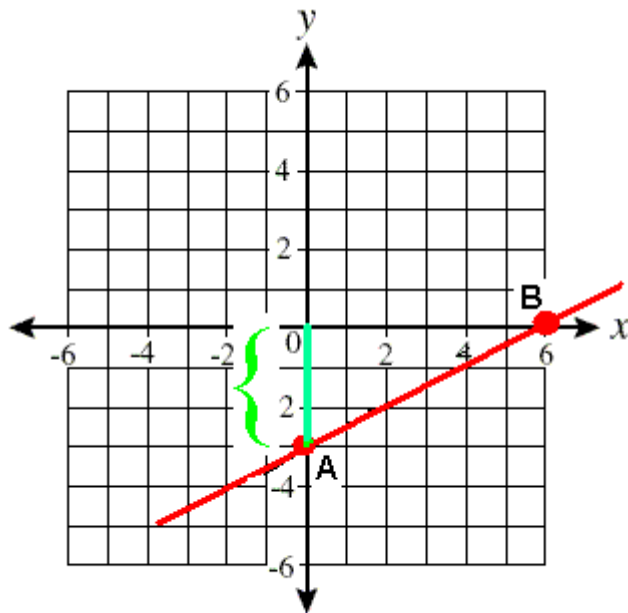
✓ $3/6$

Scaffold:

Slope is a number that measures the steepness of a straight line.

We measure slope by picking 2 points and dividing the change along the y-axis by the change along the x-axis.

What is the change along the y-axis for this line from point A to point B?



Algebra:

✓ 3

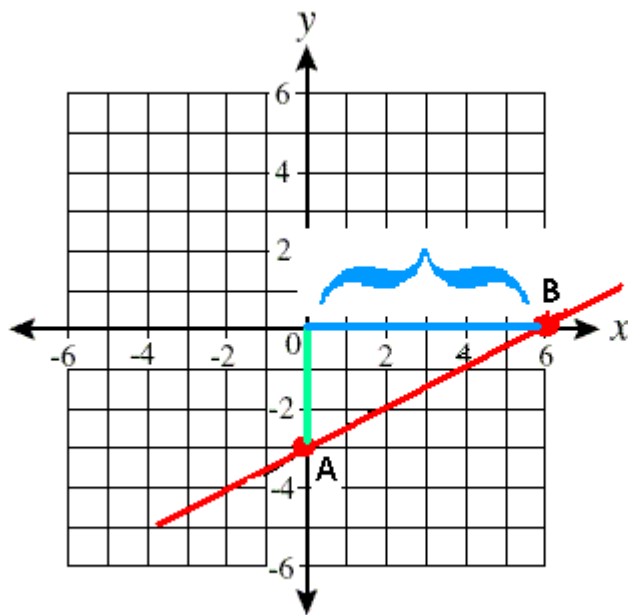
✗ 4

Hints:

- The change in y is the amount the line "goes up" from point A to point B. How many spaces does the green line go up?
- The answer is 3.

Scaffold:

What is the change in x for this line from point A to point B?



Algebra:

✓ 6

Hints:

- The change in x is the amount the line "goes over" from point A to point B. How many spaces does the blue line go from point A to point B?
- The change in x is 6.

Scaffold:

Good. Remember, the slope is the change in y divided by the change in x. What is the slope for this line?

Algebra:

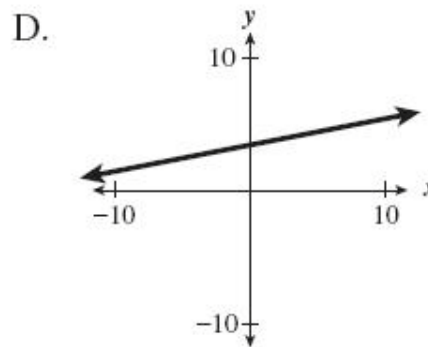
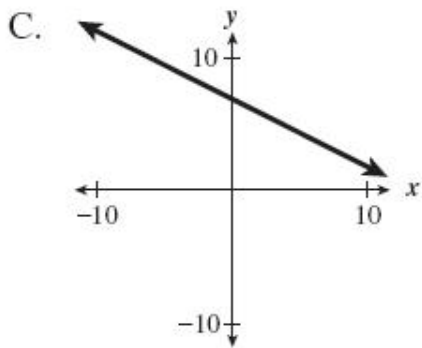
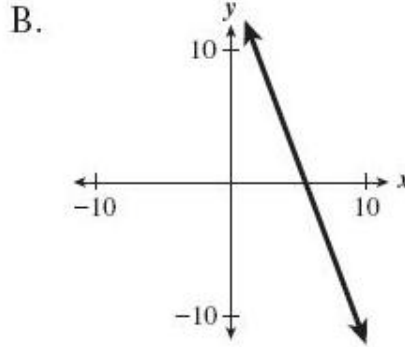
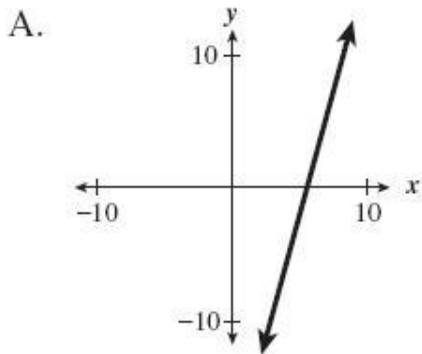
✓ $\frac{1}{2}$

Hints:

- The change in y from point A to point B is 3. The change in x from point A to point B is 6. The slope can be found by dividing the change in y by the change in x.
- The slope is $\frac{3}{6}$.

6) Assistentment #4247 "4247 - 2005-38"

Which of the lines graphed below has the **greatest positive slope**?



Multiple choice:

- Graph A
- Graph B
- Graph C
- Graph D

Scaffold:

Which of the graph(s) shows a line that has a positive slope?

