What can students learn from simple simulations about accumulations?

Krystyna A. Stave, Ph.D.
School of Environmental and Public Affairs, University of Nevada Las Vegas
4505 Maryland Parkway Box 454030
Las Vegas, Nevada 89154-4030
krystyna.stave@unlv.edu

Abstract

This paper presents the second phase of research about designing and testing the effect of systems simulations for building systems understanding. It builds on work presented last year in which we used a paired experiment in an introductory level college course to examine the effect of a simulation on understanding of simple accumulation principles. Previous results showed significant differences in some measures of understanding of systems principles but also highlighted issues with simulation design, comparability of subject groups, and measures of systems understanding. In this phase, we revised the simulation, learning measures, and study design. All students used the simulation. We compared the extent to which students interacted with the simulation with their performance on a set of systems thinking measures. Pre-test/ post-test measures showed strong improvement in understanding of accumulations principles among users who ran the simulations a moderate number of times, but analysis of the data shows it is not a linear relationship. Mid-range users (total run count between 10 and 20 runs on two different simulations) did significantly better than both low-range (1-9 runs) and high-range users (20+), indicating the simulations improved scores, but that there may be both a threshold and a saturation point in the effect of simulations on systems learning.

Introduction

This paper describes the continuation of previous work examining the potential of systems simulations for building systems understanding (Skaza and Stave 2010, Stave 2011). The first phase of the work used a paired experiment in an introductory level college course to examine the effect of simulations on understanding of simple accumulation principles. One section of students used simulations in their homework assignments and one section of students was given the same material in a standard non-simulation homework assignment. Regression results showed significant differences in some measures of understanding of systems principles between students in the simulation (treatment) group and the non-simulation group. However, the previous study highlighted some issues with simulation design, comparability of subject groups, and measures of systems understanding.

For this study, we modified the simulation design and systems learning measures. More importantly, we changed the structure of the study. Instead of using paired groups of students, with one receiving the simulation and one not, all students in this phase used the simulation. We measured the extent to which students interacted with the simulation and how they navigated through the simulation against their performance on a set of systems thinking measures. We asked: given minimal experience with stock and flow behavior, can students develop an operational understanding of accumulation principles? If so, what aspects of the simulations facilitate their learning?

Our overall motivation is to shed light on best practices for using simulations to facilitate development of intuitive understanding about basic principles of accumulations. Ultimately, we would like to have an approach for building operational systems understanding that could be used in a variety of settings, with a variety of audiences. It should not require any formal training in systems concepts, or any particular background in science or math.

Method

We have been using an introductory college level environmental science course as the setting to develop and test the simulations since Fall 2009 with progressive refinements of simulations and assessments. We are examining both the effect of the simulations on systems understanding and the ease of use of the simulation design. Although these two aspects are clearly related and make the study somewhat messy, the applied setting provides opportunities for insight about "real-world" audiences that are not always anticipated.

This paper reports the results of the study from Fall 2011. Two sections of the course were run exactly the same way. The same instructor taught the lectures for both sections, classes were held in similar sized classrooms at roughly the same time of day, with a total of 151 students. Of those, 136 students completed both the baseline and final assessments. The students in both sections were demographically similar. Approximately 85% of the students were between 18 and 24 years old. Nearly half (46%) said they were taking the class only to fulfill a general education science requirement or because it was required by their major. Only 16% reported taking the class primarily because they were interested in the subject. Almost half (45%) the students were social science and humanities majors, 34% were hospitality and business majors, 10% had not yet declared majors. Only 2% were environmental studies majors and 8% were science and engineering majors. 60% of students had not taken any previous science courses.

The study is a pre-test, treatment, post-test design, in which the treatment consists of two simulation exercises in the course of a 16-week semester. Figure 1 shows the timeline of assignments and assessments during the term. We conduct a baseline assessment before the first content lecture of student knowledge of the course content, plus their ability to read graphs and apply basic systems concepts.

During the course, we give no formal lectures or instruction on systems concepts. The only direct interaction students have with systems principles is in two self-guided assignments during the term. The treatment consists of two stand-alone assignments. Students access the assignments from the course website. Both assignments, shown as Assignment 2 and Assignment 3 in the timeline below consist of two parts: the first part is a simulation hosted on the Forio website and the second part is a set of graded questions they answer on the course website after they have completed the simulation. The in-class quizzes during the term each include 1-2 questions testing systems knowledge, and a systems question is included on the final exam.

