Lessons Learned From a Failed Experiment:

A Very Brief Introduction of System Dynamics Modeling in

Two Algebra II Classes

Diana M. Fisher, Ph.D., Portland State University

Abstract

The purpose of this paper is to present the results of an experiment conducted at a secondary school in the US in which two mathematics teachers were enlisted to teach a series of 6 lessons introducing System Dynamics modeling as a method to help students better understand the difference between the structures that produce linear versus exponential change over time in real world scenarios. There are very few formal research studies attempting to document learning outcomes from using System Dynamics modeling in mathematics classes at the secondary school level. Four algebra II classes were involved in the study, an experimental and control class for each of the two participating teachers. The teachers administered pre- and post-assessments to each of the four classes and taught 6 experimental lessons to their classes that were designated as experimental. Due to a number of difficulties that contributed to a very limited time to conduct the assessments and lessons the pre- and post-assessments did not produce significant differences between the experimental and control groups. However, the results collected on some of the lessons indicated that SD modeling might potentially reduce the gap in performance between students who are more adept at traditional math and those who are not. Videotaped think-aloud protocols used with two pairs of students, one from each experimental group indicated that, even with such a brief introduction, students can build correct SD models and correctly analyze the behavior for a problem that is slightly beyond the typical problem studied an algebra II. The experiment should be repeated over a longer period of time

and incorporated into the regular curriculum to overcome most of the problems that arose with the execution of this experiment. This paper includes a great deal of information about the experiment that could be used by others interested in improving the method described here. It is hoped future experiments will be conducted to add to the literature documenting the learning outcomes that SD modeling and analysis provides.

A. Introduction

> "Concepts and methods enabled by rapid advances of information technologies are enabling us to understand aspects of the real world where events and actions have multiple causes and consequences, and where order and structure co-exist at many different scales of time, space and organization. Within this complexity framework, critical behaviors that were systematically ignored by classical science can now be included as essential elements that account for many observed aspects of our world–for example, global phenomena that require multiple physical, biological, social, and mathematical perspectives" (Kaput, Bar-Yam, Jacobson, Jakobsson, Lemke, Wilensky, 2000, p 1).

The literature indicates a pressing need for adults to become better at understanding complex systemic behavior (Homer & Hirsch, 2006; Hung, 2008; Sterman, 2000, 2011). If understanding complex systemic behavior is so important for adults, it seems reasonable to expect mainstream K-12 education to begin to incorporate activities that give students experience working with "wicked" problems. The importance of including the study of complex systems in mainstream education is increasingly appearing in research articles within the last decade, even for elementary school students (English, 2007). Jacobson and Wilensky (2006) state,

> "The conceptual basis of complex systems ideas reflects a dramatic change in perspective that is increasingly important for students to develop as it opens new intellectual horizons, new explanatory frameworks, and new methodologies that are becoming of central importance in scientific and professional environments" (p. 12).

Although there are a variety of analytical methods that can be applied to the study of complex systems problems, the focus of this paper is on the system dynamics method.

Some Current Efforts in K-12 Education in the US

The Waters Foundation[1] has had success working with elementary and middle school teachers, training them to use behavior over time graphs, their "Habits of a Systems Thinker" cards, and Iceberg and Ladder of Inference analysis techniques to get teachers to begin using more holistic approaches to their content. Their goal is to help "build systems thinking capacity" in the teachers they train. This effort is scaling up.

The Creative Learning Exchange[2] has increased curricular offerings for K-12 teachers by commissioning curriculum development of models and lessons based on Forrester's statements regarding the nature of systems (Forrester, 2009). These lessons can be accessed on their website. Quaden, Ticotsky, & Lyneis (2008) have developed the Connection Circle to help elementary and middle school teachers and students analyze feedback loops in scenarios involving more than two components. These three gifted teachers have developed simple System Dynamics (SD) modeling lessons that help students in grades 5 – 8 (ages 11– 14 years) understand the importance of feedback in certain real-world scenarios highlighting the difference between linear and exponential patterns of change over time. CC Modeling Systems[3] also offers online courses for math and science teachers who want to learn how to create small SD models to use in their curriculum.

What is missing? Research on the efficacy of the SD modeling approach for K-12 education is missing. Mandinach and Cline (1993) conducted quite a few studies working with over 32 secondary school math, science, and social studies teachers incorporating SD into their curriculum. They discussed three levels of modeling using the STELLA software with high

---

[1] www.watersfoundation.org
[2] www.clexchange.org
[3] www.ccmodelingsystems.com

school students (ages 15 – 18 years). The first they called parameter manipulation and referred to this as the least cognitively demanding. Next they describe what they call constrained modeling, where students build a model to solve a specifically assigned problem. Finally, they describe "epitome" modeling, the most cognitively demanding, where students build original models for an idea they conceive. The needed modeling expertise of the teacher, as well as the student, increases significantly from the first to the third type of modeling, as does the amount of time that must be dedicated to the modeling activity in the curriculum. In a later article, Mandinach and Cline (2000) discuss the many difficulties that they encountered in trying to conduct their research and why they felt that this innovative approach was not destined to appear in mainstream education any time soon.

Yet, the need to document the value of the SD approach to learning about complex systems is reaching a critical stage. Those who are familiar with the approach appreciate what SD modeling can offer, but without the strength of research support to document what, heretofore, have been mostly anecdotal success stories, it will not be possible to convince educational decision-makers that efforts should be made to include this "new" approach in classrooms.

So this author undertook the design and execution of a classroom experiment to determine if the use of SD modeling might show an improvement in understanding a concept that is already in the algebra II curriculum, the difference between the underlying process that produces linear versus exponential change over time. This effort was to stay close to the current curriculum as a bridge to moving students toward the study of a slightly more difficult problem that would be outside their normal algebra II curriculum. It was hoped this might indicate a more immediate way to begin incorporating SD modeling into a traditional algebra II course without requiring much modification on the part of the teacher.

The rest of this paper describes the learning theory supporting the design of the experimental lessons, the importance of having students understand the function concept in mathematics, the school environment in which the experiment was conducted, the research question, the method used to conduct the experiment, the various data collected and analyzed, and a discussion of some of the results and lessons learned.

B. The Experiment

The study comprises an account of the student experience and the learning of the students as an introductory sequence of lessons was used to *prepare* students to build (in their final lesson) a model of a problem involving a combination of linear and exponential change over time (a situation not covered in algebra II). Design of the experimental lesson sequence and classroom environment were guided, in broad terms, by the learning theories developed by Lev Vygotsky and Jerome Bruner, as described below.

Theoretical Foundations

Learning Theory

Lev Vygotsky suggested that learning should be a socially active endeavor, where students are expressing their thinking, and the teacher is facilitating the process. This interaction should be cooperative and collaborative (i.e., the teacher uses demonstrations and leading questions) to be effective. Teachers do not transmit concepts. "If concept development is to be effective in the formation of scientific concepts [those new ideas learned in school] instruction must be designed to foster conscious awareness of concept form and structure and thereby allow for individual access and control over acquired scientific concepts" (Vygotsky in Daniels, Cole, & Wertsch, 2007, p. 312).