Multiple choice:

- Graph A

You are right that Graph A has a positive slope, but there is one more graph that also shows a positive slope.

- Graph B
- Graph C
- Graph D

You are right that Graph D has a positive slope, but there is one more graph that also shows a positive slope.

- Graphs A and B
- Graphs A and D

✗ Graphs B and C

✗ Graphs B and D

Hints:

- Graphs with a positive slope go up from left to right. They look like they're "climbing" rather than "falling"

Which of the graphs look like they are climbing?



The line is sloping upward (left to right), the slope is positive.

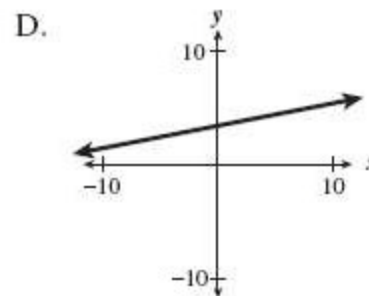
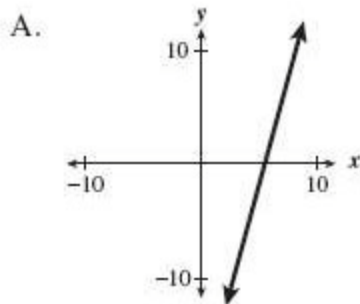


The line is sloping downward (left to right), the slope is negative.

- Graphs A and D both have positive slopes

Scaffold:

Of Graph A and Graph D, which one has the greatest slope?



Multiple choice:

✓ Graph A

✗ Graph D

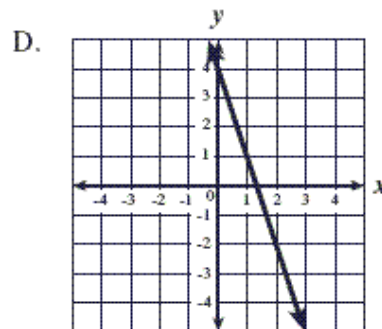
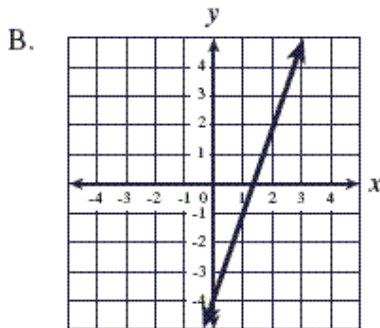
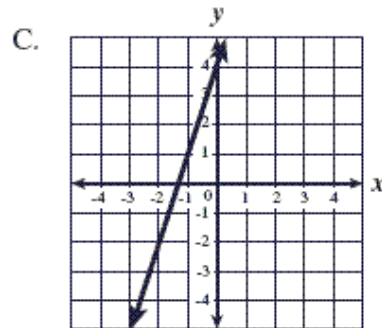
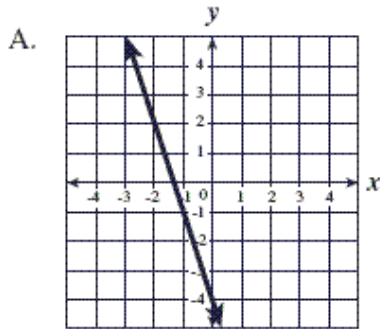
Hints:

- Which of the two graphs has a steeper line?
- Which of the two graphs looks like its climbing the fastest?
- Graph A has the greatest slope

Appendix B.3 Hints on demand condition

7) Assistentment #9771 "9771 - hint_item1"

Which graph below best represents
 $y = -3x + 4$?



Multiple choice:

- A.
- B.
- C.
- D.

Hints:

- The equation of a line can be written as follows:

$$y = mx + b$$

where **m** is the **slope**

and **b** is the **y-intercept**.

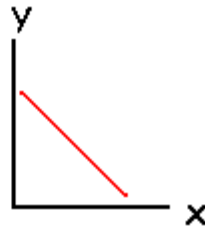
According to the equation $y = -3x + 4$, what is the slope of this line?

Is the slope of the line positive or negative?

- The slope of the line is -3, so it is negative.
Which graph(s) show a line with a negative slope?



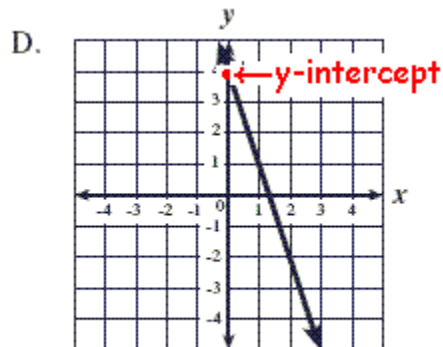
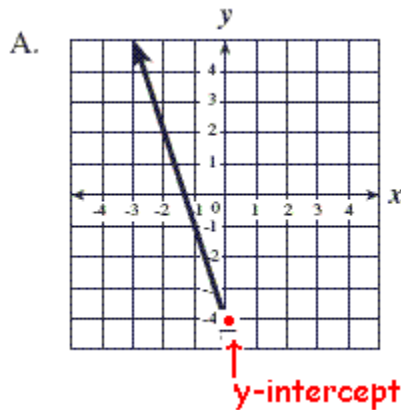
The line is sloping upward (left to right), the slope is positive.



The line is sloping downward (left to right), the slope is negative.

Graphs A and D have negative slopes so you can eliminate Graphs B and C.

- According to the equation, the y-intercept is 4. Which graph shows a line with a y-intercept of 4? Well the y-intercept is the y value of the point where the line crosses the y-axis.



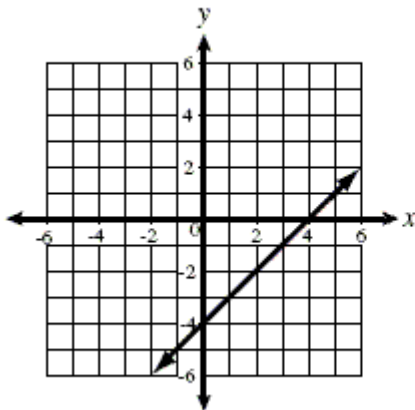
Graph D has a y-intercept of 4. Choose Graph D.