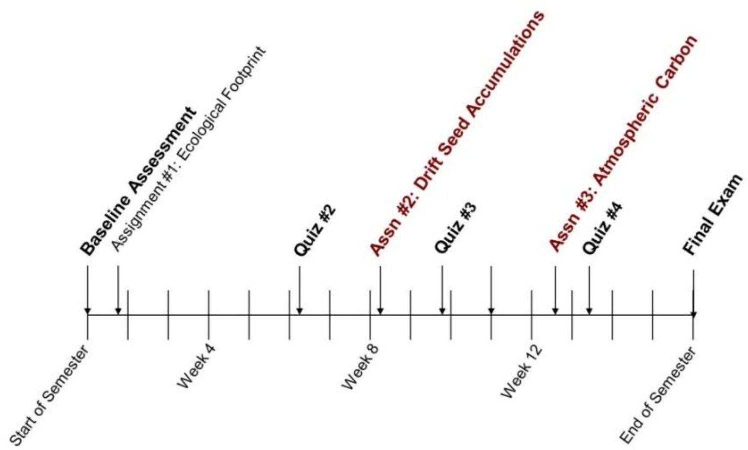


Figure 1 Timeline of Assignments and Assessments

Simulations

The simulations are based on discovery learning principles (Stave 2011). They present a hypothetical situation and allow students to experiment with the simulation to achieve a given task. Students are not expected to have any specialized knowledge to use the simulation. Simulations were based on simple system dynamics models developed in Vensim and run on the Forio Simulations platform.

The first simulation, **Drift Seeds on the Shore**, is a simple one-stock, two-flow model representing the accumulation of drift seeds on a hypothetical island. Seeds wash in to shore and wash out to sea, accumulating on the island. The student's task is to adjust the rates of inflow and outflow to achieve a sustainable level of seeds on the island. Figure 2 shows the introduction screens and simulation interface for the Drift Seeds simulation.