One of Vygotsky's major contributions to learning theory he called the "Zone of Proximal Development" (ZPD). The ZPD is conceived of as a gap between what the student could learn by him/herself and what he/she could learn with the help of more knowledgeable peers and/or the teacher. Vygotsky indicated that the trajectories for individual student learning in this zone are quite open and will follow dynamic and divergent paths. The objective of the "instruction" is, however, to help the student eventually internalize the new knowledge. Vygotsky (1978) indicated that essential (good) learning should create a ZPD ("awaken a variety of internal developmental processes in the child that are activated by working cooperatively with peers and other people in his/her environment", p. 90) that is forward looking, developmentally, rather than testing, which is backward (*ineffective*) looking. In this way, Vygotsky said, once the processes within the child become internalized they lead to *independent developmental achievement* (Vygotsky, 1997, italics added).

Jerome Bruner is an important interpreter of Vygotsky, developing Vygotsky's theory in certain directions. Still relevant today is his early book "Toward a Theory of Instruction" (Bruner, 1966) in which he presented three modes of representation that are needed to help students acquire new ideas with understanding. The first mode is enactive, wherein students manipulate concrete objects to gain an understanding of the elements in the system and how they might be related. The second mode is iconic. In this mode students use some pictorial representation of the system they experienced in the enactive mode, to capture the structure or behavior that was present in the activity. Creating and reading graphs are examples within this mode, as is the construction of various diagrams. This mode is still quite concrete – the iconic representation is directly connected to a physical activity. The final mode is symbolic, wherein students use symbols, such as numbers, computational symbols, or words to start to abstract the ideas from the concrete to other similar patterns existing in problems they do not physically experience. STELLA modeling, with pre-activities involving physical simulation, constitutes part of a learning experience that exploits all three of Bruner's modes.

The team method of building System Dynamics models, used consistently in K-12, is well grounded in these learning theories as an effective learning strategy. Students work collaboratively with each other (and with teacher facilitation) to determine what components to include in the model, how they should be connected, whether the simulation results are reasonable, how modifications to the values or to the structure modify the system behavior, and recommend possible "solutions" or policies that might transform the system to produce more desirable behavior. This focused, active interaction with the modeling process aligns well with Vygotsky's description of lessons that would produce effective concept development.

The experimental lessons focus on strengthening student conceptual understanding of linear and exponential functions, the concept of function being foundational in algebra.

The Importance of the Function Concept in Mathematics

> "The concept of function is central to undergraduate mathematics, foundational to modern mathematics, and essential in related areas of the sciences. A strong understanding of the function concept is also essential for any student hoping to understand calculus – a critical course for the development of future scientists, engineers, and mathematicians" (Oehrtman, Carlson, & Thompson, 2008, p. 27).

Many mathematics education researchers present studies attempting to determine whether students understand the abstract definition of function, i.e., that a function is a special relationship between two variables whereby each value of one (the independent variable) within a specified range maps to at most one value of the other (the dependent variable). System Dynamics fits into a broader, more dynamic, more applications oriented view of a function. Using the stock/flow diagramming structure of SD many of the continuous elementary functions used in an algebra classroom can be constructed. Those stock/flow structures can serve as an alternate, two-dimensional symbolic representation for the elementary functions. The *connection* between different representations of functions is essential in helping student build a stronger understanding of the function concept (Leinhardt, Zaslavsky, & Stein, 1990; Keller & Hirsh, 1998; Yerushalmy, 1991).

The elementary functions that are studied in most second year algebra classes in secondary school represent different patterns of change over time that occur often enough that mathematics educators think students should be able to recognize them whenever they arise in class, whether by viewing the graph or by solving a story problem. But story problems tend to be simplified due to the limitations students have in understanding closed form equations.

> "What makes teaching (and learning) of the translation skills so difficult is that behind them there are many unarticulated mental processes that guide one in constructing a new equation on paper. These processes are not identical with the symbols: in fact, the symbols themselves, as they appear on the blackboard or in a book, communicate to the student very little about the processes used to produce them" (Clement, Lockhead, & Monk, 1981, p. 289).

These functions are generally only called to represent change in one direction (increase or decrease, never both). But systems are not so restricted. So for students to be able to study problems involving systems, it should not be expected that they jump from the graph or story description directly to a "correct" symbolic representation immediately. Kaput (1999) indicated that functions are used to build mathematical systems through successive approximations, where each iteration tries to improve the structure built in an attempt to understand a problem under study. This process involves modeling and some researchers argue that modeling problems/phenomena is the primary reason to study algebra.

The SD stock/flow function diagram is built with a focus on how the structure produces the function's characteristic pattern of change over time. This structure opens up a representation that has been used successfully to provide a conceptual introduction to elementary calculus concepts within algebra classes. It also can provide a vehicle for enabling students to use the elementary function structures as building blocks that will allow them to study more realistic applications by having them build models that combine these elementary function structures, as small Lego structures can be used to build larger Lego systems.

The lessons for this study focus on increasing student understanding of linear and exponential change over time. Two lessons are devoted to introducing and then enhancing student understanding of the behavior of each function. The first of these two lessons uses a kinesthetic activity to introduce the primary dynamic characteristic of the function, that is, coupling the behavior with a specific type of rate of change (constant for linear, proportional to current value for exponential).

Students follow specific directions in these kinesthetic lessons initially, then are expected to produce their own actions to generate the same function behavior emphasized in the lesson. The attempt is to move students from a "following directions" mode of thinking to a "create your own" understanding of the function. The second function lesson (for each function) has students build an SD model that produces the characteristic function behavior (with focus on rate of change), following directions given in the lesson. Students experiment with the models and explain the reason the model output changed. Students, again, are expected to create an original model that produces the function behavior under study.

Modifying the models gives students more experience with each function's behavior. Finally, students are given one problem that involves both a constant and a proportional rate of change and students, it is hoped, understand the type of function structure needed to capture these dynamics as they create one model structure to mathematize the problem. This lesson is designed to move students from the "create your own" understanding on a single function type to the "transfer to a new situation" concept of function where they recognize, from the description, the types of functions that must be used to solve the new, more complicated problem (involving more than one function type).

In addition to doing the tasks, students will also be expected to explain what they create as a culminating activity in each stage of the experimental/intervention instruction. Explaining one's creations has been shown to "enhance learning and understanding of new knowledge" (Chi, DeLeeuw, Chiu, & LaVancher, 1994, p. 469). Students will be working in teams on most of the activities. Team problem solving is considered a positive learning environment, improving student attitude toward mathematics, and fostering more student engagement in lessons and with each other (Davidson, 1990; Springer, Stanne, & Donovan, 1999).

The Teachers and School Environment

This study was conducted in partnership with the curriculum vice-principal and two algebra II teachers at a local high school in Portland, Oregon. This school serves primarily a middle socio-economic population that is 22% minority, and about 24% of its students are considered disadvantaged. It has 56 teachers who serve about 1250 students. The school has no Title I funding. The students enter the school generally at normal grade level in mathematics. Reporting on state mathematics tests indicate that 79% of the students are proficient in mathematics.

The two teachers are experienced secondary mathematics teachers with between 14 and 32 years of experience, respectively, teaching at their current high school. Each has taught algebra II between 10 and15 years, respectively, each is quite comfortable with technology, and each has used STELLA model-building briefly in their algebra II classes in the past, and then quit using the STELLA software.

The student segment of this study attempted to determine how well students understood an extension of the elementary functions (linear and exponential) as they designed a new function that combined the change dynamics associated with the two initial functions.