8) Assistent #9772 "9772 - hint_item2"

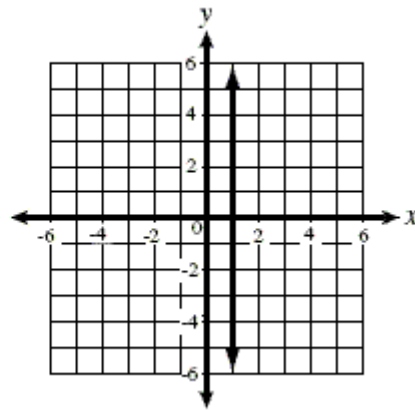
Which graph contains the points given in the table below?

x	y
-4	1
0	1
4	1

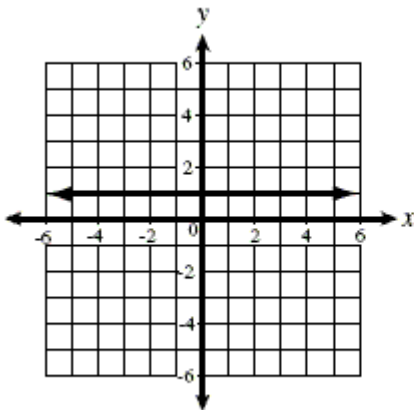
A.



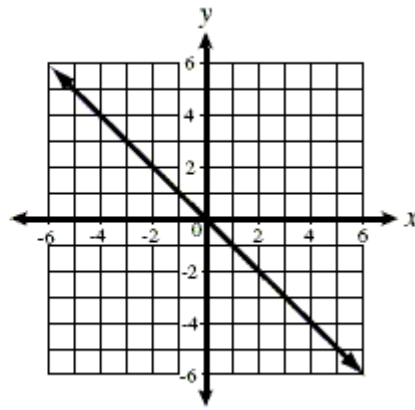
C.



B.



D.



Multiple choice:

A.

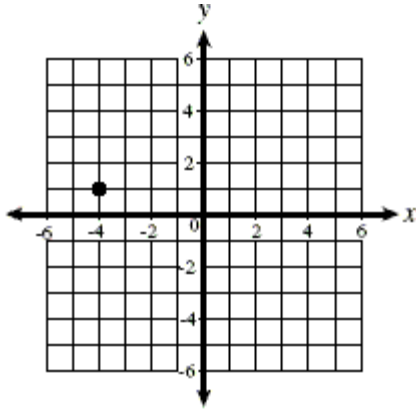
B.

C.

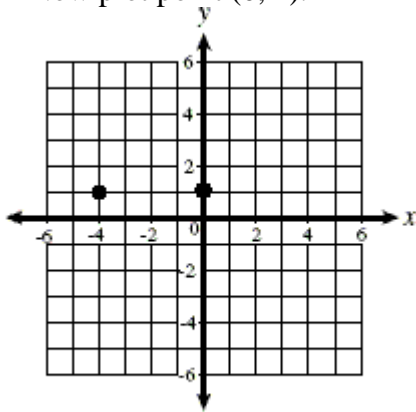
D.

Hints:

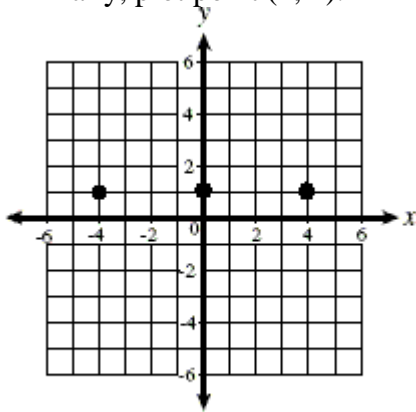
- Try plotting the points in the table. Start with point (-4, 1)



- Now plot point (0, 1).



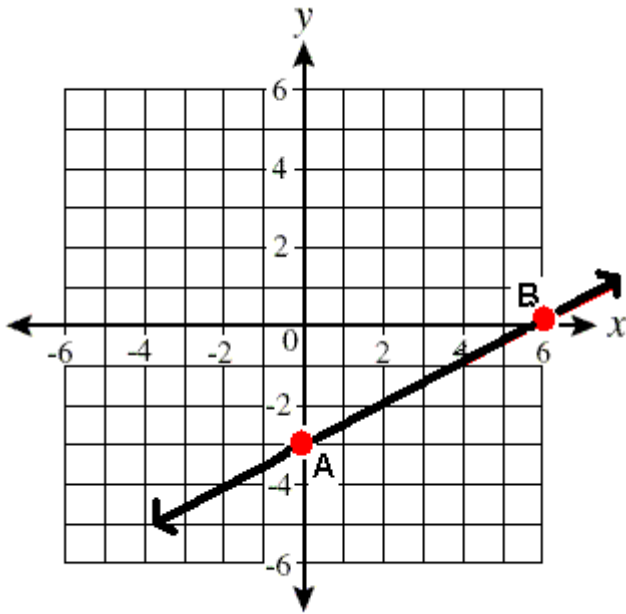
- Finally, plot point (4, 1).



- Which graph contains these 3 points?
- The answer is B.

9) Assistentment #9773 "9773 - hint_item3"

What is the slope of the line graphed below?