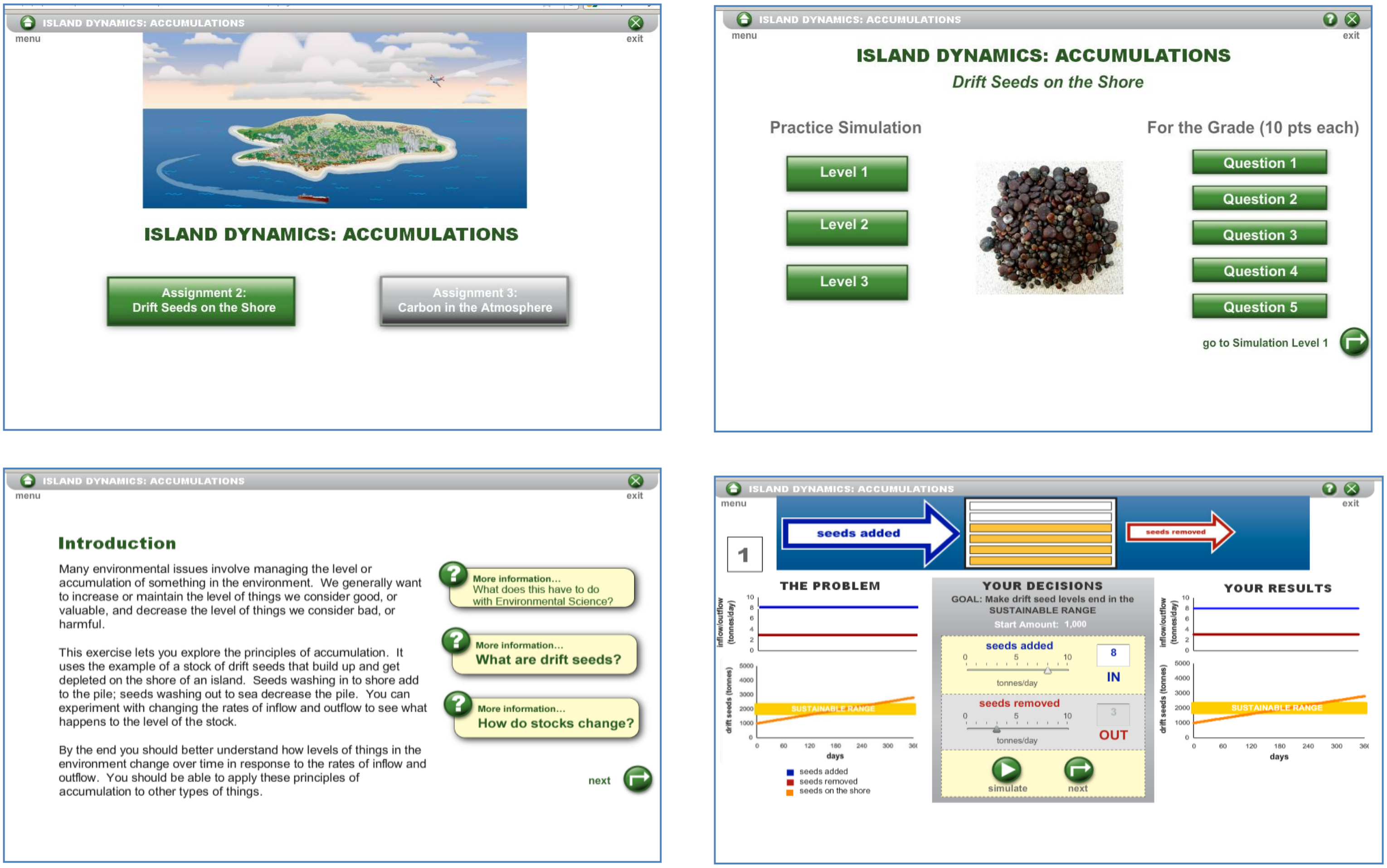


Figure 2 First three introductory screens of Drift Seeds simulation, plus simulation interface

examine the effect of changing emissions and removal rates on carbon accumulation. This simulation also breaks the inflow into two parts – one representing large-scale emissions and one representing emissions from individual-scale activities. The task is the same, however, adjusting the rates to achieve a sustainable level of carbon in the atmosphere. Figure 3 shows selected screens for the Carbon simulation.