The Research Question:

Can System Dynamics model-building activities aid students in identifying and differentiating linear and exponential function behavior over time in the context of a real-world scenario?

Method

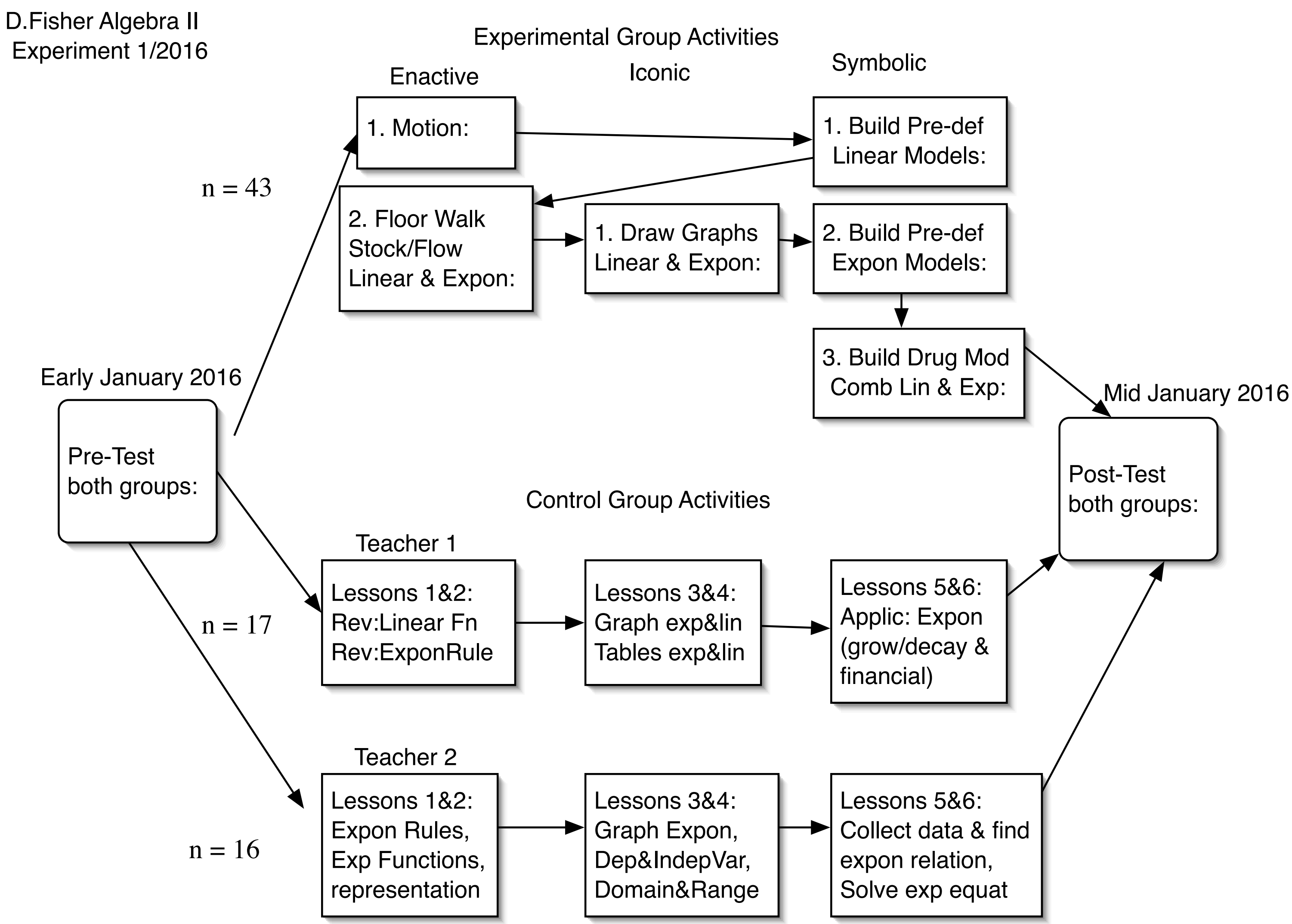

Figure 1: Quick overview of the method for the algebra II student experiment discussed in this paper.

Figure 1 displays a quick overview of the method for the student experiment. There were 6 lessons for the experimental algebra II class. Two of these lessons fall into the "enactive" mode of Bruner's learning theory, two in the "iconic" mode, and two in the "symbolic" mode. (See supplemental file.) All lessons require students to work with the topic for the lesson and three of the six require the creation of original scenarios and explanations dealing with the focus topics. The enactive mode lessons are intended to determine if students have moved from "following directions" mode to the "create your own" thinking about linear and exponential functions. The iconic and symbolic lessons are intended to determine if students have moved from "create your own" thinking to being able to "transfer" the function concept to a new situation involving more than one function type.

A quick overview of the lessons to be used with students will be described below, with a more detailed explanation of the lessons in the supplemental file (Appendix A).

1. (Motion) As a full-class activity the teacher had students move, to produce different types of linear motion, in front of a motion detector connected to a computer that projected the motion graph on the overhead screen. As a final problem, an original walking scenario was designed by each student and the students explained how one would walk to produce the graph they drew. This problem was "graded" by the researcher.

2. (Linear Models) Using the STELLA software on a classroom set of netbooks students (in teams of two) built linear models matching scenario descriptions given on a handout. A final assignment asked students to sketch the stock/flow diagram for an original linear model devising their own scenario. The original model was "graded" by the researcher.

3. (Floor stock/flow activity: Linear functions and exponential functions) The teacher created a large stock/flow structure on the floor of the classroom and had the students enter and leave the stock in different linear and exponential patterns. A table of values (number of students in the stock) was recorded for each different pattern.

4. (Graphs) Students graphed the values that were recorded in the tables that were produced in activity 3 above.

5. (Exponential Models) After the teacher built a stock/flow bank account with simple interest model in class students used the classroom set of netbooks to build (in teams of two) various

exponential models based on problem descriptions given in their handout. A final assignment asked students to sketch the stock/flow diagram for an original exponential model devising their own scenario. The original model was "graded" by the researcher.

The following day the teacher built a stock/flow bank account model with interest added (inflow) and constant withdrawals (outflow) with the aid of students in the class, asking students how to build the model and asking them to predict behavior based on changing values in the model.

6. (Combine Functions: The Drug Model) A random selection of two pairs[4] of students were removed from the experimental classes when the rest of the experimental classes were working (in teams of two students each) on the drug model lesson. Each randomly selected pair of students built a drug model requiring a constant inflow from an IV drip and exponential outflow simulating the body metabolizing and eliminating the drug. These students were videotaped by the researcher. The rest of the class built the same model, while the two pairs of students were being videotaped. The researcher "graded" the entire drug packet for all of the experimental students.

   The STELLA diagram for the Drug Model is shown in Figure 2 below.

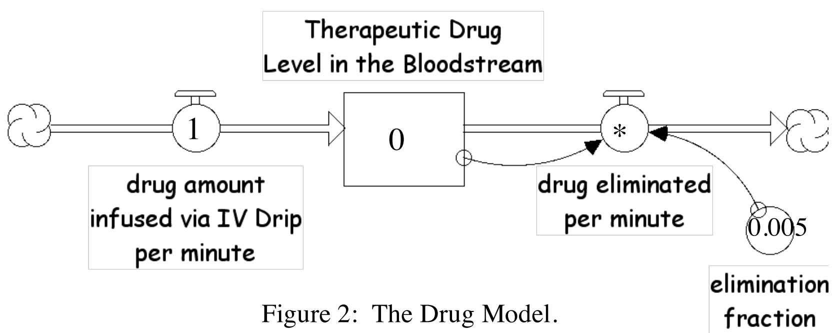

Figure 2: The Drug Model.