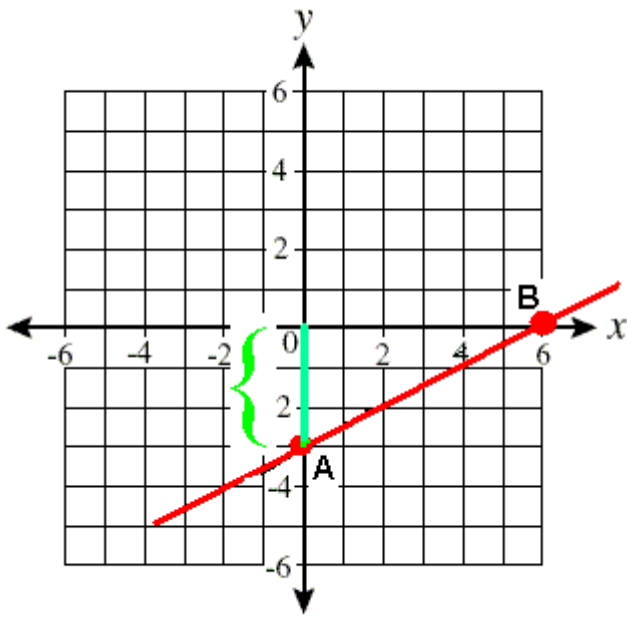


Multiple choice:

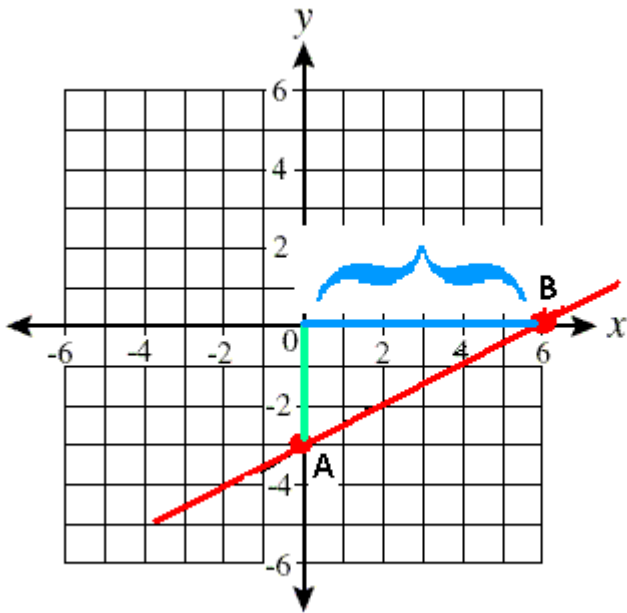
- A. 2
- B. 3
- C. $\frac{3}{6}$
- D. $\frac{6}{3}$

Hints:

- Slope is defined as the change in y divided by the change in x from point A to point B.
The change in y from point A to point B is 3 as shown below.



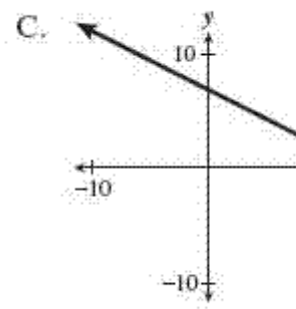
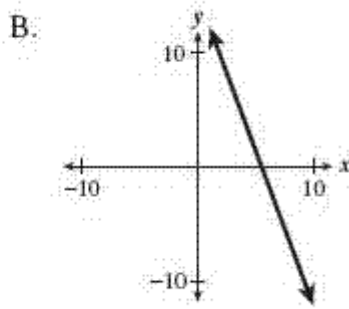
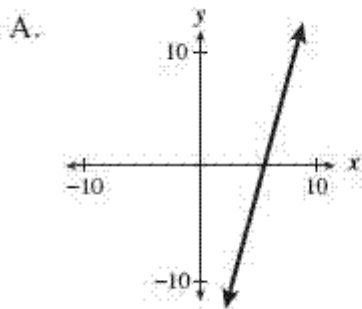
- The change in x is 6 from point A to point B as shown below.



- The change in y is 3.
- The change in x is 6.
- What is the slope?
- The slope is $\frac{3}{6}$

10) Assistentment #9774 "9774 - hint_item4"

Which of the lines graphed above has the greatest positive slope?



Multiple choice:

- Graph A.
- Graph B.
- Graph C.
- Graph D.

Hints:

- Which graph(s) have a positive slope?
-



The line is sloping
upward (left to right),
the slope is positive.



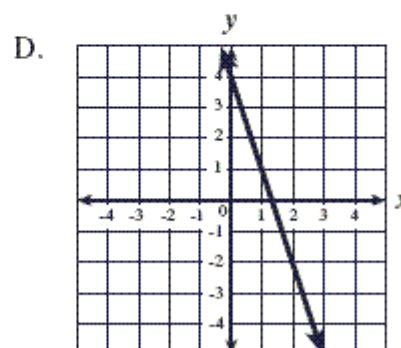
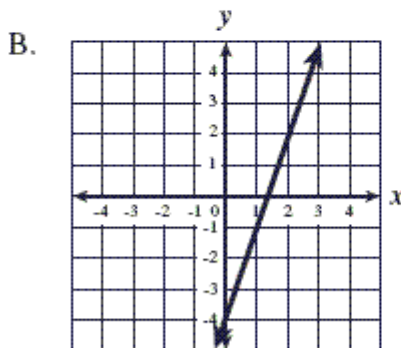
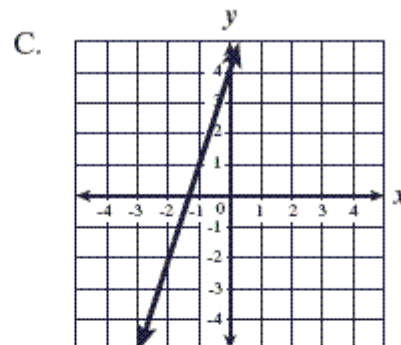
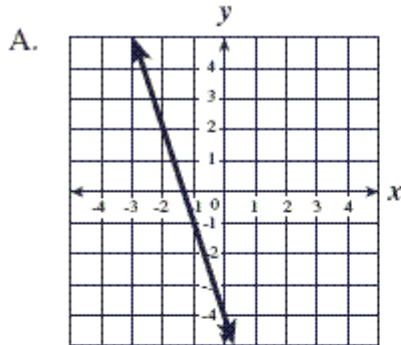
The line is sloping
downward (left to right),
the slope is negative.