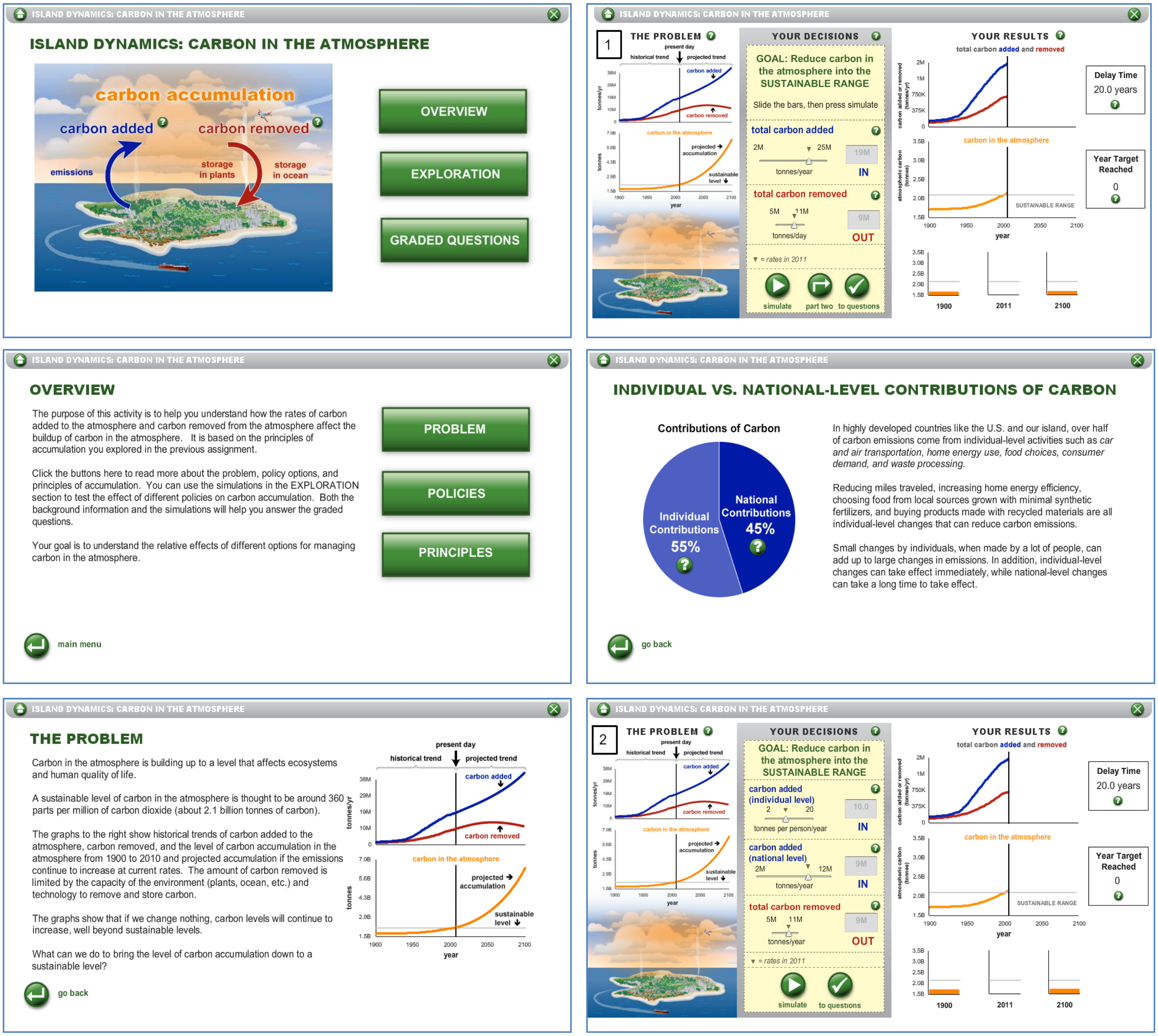


Figure 3 Selected interface screens from Carbon Simulation

Students are not required to read the background information provided under the “OVERVIEW” button. They can go directly to the simulation under “EXPLORATION” or skip the simulation altogether and go first to the quiz under “GRADED QUESTIONS”. An explanation of the principles of accumulation is provided in the Drift Seeds simulation under the “How do stocks change?” button, and in the Carbon in the Atmosphere simulation under the “Principles” button. The information is shown in Figure 5. In our experience, most students do not read the background information before attempting the simulation or quiz questions. In both simulations users can return to the principles information from the simulations or questions, and those who do read the information tend to do so after attempting the quiz questions.