Data and Analysis

Originally it was intended to conduct these lessons throughout the first semester of the year, but the teachers indicated that the curriculum structure of the algebra II course had changed recently and they no longer taught linear and exponential functions as separate units in algebra II. So they wanted to conduct the experiment in a brief 3-week unit review just before the first semester exam. The 3-week period was to occur directly after winter break. Another situation

[4] Student names were placed in pairs on small papers and selected at random from a container. The pairing was designed to maintain student comfort in working with another student they usually choose to work with, if possible, on the task for activity 6.

occurred that pushed the semester exam forward by a week. So now there would be only two weeks between winter break and semester exams. Yet another situation occurred in that the school changed from a modified block schedule (3 – 50 minute class periods and 1 – 90 minute class period per week) to alternating block schedule, (alternating 3 – 93 minute classes per week and 2 – 93 minute classes per week). Moreover, there was not a guarantee of 93 minutes, since some days involving early dismissal, late arrival, tutor times, or assembly schedules changed class periods to either 63 or 75 minutes instead of 93 minutes.

Due to circumstances beyond the control of the teachers and researcher 2 snow days occurred immediately after winter break. The 6 lessons and the pre- and post-assessments had to be executed in 4 class periods of between 63 and 93 minutes each. Consequently, students were able to complete only the first 3 (of 4) exercises in the linear and exponential modeling lessons. For one teacher the graphing lesson had to be eliminated due to a very shortened class period. In the drug model lesson there were significant difficulties with the computers and many students were only able to complete ¼ of the lesson.

It is useful to know the math proficiency level of each of the four classes involved in this experiment. The math proficiency level of the students involved in this experiment was measured (only) by using the class mean scores of their first semester assessments in algebra II. The results were:

Table 1: Class mean scores, first semester 2015-2016, for each of the four classes in this experiment.

	Teacher 1	Teacher 2
Experimental Class	83% n=21	69% n=22
Control Class	74% n=17	85% n=16

The Pre- and Post-Assessments

The pre- and post-assessment structures were almost identical in that they contained the same number and types of questions, only differing in numeric values or in slight modification of question asked (i.e., growth questions changed to decay questions, etc.) The assessments were identified as form A or form B and half of the students in each class received form A and half form B for pre-assessment. Students received the other form for their post-assessment. There were 12 questions on each assessment form. Since the same assessments were to be used for both the experimental and control group students there were some traditional questions (questions 5 and 11[5]) as well as some more conceptual questions regarding linear and exponential function behavior, although no questions required the use of mathematical equations, to answer. More detail about the assessment questions can be found in the supplemental file (Appendix B).

Results: The following tables show the results on the pre- and post-assessments as coordinates on a grid. Each student has his/her own point on the grid. The first coordinate of each point represents the pre-assessment score for the student. The second coordinate represents the post-assessment score for the same student. Students above the y = x diagonal line indicate an improvement in their performance from the pre- to the post-assessment. Each of the four classes, two experimental and two control, has its own plot.

---

[5] Questions 4, 7, and 9 favored the experimental group. The rest of the questions 1, 2, 3, 6, 8, 10, and 12 were intended to be neutral.

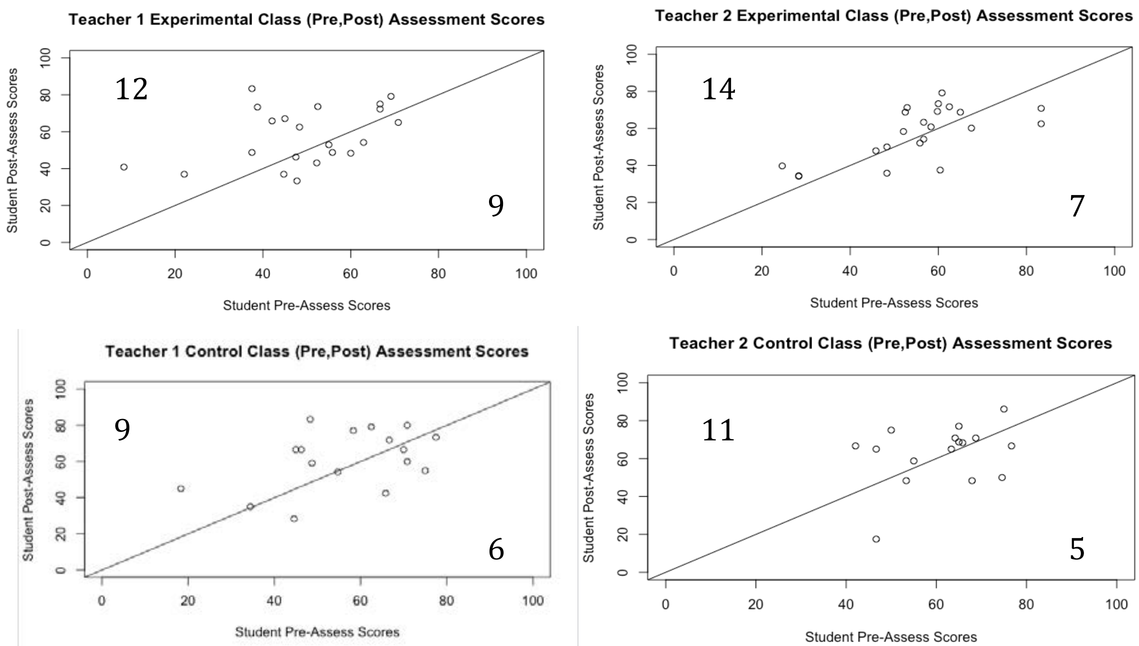

Figure 3: (Pre, Post) assessments plots for each student in each of the four classes involved in this experiment.

If we combine the two experimental groups and the two control groups and analyze individual questions, there was interesting information:

Table 2: Improvement in scores from pre- to post-assessments for the combined experimental and combined control groups, and then the improvement of the experimental over the control group on each question.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
Post-Pre Experimental Group n = 43	2.3	-1	-1.7	2.3	-8.1	18.6	-8.1	15.4	14.9	3.8	4.1	21.4
Post-Pre Control Group n = 31	6.1	5	5.5	-1.5	6.1	-3.0	0	7.5	1.5	8.5	-1.8	6.1
Experimental-Control	-3.8	-6	-7.2	3.8	-14.2	21.6	-8.1	7.9	13.4	-4.7	5.9	15.3

Table 2 showed promising results on a few questions, 6, 9, and 12. Yet Looking at each experimental group compared to its control group the findings tell a slightly different story.