- Graphs A and D have positive slopes.
- Between Graphs A and D, which graph has the steepest line?
- The answer is A. Please Choose A.

Appendix B.4 Worked Solutions Condition

11) Assistentment #13953 "13953 - no_feedback_1"

Which graph above best represents
 $y = -3x + 4$?



Multiple choice:

- A.
- B.
- C.
- D.

Scaffold:

The system will not tell if you got the answer right or wrong, right now. This problem will be explained to you at the end of this assignment. Please choose 'Ok' to proceed.

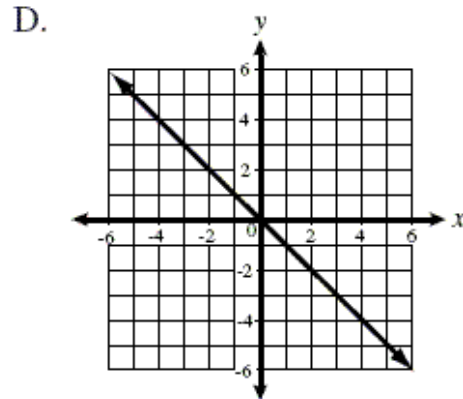
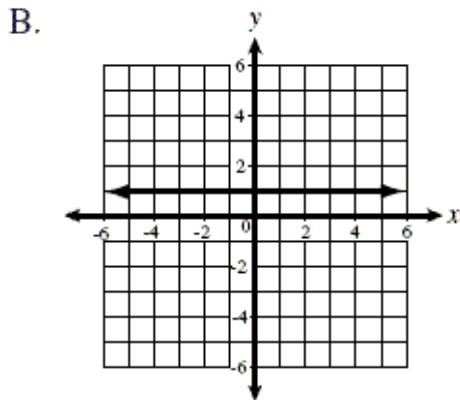
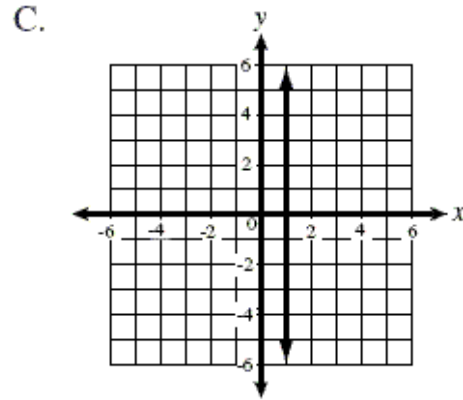
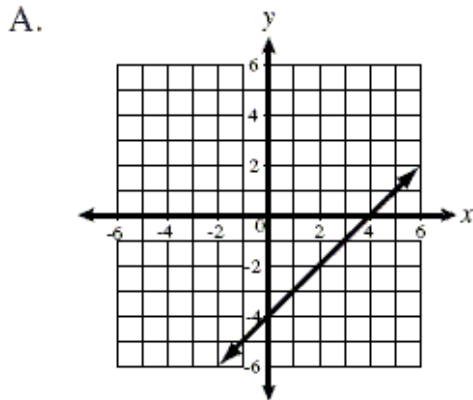
Multiple choice:

- Ok

12) Assisment #13954 "13954 - no_feedback_2"

Which graph contains the points given in the table below?

x	y
-4	1
0	1
4	1



Multiple choice:

A.

B.

C.

D.

Scaffold:

The system will not tell if you got the answer right or wrong, right now. This problem will be explained to you at the end of this assignment.

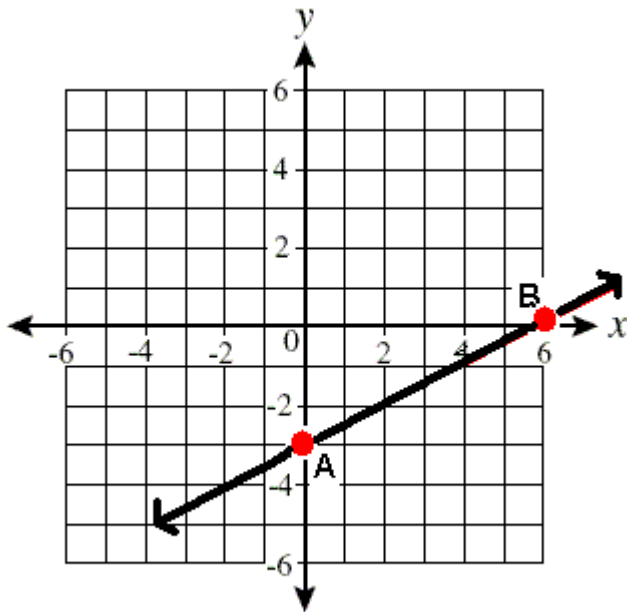
Please choose 'Ok' to proceed.

Multiple choice:

Ok

13) Assisment #13955 "13955 - no_feedback_3"

What is the slope of the line graphed above?



Multiple choice:

- A. $\frac{3}{6}$
- B. $\frac{6}{3}$
- C. 2
- D. 3

Scaffold:

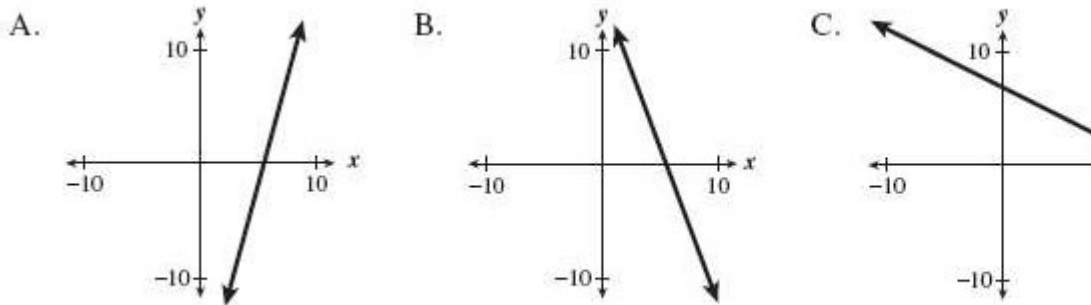
The system will not tell if you got the answer right or wrong, right now. This problem will be explained to you at the end of this assignment. Please choose 'Ok' to proceed.