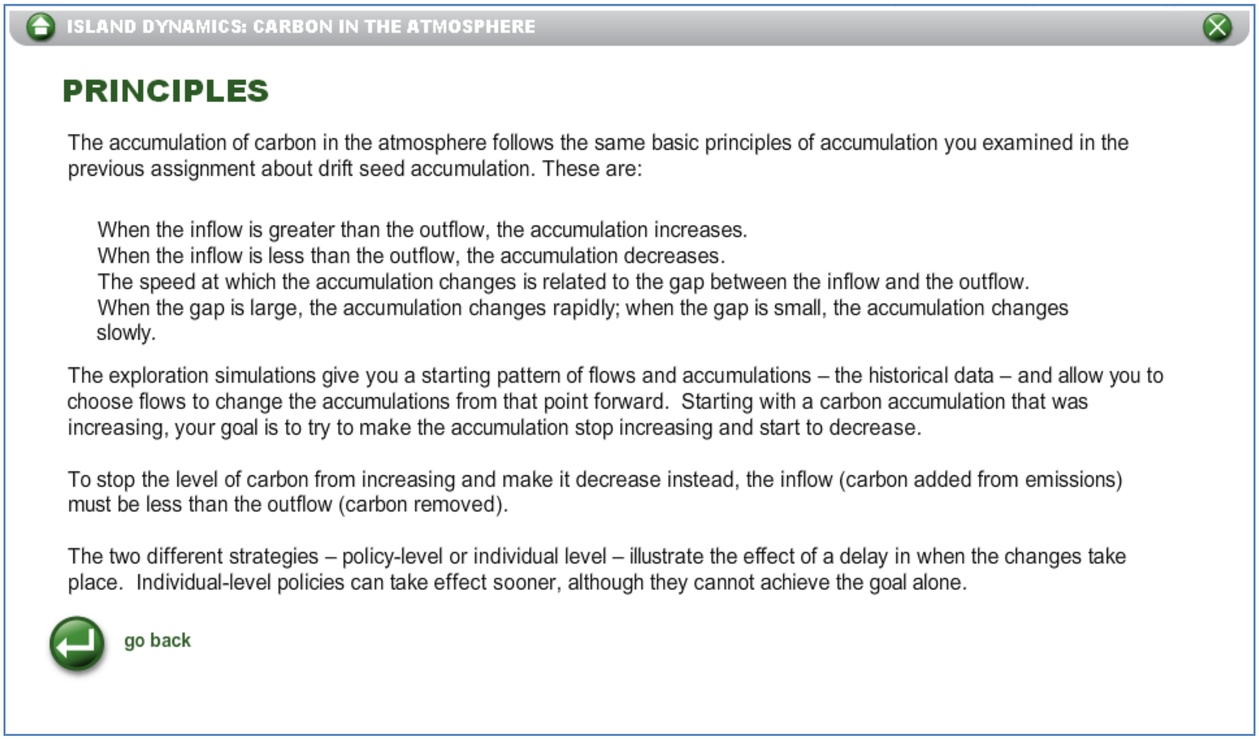


Figure 5 Optional Principles of Accumulation information page in Carbon simulation

Measures

All students were required to take the baseline assessment and were graded on the number of questions they answered, not the answers themselves. For full credit, students were required to mark an answer for all questions, but each question included an option to answer “I don’t know.” Students were told that know particular knowledge was expected at the beginning of the course, but that the baseline information would be used to help tailor the material to their interests and concerns. Any question that was considered to be possibly embarrassing or uncomfortable included a “Prefer not to answer” response choice. The baseline measured a wide variety of things for the purpose of the course, including environmental knowledge, subject area interest, environmental attitudes and behaviors, as well as student ability to understand graphs and their grasp of basic systems concepts, which was used for this study. We wanted to be able to separate poor graph-reading skills from measures of systems understanding. Students completed the baseline assessment on their own on the course website.

Baseline: ability to work with graphs

Six multiple choice graphing questions measured two different graph skills: identifying specific points on a graph and identifying specific trend lines relative to other lines. Figure 8 shows one of the point identification graph questions. Figure 9 shows one of the trend line questions.

We also used a modified version of Sterman's People In the Store graph and questions (Cronin et al. 2009), in which the first two questions serve as measures of graph skills and the second two measure understanding of accumulation principles. Figure 7 shows the modified graph and questions.

The number of correct answers on the graph questions was summed to create an overall **Graphing Score** with a maximum possible value of 6.

Baseline: systems understanding

Seven multiple choice questions measured systems understanding: two questions on population change, three questions on carbon accumulations, and the two accumulation questions in the People In the Store problem. The total number of correct responses was the baseline **Systems Score**, with a maximum possible value of 7.

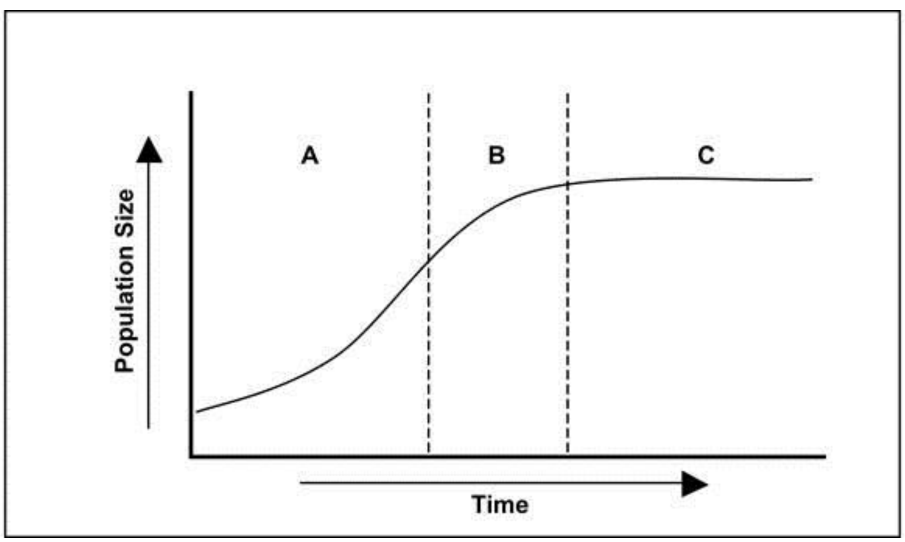


Figure 6 Graph for systems questions 1 and 2

Figure 6 shows the graph of population change used to test understanding of the relationship between the change in a stock and changes in the related flows. The two questions below were asked about Figure 6.

SQ1: How is the relationship between birth rate and death rate changing in section A (of Figure 6)?
SQ2: How is the relationship between birth rate and death rate changing in section B?

a. Birth rate is increasing and/or death rate is decreasing; they are getting further apart from each other.
b. Birth rate and death rate are the same as in part A; the relationship is not changing.
c. Birth rate and death rate are equal.
d. Birth rate is decreasing and/or death rate is increasing; birth rate and death rate are getting closer together.
e. I don't know.