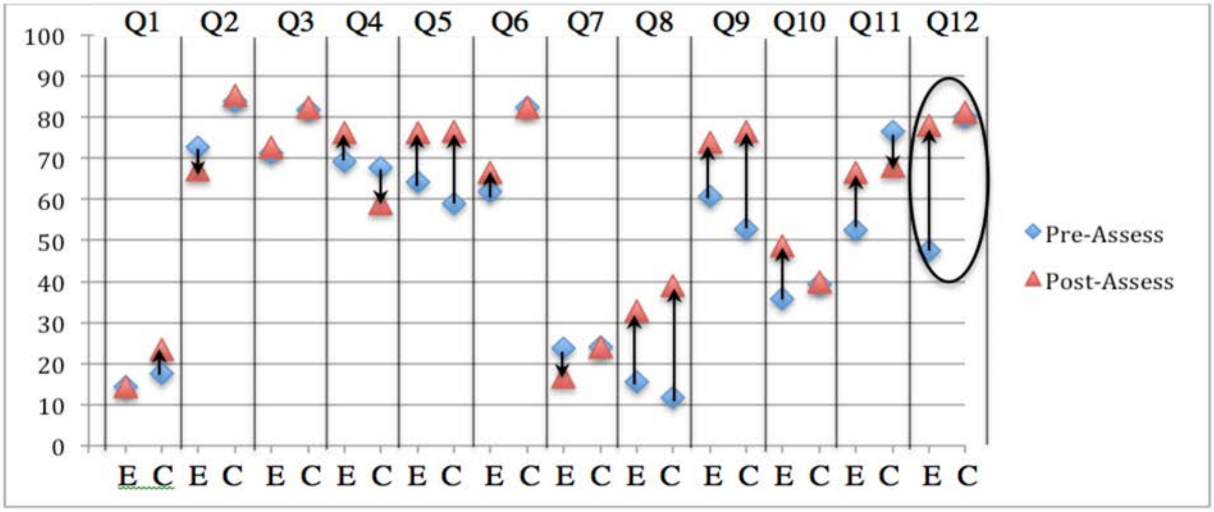

Figure 4: Teacher 1: Experimental pre & post-assessment scores compared to control pre & post-assessment scores for each quesion. Note: question 12 has significant improvement for the experimental group but the control group performance was already high and remained high.

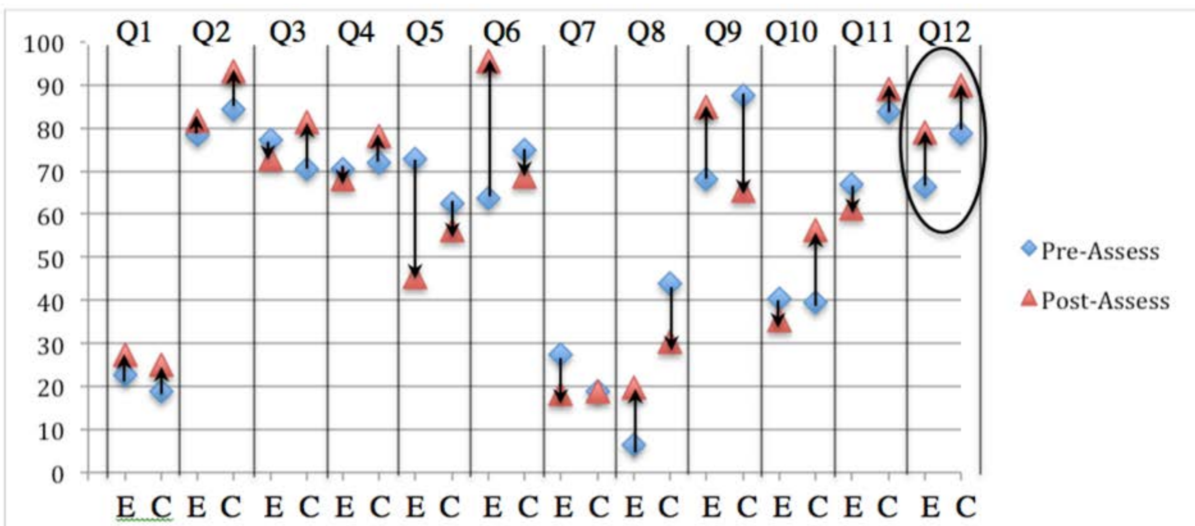

Figure 5: Teacher 2: Experimental pre & post-assessment scores compared to control pre & post-assessment scores for each question. Note: : question 12 shows experimental and control groups improving by about the same amount.

It is apparent from both graphs (Figures 4 and 5) that it is not possible to draw conclusions about any overall success of the experimental groups over the control groups on any particular question.

On the given assessment questions 2, 4, 6, 8, 9,10, and 12 fell into the category directly relating to an understanding of linear vs exponential change over time.

- The scores on question 2, draw the graph of the water level in a container when someone is filling (emptying) a container at a steady constant pace for 5 seconds, then stopping for 5 seconds, were pretty high. There is no reason the experimental group should have done more poorly than the control group. In point of fact, the experimental group for teacher 2 (with a lower semester 1 algebra II average) performed better on this question than the other experimental group (with the higher class average). Most of the errors involved students drawing a diagonal line up instead of down (or vice versa) and/or continuing the diagonal line for the second 5 seconds instead of drawing a horizontal segment. All students did, however draw linear segments and not exponential ones. So, for the purposes of the research question, the students did choose the correct function (linear) representation for their curves.

- Question 4 gave students a description of a person walking and asked them to select the correct distance versus time graph (given 4 choices). Most of the errors involved not paying close attention to the time duration associated with a specific walking strategy. So this question did not really differentiate between linear and exponential behavior, since all the choices were linear. It would have been better to include an exponential curve as part of one graph choice.

- Question 6 involved describing the value of the slope of an exponential graph over time. This was a good question for differentiating between linear and exponential functions, but, there was not significant improvement by the experimental group compared to both control groups on this question.

- Question 9 involved using a labeled linear STELLA model diagram and asking students if the stock value would increase or decrease, and then asked them to select one of 3 graphical patterns to represent the increase or decrease. This was a good question differentiating linear and exponential change for the experimental group, but was not really a fair question for the control group. The experimental groups improved, as was expected, but one of the control groups improved even more.

- Since students had insufficient time to complete all of the problems for the linear and exponential model lessons. Some important relationships (assessed, for example, in questions 8 – does doubling the growth rate double the growth amount in 5 years - or question 10 – will a linear payment plan or an exponential payment plan pay off a loan faster) were not experienced.

- Question 12 held great promise. It asked students to identify each of 5 short scenarios as representing linear or exponential change over time. Both experimental groups improved, one significantly (over 30%), but one control group had high scores already (and remained high) and the other control group improved almost as much as the experimental group for that same teacher.

- Questions 3 and 7 tried to assess how a quantity would change based on relative size of inflow versus outflow and so were not directly useful for the research question. With the problems that occurred on the drug lesson (the only lesson that contained both an inflow and an outflow that all students were required to complete) there was insufficient time to practice, and so to develop an intuition for that concept.

Data from The Lessons

Motion Detector: The graded question involved giving each student a graph containing two points, asking the students to connect the two points using between 2 and 4 linear segments, and then to explain the walking motion a person would have to complete to produce the graph drawn. 3 points were awarded for drawing a correct graph and 4 points were awarded for describing the walking motion correctly. The results of this assessment are shown below.

Table 3: Mean scores for each experimental class on the motion detector problem.