Multiple choice:

- Ok

14) Assistentment #13959 "13959 - no_feedback_4"

Which of the lines graphed above has the greatest **positive** slope?



Multiple choice:

- Graph A
- Graph B
- Graph C
- Graph D

Scaffold:

The system will not tell if you got the answer right or wrong, right now. This problem will be explained to you at the end of this assignment. Please choose 'Ok' to proceed.

Multiple choice:

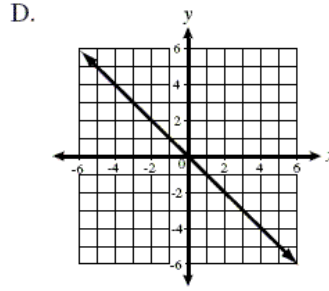
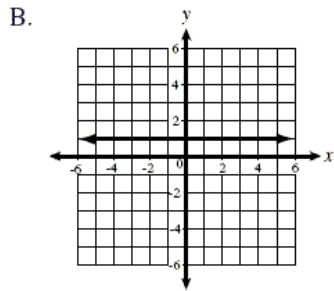
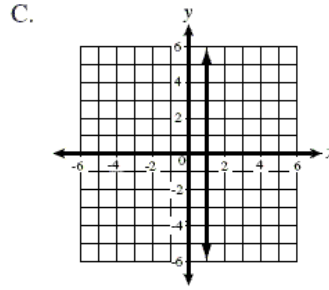
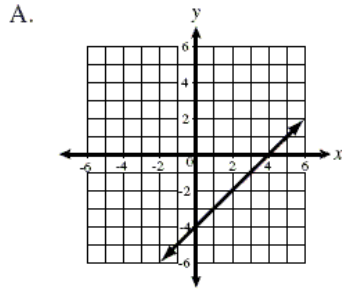
- Ok

15) Assistentment #14074 "14074 - explanation1"

You have finished all of the homework problems. Click Ok to see the correct answers for all of the problems you have just done. Please read through them carefully and make sure that you understand them. Don't forget to scroll up to see each complete problem and explanation.

Multiple choice:

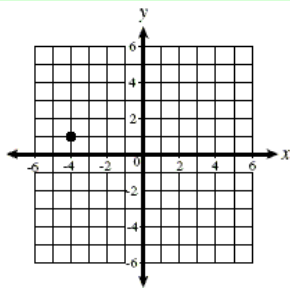
- Ok, I am ready to read the explanations.



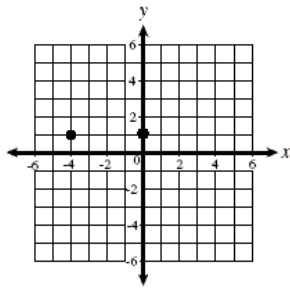
Which graph contains the points given in the table below?

x	y
-4	1
0	1
4	1

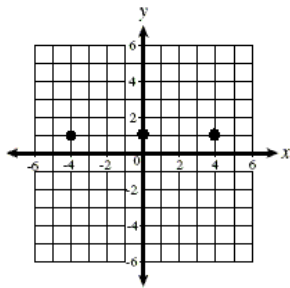
The correct answer is Graph B. Read the explanation below to see why the answer is Graph B.



Try plotting the points in the table. Start with point $(-4, 1)$



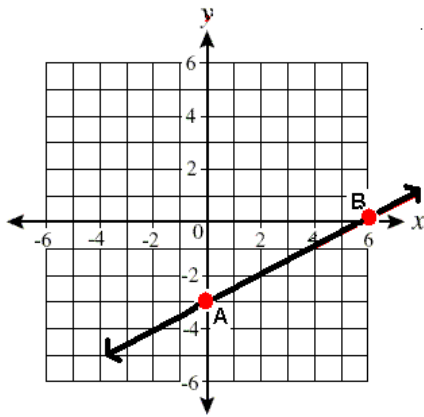
Now plot point $(0, 1)$.



Finally, plot point $(4, 1)$.

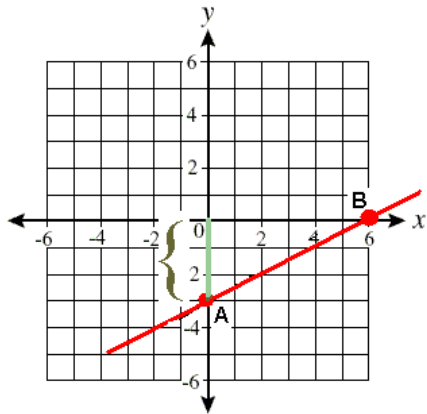
Which graph contains these 3 points?

The answer is B.

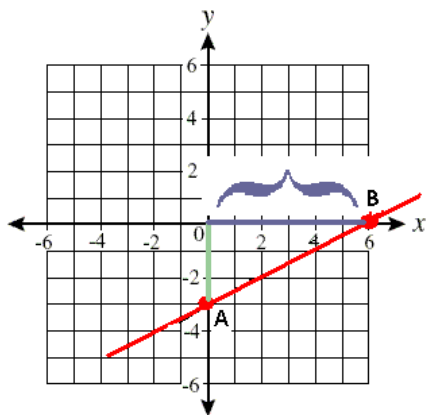


What is the slope of the line graphed above?

The answer is $3/6$. Read the explanation below to see how to find the slope.



Slope is defined as the change in y divided by the change in x from point A to point B .
What is the change in y from point A to point B ?