Three questions were asked about carbon accumulation:

SQ3: Carbon accumulation in the atmosphere is a growing concern. Suppose the carbon emission rate (the rate at which carbon is added to the atmosphere) and the carbon removal rate both remain constant over a period of time, but the rate at which carbon is removed from the atmosphere is greater than the rate at which carbon is emitted. What will happen to the total amount of carbon in the atmosphere over this time period?

a. It will remain constant at a very low level.
b. It will remain constant at a very high level.
c. It will increase over time.
d. It will decrease over time.
e. None of the above.

SQ4: Carbon in the atmosphere is currently increasing rapidly. Which of the following could be true about the relationship between carbon emissions and carbon removal?

a. Carbon emissions and removals are both increasing, with removals higher than emissions.
b. Carbon emissions and removals are both decreasing, with removals lower than emissions.
c. Carbon removals are decreasing, and are consistently more than carbon removed.
d. Carbon emissions are increasing, and are consistently less than carbon removed.
e. None of the above.

SQ5: To reduce the amount of carbon in the atmosphere, we need to …

a. reduce the amount we add to the atmosphere each year by about 10 percent.
b. do nothing; the level of carbon in the atmosphere is decreasing naturally.
c. make sure the amount added to the atmosphere is less than the amount that is removed.
d. It is not possible to reduce the amount of carbon in the atmosphere.
e. None of the above.

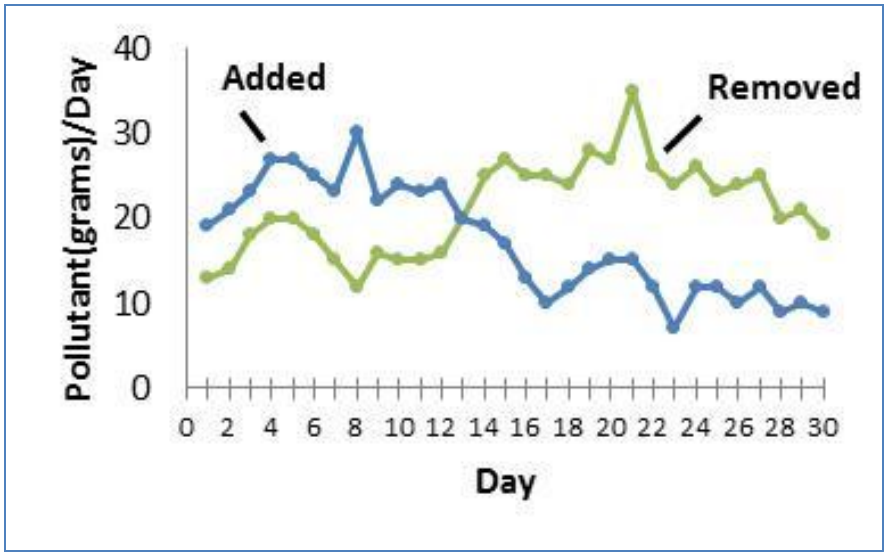


Figure 7 PINS problem graph modified to represent Pollutant in a Pond

In addition to the two graphing questions (GQ1 and GQ2), the two systems questions were asked as follows for Figure 7:

GQ1: On what day was the most pollutant added to the pond?
GQ2: On what day was the most pollutant removed from the pond?
SQ6: When was the most pollutant in the pond?
How did you determine your answer?
SQ7: When was the least pollutant in the pond?
How did you determine your answer?
Describe what happens to the amount of pollutant in the pond over the entire time period.

We also added three qualitative questions to the PINS problem, asking how respondents determined their answers to the systems questions, and to describe what happened to the accumulation over the entire time period.

Students took the baseline as an online assessment through the course website (WebCampus platform.) Students were not allowed to revisit questions once they moved on.

Intervention: simulation use

The Forio platform allowed us to capture data about how students used the simulations. For this analysis, we tracked how many times they ran each simulation (**Run Count**). For the second simulation (Carbon) we also tracked the number of times they visited different screens, including the page that presented the accumulations principles (**Principles Page Visits**).

Since each simulation contained graded questions within the simulation plus graded questions on the course website, we also measured their performance on systems questions directly after the assignment. Some students did not access the simulation at all, but completed the second part of the questions. These students provide important data, essentially representing a self-selected "non-treatment" group.