Experimental Groups	
Teacher 1 n = 20	89%
Teacher 2 n = 22	79%

Constructing an original linear STELLA model: The graded question required that students select a scenario that increased or decreased in a linear fashion over time. They were to sketch the appropriate STELLA diagram and label it (3 points), indicate the values and units for the stock and the flow (3 points), indicate whether the stock value would increase or decrease over time and how they could tell from the diagram (2 points), and how they knew the stock would change linearly (2 points). Note: students did not actually build this model on the computer.

Table 4: Mean scores for each experimental class on the linear modeling problem.

Experimental Groups	
Teacher 1 n = 20	80%
Teacher 2 n = 22	75%

Constructing an original exponential STELLA model: The graded question required that students select a scenario that increased or decreased in an exponential fashion over time. They were to sketch the appropriate STELLA diagram and label it (3 points), indicate the values/formula and units for the stock and the flow (3 points), value for the converter (1 point), indicate whether the stock value would increase or decrease over time and how they could tell from the diagram (2 points), and how they knew the stock would change exponentially (2 points). Note: students did not actually build this model on the computer.

Table 5: Mean scores for each experimental class on the exponential modeling problem.

Experimental Groups	
Teacher 1 n = 20	75%
Teacher 2 n = 22	76%

Students worked on the first 2 models for the linear and exponential lessons in pairs. That work would be considered at the "following directions" stage of the learning framework. The 3rd problem in each lesson, however, was to design an original growth or decay model of their choosing, label the model, indicate how the model should be defined (i.e., the values or formulas needed) and explain why the model would show growth/decay and why the stock should increase/decrease in a linear/exponential fashion. Success on these exercises would demonstrate that those students had progressed to the "create your own" stage of the learning framework, since they are transferring their understanding of linear and/or exponential structure to a new situation, defining the model correctly, and correctly indicating why the model should produce the desired pattern of growth/decay over time.

Building the drug model: Due to technical problems only the first 2 questions on the student packets were graded. The first question required that students read the description of the scenario and sketch the STELLA diagram that would capture the details of the scenario. (The scenario required a constant inflow and an exponential outflow for the stock.) The diagram was graded (4 points) as indicated in the table below (1 point for a diagram that contained an inflow and an outflow, 1 point if the inflow was constant and the outflow was exponential, 1 point if the diagram was labeled correctly, and 1 point if the correct values were indicated for each icon). Students were then to predict what they thought the graph of the stock value would be over time. Then students were to build the model on the computer and copy the simulation results on the same grid as the prediction graph. The 2 graphs were given a total of 1 point.

Table 6: Mean scores for each experimental call on the sub-parts of the first and second problem of the drug model lesson.

Experimental Groups					
	Stock, In/outf	Con in, ExpOu	labels	values	Pred Run
Teacher 1 n = 21	97%	80%	34%	39%	57%
Teacher 2 n = 17	68%	62%	26%	32%	44%

Theoretically, those students who correctly designed the drug model and simulated it would have moved to the "transfer" stage of the learning framework because they demonstrated that they understood linear and exponential behavior sufficiently to create a diagram that contains a constant flow (and linear behavior in the stock, if there were no outflow) and an exponential flow. Moreover, they understood that the outflow had to be exponential and the inflow constant. This would satisfy the "transfer" stage since students are applying their understanding of linear and exponential behavior to a completely new scenario involving both functions, something they had not dealt with in the past.

Results of Videotape of two pairs of students from each experimental class

Additionally, for the drug model, two pairs of students (one from each experimental class) were videotaped, separately, using a think-aloud task-structured protocol to determine how they thought about the model as they were creating it. The videotapes were transcribed and analyzed using Lesh and Lehrer's (2000) videotaping analysis framework to determine if students understood the difference between linear and exponential functions and could interpret a more sophisticated problem involving both types of functions.

The pair was working on the drug lesson (and being videotaped by the researcher) at the same time their algebra II class was working on the same lesson with their algebra II teacher. Each student had a "Drug Model Lesson" packet to read and work from for this exercise. Students were to build and exercise the drug model while using think-aloud-protocol.

1. Overall Purpose 1: To capture the thinking of the students as they decided how to construct the stock/flow diagram to capture the behavior of the following problem:

    "You are continuing your work as a medical resident at a local hospital. You are again working in the emergency room when a patient arrives and needs medical attention. For this patient you decide you must insert an IV drip into the person's arm in order to administer a therapeutic drug. You set the IV drip so it will allow a constant inflow of 1 g/min of the drug into the person's blood system. The patient, you estimate, will eliminate 0.55 % of the drug in his system each minute. (Be careful, 0.55% is less than 1%)."

a. A subordinate purpose was to determine if the students could construct a stock/flow diagram that had the students select a stock and identify it as the amount of drug in the body, and construct a constant inflow toward the stock and an exponential outflow from the stock. Note that the generic exponential growth and exponential decay stock/flow diagrams were shown in a boxed display at the top of the paper, for student reference.

- Team 1 produced a correct stock/flow structure on their initial attempt.

- Team 2 started out with an incorrect inflow structure (exponential instead of constant), but realized since the inflow was a constant they need to remove the exponential inflow structure and replace it with a constant flow structure. (They redrew the inflow correctly.) They also defined the model components correctly.

2. Overall Purpose 2: To determine if students can explain what is happening with the dynamics of the problem that causes the shape of the stock graph to be produced when the model is simulated? That is, can the students interpret the model output, relating it to the real world problem?

a. It is typical in System Dynamics modeling lessons to request that students anticipate model behavior before simulation runs are executed. Most students, initially, have a great deal of trouble doing this, as it is not something that is typically asked of them in their math classes. This situation was no different.

- The first pair of students drew an exponential growth curve. The curve was incorrect.

- The second pair created a linear graph where they moved by blocks, in time, and increased the next point vertically by 1 block but then subtracted .5 of a block from that, creating a line that had essentially a slope of .5 of a block. This curve was incorrect on many levels. (Note: The default STELLA graph pad is divided into 16 grid blocks covering the entire domain and range for a given model component output.)

b. Students were then asked to explain the result the simulation produced (whose appearance showed an exponential convergence upward from 0 to about 173 g, reaching steady state about 1/3 the way through the simulation run).

- The first pair of students reached a reasonable conclusion. In the packet the students wrote: "The medicine being eliminated slowly approaches the value of the medicine entering but never will reach the same value or decrease." Unfortunately, they did not explain why the medicine level should not decrease. But it is a valuable insight that they did consider this option.

- The second pair, after exclaiming how different the simulation graph was from their prediction, came to a reasonable explanation for the shape of the simulation graph.

The students observed the graph, then set up a table with the drug in the system, the inflow value and the outflow value. In the packet the students wrote: "The inflow is constant but the outflow is exponential so as more drugs enter the system the outflow at a rate of .55% the outflow will grow to be the same as the inflow."

3. Overall Purpose 3: The final objective was to determine if the students were becoming comfortable enough understanding the model dynamics that they could correctly predict what would happen to the graphical model output if a modification of the model was made. The following description was given to the students:

   "A complication occurs with this patient about 8 hours after the IV is administered. One of his kidneys quits functioning, causing the elimination rate to reduce to half. Predict what you think will happen to the drug level in the patient's body, recording your prediction on the grid below."

Both pairs correctly predicted the new behavior and were able to identify where the model needed to be modified, and how the definition of the selected component needed to be modified.