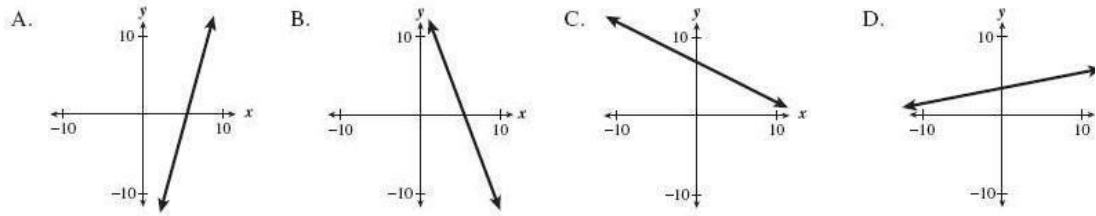
The change in y is 3.

What is the change in x from point A to point B ?

The change in x is 6.

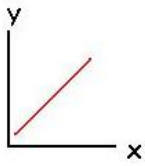
What is the slope?

The slope is $3/6$

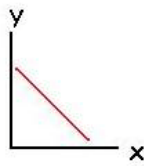


**Which of the lines graphed above has the greatest positive slope?
The answer is Graph A. Read the explanation below to see why.**

If you look at a positive slope from left to right, does it go up or down?

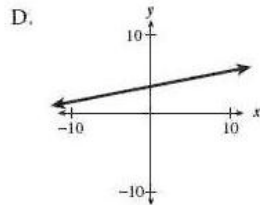
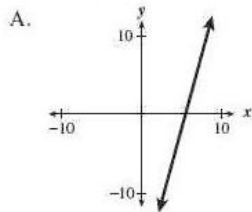


The line is sloping
upward (left to right).
the slope is positive.



The line is sloping
downward (left to right).
the slope is negative.

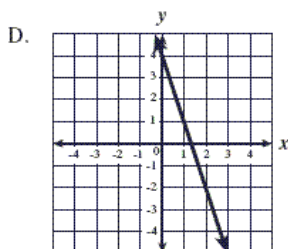
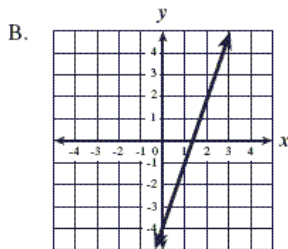
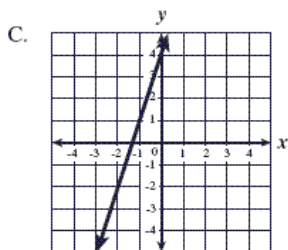
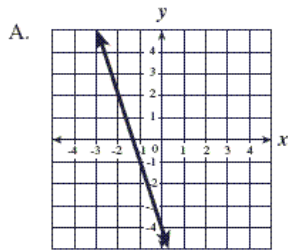
Graphs with a positive slope go up from left to right. They look like they're "climbing" rather than "falling"



Graphs A and D both have positive slopes.
Which of the two graphs has a steeper line?

Which of the two graphs looks like its climbing the fastest?

Graph A has the greatest slope



Which graph above best represents $y = -3x + 4$?
The correct answer is Graph D. Read the following explanation to see how to find the answer.

The equation of a line can be written as follows:

$$y = mx + b$$

where **m** is the **slope**

and **b** is the **y-intercept**.

According to the equation $y = -3x + 4$, what is the slope of this line?

Is the slope of the line positive or negative?



The line is sloping upward (left to right), the slope is positive.



The line is sloping downward (left to right), the slope is negative.

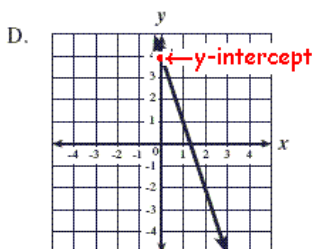
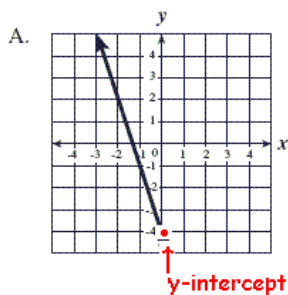
The slope of the line is -3 , so it is negative.

Which graph(s) show a line with a negative slope?

Graphs A and D have negative slopes.

According to the equation, what is the y-intercept of the line?

The y-intercept is 4. Which graph shows a line with a y-intercept of 4?



The y-intercept is the y value of the point where the line crosses the y-axis.

Which graph shows a line with a y-intercept of 4?

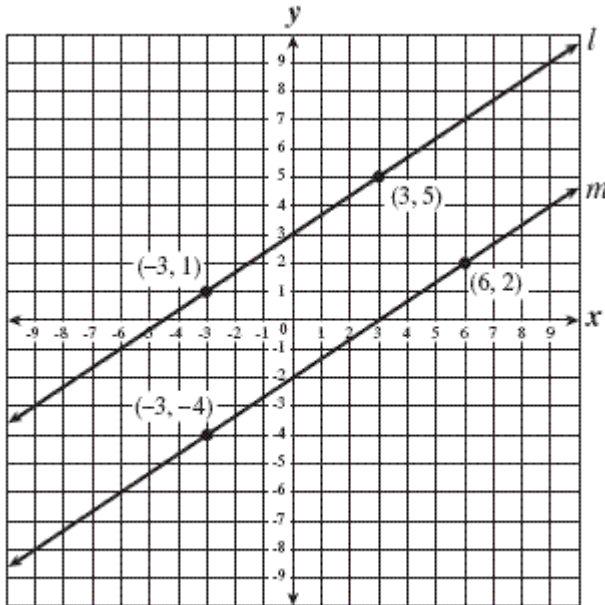
The answer is Graph D.

Appendix B.5 Post-test

1) Assistentment #9795 "9795 - 4_2005_gr8_transfer"

The coordinate grid below shows the graphs of two lines: line l and line m .

Which of the following is a true statement about the relationship between line l and line m ?