Final Exam: systems understanding

We asked the two-part question shown below on the Final Exam (**Water in the Pool**). Students had not seen this pattern of flows in the course of the term or in the simulations and they had not been asked to draw the level of the stock based on the flows. The pattern of flows is a simplified version of the PINS graph, however, and was developed as a way to compare performance on similar questions without repeating the baseline question.

The drawing and text description were coded using the coding scheme shown in the appendix. A grade out of 5 points was assigned to each part of the question (graph and text description) based on the level of understanding of accumulation principles demonstrated by each separately. The total grade is the sum of the graph and the text scores. The maximum grade of 10 points demonstrates the respondent's ability to both apply and explain accumulation principles.

Results

The punch line: what students learned from simple simulations about accumulations

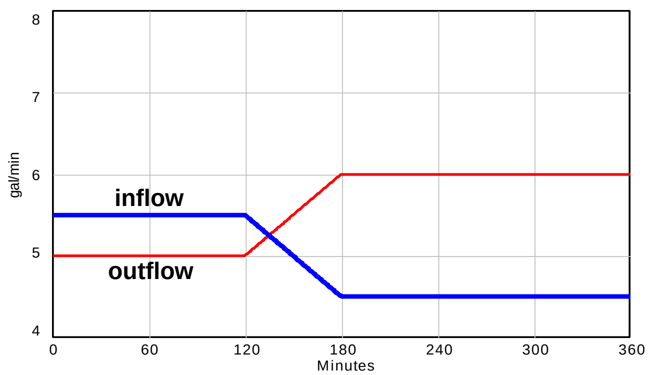
Student understanding of and ability to apply basic principles of accumulation increased markedly between the baseline and final assessment. By the end of the course, after having been introduced to accumulation principles only in two self-guided simulation assignments, 43 of the students (32%) got a perfect score, demonstrating a solid grasp of how the relationship between inflows and outflows affects the level of a stock. The average grade for 136 students was 6.29 out of 10 points.

The drawings and text descriptions together show strong evidence of their understanding. Forty percent drew the accumulation graph completely correctly, with the first section increasing, the second section decreasing, with the end of the graph lower than the start. 34% explained the reasons for the behavior completely correctly (relationship between stock and flows). By contrast, only 6% answered the PINS question "when was the most pollutant in the pond?" correctly at the beginning of the course and 8% answered the question "when was the least pollutant in the pond?" correctly. Only 2% answered both correctly.

FINAL EXAM QUESTION

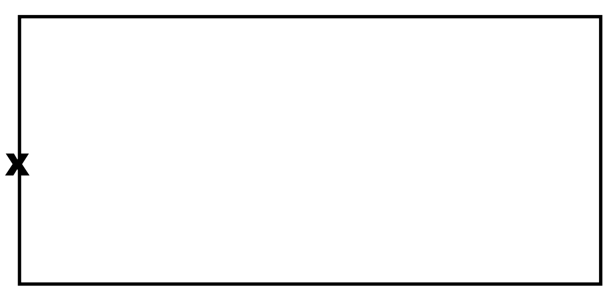
The graph below shows the pattern of water flow into and out of a pool over time. We could imagine that this "pool" is really water in Lake Mead, for example. Understanding the relationship between the flows and the amount of water in the pool helps us think about how we might predict, or manage similar environmental accumulations.

water in the pool -- flows



a. In the box below, draw the line representing what happens to the amount of water in the pool given the flows shown below. The initial amount of water is indicated by the **X**.

amount of water in the pool



b. Explain why you drew the line this way:

Further, comparing qualitative student responses between the baseline and final exam shows a marked change in their ability to explain why they answered the way they did.

The table below shows a selection of student responses to the baseline qualitative question and the final exam question. It includes responses from those students who received a perfect score on the final exam (grade here shown as 100% correct) and indicates whether the respondent answered the initial PINS systems questions correctly (1) or not (0). The text descriptions illustrate a general increase from BASELINE to FINAL in ability to explain the structural reason why the system behaved the way it did. Even students who got both or one of the systems questions correct on the baseline could generally not explain why they chose their answer, but provided a sophisticated description on the final. Baseline explanations of why the stock changed were quite simplistic and often not related to the questions. The diversity in responses increased confidence that these descriptions were thoughtful, not rote responses.