- Team 1: [The students ran out of time. They did not have time to see the simulation run, so they were not able to check their prediction. Both students drew prediction graphs in their lesson packet that indicated an initial jump upward at 8 hours followed by a leveling off of the drug at a higher equilibrium. This is the correct behavior the model would have produced.]

- Team 2: These students drew both the prediction (using a dotted curve) and the actual simulation run (using a solid curve). The dotted curve increased in a more gradual fashion, indicating a smoother upward transition at 8 hours, than the simulation graph that made a more pronounced upward jump at 8 hours.

The transcripts[6] indicate that the students were able to mathematize the problem without much difficulty, even after a faulty start with the second pair. They were able to use the software to explain the dynamic behavior represented in the simulation graph of the initial model. That was a key point. Once they had an understanding of the cause of the dynamic behavior pattern of the original model they were able to modify the model and were reasonably successful predicting the new behavior of the modified model. All of this work

[6] Conversation clips from the transcripts can be found in the supplemental file - Appendix C.

was on a model whose behavior was not typical of functions they had seen in class at that point.

The student problem solving scenario captured in these two videotapes supports the claim that students are able to mathematize new scenarios that are combinations of behaviors they already know and are able to analyze and modify and reanalyze the problem with relative ease. The videotaping lasted about 33 minutes for each pair of students.

This analysis was done comparing three problems across both student pairs (horizontal analysis) and also analyzing the improvement in thinking over the full time frame (vertical analysis) for each pair of students.

Algebra Questionnaire

A Likert scale was used to give students in the experimental groups an opportunity to express how useful they felt the lessons were in helping them understand the concept of linear and exponential functions. The lowest indicator was labeled “Not very helpful” and the highest indicator was labeled “Very helpful.” There were five points on the scale. Students were to choose one of the five points to assess each lesson (motion detector, building linear STELLA models, tape-walking floor activity, building exponential STELLA models, building a STELLA bank model with the teacher, building a STELLA drug model).

Table 7: Likert scale results, student opinions about the lessons.

Activity	Teacher	1 (Not_very helpful)	2	3	4	5 (Very helpful)
1. The motion detector activity	1 2	1 4	2 3	10 6	10 5	2 1
2. Building linear models on computer with STELLA	1 2	2 6	1 3	11 3	5 6	6 1
3. The walking into/out of the box made with tape on the floor	1 2	0 6	3 2	13 5	5 5	4 1
4. Building exponential models with STELLA	1 2	2 5	2 6	11 2	7 5	3 1
5. Building bank model in class with the teacher	1 2	1 4	0 3	8 5	8 6	8 1
6. Building the drug model with STELLA	1 2	1 5	1 5	7 5	12 3	4 1

Students were then asked to make comments about the lessons in general.

Table 8: Student comments about their impression of the usefulness of the experimental lessons.

Statement	Teacher	Negative	Neutral	Positive
Comments about items 1-6	1 2	3 5	2 0	9 3
7. From these six lessons, did you feel you learned important/useful information about linear and exponential functions?	1 2	5 12	6 2	14 5
8. Do you think the STELLA models helped you understand why the graphs of the situation you built models for had a linear or exponential shape?	1 2	3 7	1 3	19 8
9. Do you think this activity (all 6 lessons) were worth doing?	1 2	8 12	3 1	14 6
Additional Comments	1 2	0 3	0 0	4 0

Of the comments made regarding question 1-6: 55% were positive comments (enjoyed, liked, fun, helpful), 9% were neutral (wanted more interaction with the teacher, depends on the person), and 36% were negative (took more time than it was worth, didn't understand, confusing).

Of the comments made regarding question 7: 43% were positive (yes: reasons varied from gained better understanding because I am a visual learner, liked the problems, it was hands-on). 18% were neutral (I already knew linear and exponential, learned a little). 39% were negative (no: I already knew this, confusing, not taught well, easier to just use numbers).

Of the comments made regarding question 8: 66% responded in the affirmative (it was visual so easier to understand, the models were interactive, I could see what was going on, I already understood the concept but now I know it from a different perspective), 10% were neutral (I already understood the concept), and 24% were negative (it was confusing, I still don't understand them, it was unclear).

Of the comments made regarding question 9: 45% said yes (fun, it definitely improved my understanding, they all tied together, because I learned something that actually might be useful in my life), 16% were neutral (timing was bad – right before finals, useful now but probably not in the future), and 45% were negative (no impact on my learning, hard to understand, felt like we were rushing, not practical, confusing).

Validity

This procedure (the six lessons) is assumed to have content and face validity because the lessons deal specifically with linear and exponential function concepts and have been used successfully with algebra II students in the past. The procedure is also assumed to have construct validity, hoping to support a more conceptual understanding of linear and exponential functions. The definition used to support evidence of conceptual understanding is 1) whether students can extend the models built in class to include a combination of those functions (linear and exponential), 2) whether students can comfortably transfer the linear and exponential scenarios they studied in the model-building activities 2 and 5 to a drug model scenario, and 3) whether students can correctly modify the initial drug model they create at the beginning of activity 6 to solve additional problems presented later in the drug lesson. Note: The pre- and post-assessment is assumed to have reliability due to the use of the cross-balancing[7] process for distributing the two forms of the assessment.

C. Discussion and Lessons Learned

The Assessment

- The extreme compression of time for the six lessons (one of which – the graphing lesson - was not able to be executed in the experimental group for teacher 2) impacted the overall results of the post assessment, as students did not have time to complete any lesson and there was no time to provide feedback to the students on their work before moving on to another lesson. Also, students who missed a class did not have an opportunity to make up the lesson due to the very short time frame for the experiment. Consequently, the full post-assessment scores do not represent a reasonable assessment of overall progress toward shedding light on the research question.

- For any possible meaningful results from the total scores the experiment would have to be conducted over a longer period of time.

- It was a mistake to do this experiment as a short unit (even if we had been able to dedicate a full three weeks to the unit) rather than infusing it into the regular curriculum. It seemed as if most students did not place much value on the new approach because they knew they would not see it again after the experiment concluded. Also, a short unit experiment, while easier to design and implement, gives short shrift to the fact that SD analysis is a way of *thinking* about problems. It needs incubation time to be fully realized.

[7] Half the students received form A and half form B for the pre-assessment. Then the form was reversed to use for the post-assessment for each student.

- To perform statistical test item confidence analysis it is important to ask questions that have at least 5 different possible scores. This was not considered in advance of the execution of the experiment.

- It is important to view the data using multiple perspectives. Combining the experimental classes and control classes in order to achieve a larger "n" value to assist statistical analysis actually hid important relationships that ultimately negated the assumptions that the collective data seemed to present.

The Lessons

- STELLA linear and exponential model building seems to narrow the gap in performance between students who traditionally perform well compared to students who traditionally perform more poorly in algebra II. Evidence for this statement was based upon comparing the class averages on the linear and exponential graded assignments for the two experimental groups whose performance on their first semester algebra II course was different by 14 percentage points.

- The results of the drug model lesson completed by the two randomly selected student teams who were videotaped were quite positive. Even the students who created an incorrect model initially were able to correct their mistake themselves as they built the model. Those results provide promise that, had the two classroom attempts to complete the drug model lesson not succumb to so many technical problems,[8] there might have been an opportunity to determine whether more students were able to transfer what they learned in the linear and exponential modeling lessons to a new, more complicated scenario.