Multiple choice:

- A. The slope of line l is greater than the slope of line m .
- B. The x-intercept of line m is greater than the x-intercept of line l .
- C. The y-intercept of line m is greater than the y-intercept of line l .
- D. The slope of line m is greater than the slope of line l .

Scaffold:

The explanation to how to solve this problem will be presented at the end of this assignment. Please choose 'Ok' and click Submit to proceed.

Multiple choice:

- Ok

2) Assistentment #14499 "14499 - tranfer_2003_7_gr8"

What is the y-intercept of the graph represented by the equation below?

$$y = \frac{4}{5}x - 2$$

Algebra:

- 2
- $y=-2$

Scaffold:

The answer is -2. This is because the x-value at the y-intercept is always 0 so if we plug in 0 for x then we get -2.

Please choose 'Ok' and click the Submit button to proceed.

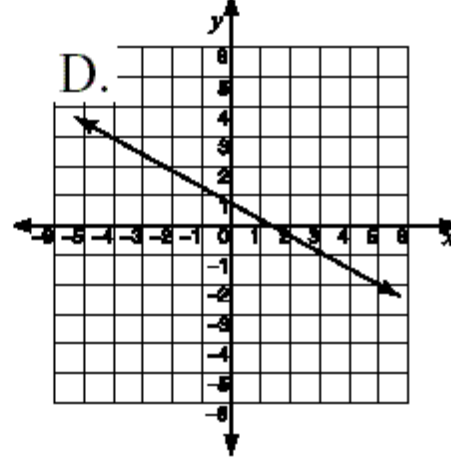
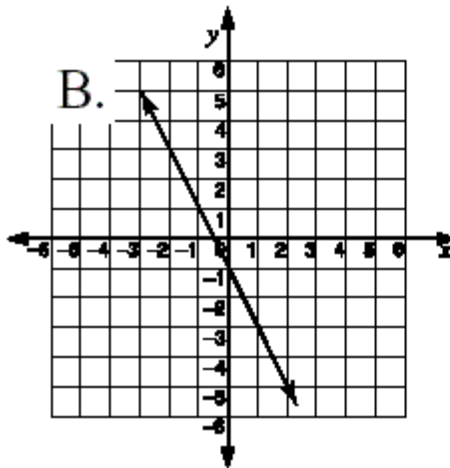
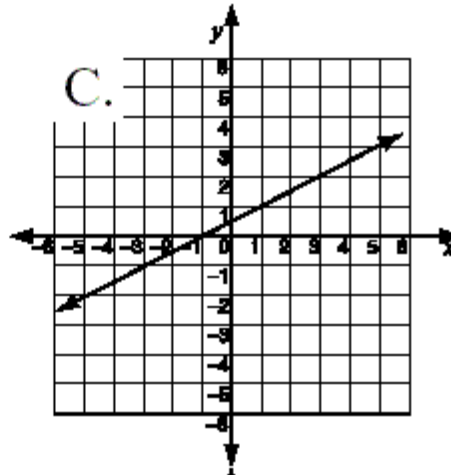
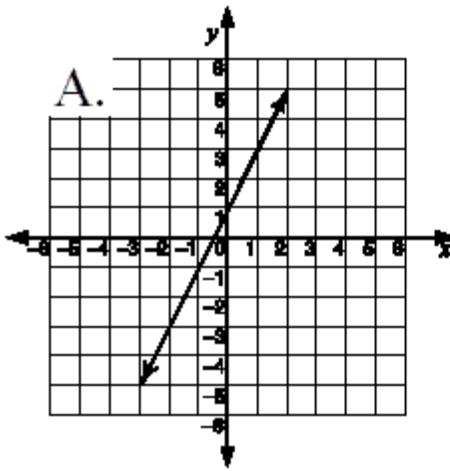
Multiple choice:

✓ Ok

3) Assistent #9800 "9800 - 3_2001_transfer"

Which graph below contains the points in the table below?

x	y
-2	-3
-1	-1
1	3



Multiple choice:

- ✓ A.
- ✗ B.
- ✗ C.
- ✗ D.

Scaffold:

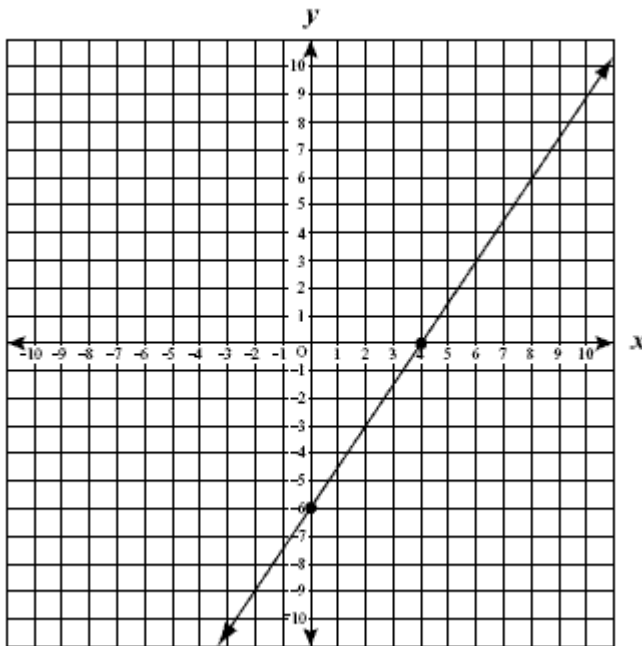
The explanation of how to find the correct graph will be shown at the end of this assignment. Please choose 'Ok' and click Submit to proceed.

Multiple choice:

- ✓ Ok

4) Assistent #9801 "9801 - 18_2003_transfer"

What is the slope of the line graphed below?



Algebra:

- ✓ 6/4

Scaffold:

The slope of the line is 6/4.

The slope is found by dividing the rise, which is 6, by the run, which is 4.

Please choose 'Ok' and click Submit to proceed.

Multiple choice:

- ✓ Ok