The Questionnaire

- Students who generally had more success with math evaluated this very quick experiment in a more positive light than the students who were generally less successful. This does not mean the lessons could not have altered the negative view had the students been given more time to work with the software. The experiment was rushed and did not provide time for students to complete lessons, nor was there time to provide feedback about their work to the students, both situations that are felt more strongly by students who are already uncomfortable with regard to course content.

- Although the teachers chose the placement of the experiment in the school calendar, some of the students indicated that they were not pleased that the experiment was conducted just before semester final exams. They wanted to work on lessons that specifically prepared

[8] The netbooks were old. Quite a few of the computers did not function properly. By the end of the experiment, the STELLA software, whose image had to be "pushed" onto the computers from the central district office, was missing from a significant number of the computers that were still operational.

them for the semester exam. The students knew their course grades would not be impacted by their performance on any of the experiment assessments.

Other

- There are other issues that should be researched. Another article by this author documenting this experiment from the perspective of the teacher is currently in the review process with a mathematics education journal. It includes an analysis of teacher beliefs, teacher reflections during the experiment and also clarifies why the two teachers, both of whom have found SD modeling useful for their students, no longer use it in their curriculum.

D. Conclusion

It is becoming abundantly clear that adults need to gain facility dealing with complex systems scenarios as informed citizens or even when making decisions about social interactions. It is becoming clear that some educational leaders also feel that educating people about complex systems should start in mainstream education, with students in K-12. Progress has been made in K-12, but research supporting the improvement in learning outcomes for students at this level is sorely lacking. The experiment described in this paper encountered some of the problems mentioned by researchers who have conducted classroom experiments with SD in the past. The difficulty with the technology and the extremely compressed time frame for executing the experimental lessons led to results showing no statistical difference between the experimental and control groups on the pre- and post-assessments. However, there were promising results from data gathered on the lessons themselves, indicating SD modeling may provide a useful method of closing the gap somewhat between students who are adept at math and those who are less adept, when analyzing dynamic problems. The results of the videotaped think-aloud protocol used with a pair of students, randomly selected from each of the experimental classes, as they built a drug model indicate that students are able to successfully build and analyze a model that is slightly more difficult than the typical problems studied in algebra II. The experiment should be repeated over a more extended time frame where the experimental lessons are infused into the mainstream curriculum. The experiment was intended to suggest a path from the current topics covered in algebra toward those topics that are more illustrative of complex systems. In this vein, the current experiment can add value to the current literature.

The student work produced in classrooms where teachers are already using Systems Thinking and System Dynamics modeling is impressive. This work will be largely ignored until documentation can be provided showing an improvement in learning outcomes. This improvement will probably not be captured by current standardized tests. Further work needs to devise appropriate assessment instruments to capture the more holistic thinking, attention to feedback analysis, structure/behavior connections, and increased depth of understanding evidenced in student explanations when they are using systems tools. We should have a discussion about what behaviors we hope to see in students studying systems that sets them apart from students in other more traditional classes. We should have a discussion about how to structure experiments to guide our studies to give them the best chance to produce results that we can document. I invite those researchers who are interested in pursuing this path to build from the mistakes of this experiment and then tell us how much further they were able to progress, than I have. We can do this. It is, after all, a complex systems problem.

References:

Bruner, J. S. (1966). *Toward a theory of instruction* (Vol. 59). Cambridge, MA: Harvard University Press.

Chi, M. T., DeLeeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439-477.

Clement, J., Lochhead, J., & Monk, G. S. (1981). Translation difficulties in learning mathematics. *The American Mathematical Monthly*, *88*(4), 286-290.

Daniels, H., Cole, M., & Wertsch, J. V. (2007). *The Cambridge companion to Vygotsky*. Cambridge, NY: Cambridge University Press.

Davidson, N. (1990). Small-group cooperative learning in mathematics. In T. J. Cooney & C. R. Hirsh (Eds.) *Teaching and Learning Mathematics, 1990, Yearbook of the National Council of Teachers of Mathematics* (pp. 52-61). Reston, VA: National Council of Teachers of Mathematics.

English, L. D. (2007). Complex systems in the elementary and middle school mathematics curriculum: A focus on modeling. In B. Sriraman (Ed.) *The Montana Mathematics Enthusiast Monograph: Vol. 3. Festschrift in Honor of Gunter Torner* (pp. 139–156). Charlotte, NC: Information Age Publishing.

Forrester, J.W. (June, 2009). *Learning through system dynamics as preparation for the 21st century*. Paper presented at the K-12 Systems Thinking and Dynamic Modeling Conference. Retrieved from http://static.clexchange.org/ftp/documents/whyk12sd/Y_2009-02LearningThroughSD.pdf

Homer, J. B., & Hirsch, G. B. (2006). System dynamics modeling for public health: Background and opportunities. *American Journal of Public Health*, *96*(3), 452-458.

Hung, W. (2008) Enhancing systems thinking skills with modeling. *British Journal of Educational Technology*, 39 (6), 1099-1120.

Jacobson, M., & Wilensky, U. (2006) Complex systems in education: Scientific and educational importance and implications for the learning sciences. *The Journal of the Learning Sciences*, 15 (1). 11-34.

Kaput, J. J. (1999). Teaching and learning a new algebra. In E. Fennema & T. A. Romberg (Eds.) *Mathematics Classrooms that Promote Understanding*. Mahwah, New Jersey: Lawrence Erlaum Associates, Publishers.

Kaput, J., Bar-Yam, Y., Jacobson, M., Jakobsson, E., Lemke, J., & Wilensky, U. (2000). Two roles for complex systems in education: mainstream content and means for understanding the education system itself. *Planning documents for a national initiative on complex systems in K-16 education*. Retrieved from http://www.necsi.edu/events/cxedk16/cxedk16_0.html.

Keller, B. A., & Hirsch, C. R. (1998). Student preferences for representations of functions. *International Journal of Mathematical Education in Science and Technology*, *29*(1), 1-17.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of educational research*, *60*(1), 1-64.

Mandinach, E. B., & Cline, H. F. (1993). Systems, Science, and Schools. *System Dynamics Review*, 9 (2), 195-206.

Mandinach, E. B., & Cline, H. F. (2000) It won't happen soon: Practical, curricular, and methodological problems in implementing technology-based constructivist approaches in classrooms. In A. P. Lajoie (Ed.) *Computers as Cognitive Tools, Vol II*, (377-395). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.), *Making the Connection: Research and Practice in Undergraduate Mathematics* (pp. 27-42). Washington, DC: Mathematical Association of America.

Quaden, R., Ticotsky, A., & Lyneis, D. (2008). *The Shape of Change and The Shape of Change: Stocks and Flows*. Acton, MA: The Creative Learning Exchange.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, *69*(1), 21-51.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Boston, MA: Irwin/McGraw-Hill.

Sterman, J. D. (2011). Communicating climate change risks in a skeptical world. *Climatic Change*, *108*(4), 811-826.

Vygotsky, L. (1978) Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.) *L.S. Vygotsky Mind in Society: The Development of Higher Psychological Processes*, Ch 6. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1997). The history of the development of higher mental functions. In R. W. Rieber. *The Collected Works of L. S. Vygotsky. Volume 4: The History of the Development of Higher Mental Functions* (pp. 1-26). New York: Plenum Press.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. *Journal of Computer Assisted Learning*, *7*(1), 42-57.