Table 1 Responses for students who scored full credit on the systems question on the Final

Water in Pool Grade	Pond 3: Most in Pond CORRECT	Pond 4: Least in Pond CORRECT	BASELINE Description (revised PINS problem) “Describe how the total amount of the pollutant changes over 30 days. Why do you think it changed in this way?”	FINAL description of how level of Water in Pool changes. “Explain why you drew the line this way:”
100	1	1	it rises until day 13, when green is more than blue. then it constantly gets smaller	Initially, inflow is only slightly greater than outflow. This changes, however, and outflow is soon a lot more than inflow, causing the amount of water in the pool to rapidly decrease.
100	1	1	I don't know	At first the inflow was larger than the outflow meaning the water level would increase. That changed and soon the outflow was much larger than the inflow explaining the steep decrease in the last half.
100	1	1	at first it is steadily increasing until day thirteen where it levels out and then starts decreasing	At first, in flow was higher than outflow so levels increased. But as inflow decreased and outflow icreased amount of water would begin to level off and then decreases so long as outflow was greater than inflow.
100	1	0	i dont know	In the beginig because there is more in flow then out flow the was full over time out flow was way above in inflow which means the pool is loosing water fast
100	0	1	At first the amount added was greater than removed so there was a surplus. At about the half way point in the graph the amount removed was greater reducing the amount added	At first it increases because in flow is greater than outflow. Then inflow drops and stays below resulting in an overall loss at a consistent rate.
100	0	1	different days different things could be going on. like some days there might be busier days at the beach which gets polluted more.	The inflow was slightly higher than outflow so I drew a slight rise line. Once the line met I made my graph flat. Outflow was higher than inflow considerably so I drew a rapid decreasing line.
100	0	1	the amount of pollutants will decrease. They are removing more than they are adding	At the beginning there was more inflow than outflow so the amount of water initially increased. After time the amount of water being put into the pool was less than

				the amount taken out and the maount of water rapidly decreased.
100	0	1	Because from day 12 on, more pollutants were being removed than added and a greater rates than there were ever added over removed.	When the level of inflow was higher, the amount of water woiuld go up on a curved line. And, as the level of outflow increased while the inflow decreased, the curve switched directions to curving downwards, indicating a decrease in the total amount of water.
100	0	0	This could have been a month of a swarm of certain animals that eat or will get rid of the pollutant, and along with the same time of less pollutant being added to the pond.	I drew the line this way because at first the inflow is more than outflow, raising the amount of water. Then the outflow becomes significantly more than inflow, making water decrease.
100	0	0	Pollution in the water began to increase from Day 1 to Day 13, but the reverse began to happen in Day 14 on when the amount of pollutants removed became significantly higher than the amount of pollutants added. Also, I can't go back and change my answer from the previous questions, but looking back on it from this approach I now realize the day with the most pollutants is Day 13, and the day with the least pollutants is Day 30.	If the outflow is currently less than the inflow, the water level will increase. However, just after 120, the outflow increases dramatically while the inflow decreases dramatically, causing the water levels to drop.
100	0	0	I got confused.	First two hours, both inflow and outflow remain the same amount, but inflow is more than outflow so that the total amount of water increases gradually. Between 120 and 180 minutes, inflow goes down 4.5 from 5.5 while outflow goes up from 5.0 to 6.0. The amount of water drops down rapidly. After 180 minutes, both flows stay the same amount but the outflow surpasses 6.0 to 4.5 in the inflow. thus, the total amount of water decreases gradually.
100	0	0	There was more removed when there was less because it's easier to remove more when it is less.	The first 140 minutes or so shows the water was increasing because the inflow was greater than the outflow. After that the water rapidly decreases because the outflow is much larger than the in flow.
100	0	0	The total amount of pollutant decreases over the 30 days. It most likely changed in this manner due to a new system of removing the pollutants and a change in the source of the pollutants.	I drew the line this way because until 120 minutes, the inflow was greater than the outflow, so the amount of water in the pool was increasing. It was still increasing but at a slower rate until about 150 minutes, where the line peaked. From then until 180 minutes, the outflow exceeded the inflow, progressively more and more as the minutes went on, so my line had a downward curve. From that point to the end, my line had a straight, sharp downward slope because the outflow rate was consistently much higher that the inflow rate.
100	0	0	At first the amount of pollutants added is greater than the amout removed, but after the first 12 days the amount removed continues to rise and the amount added substantially decreases. The amount removed spikes at	At first, when the inflow is greater than the outflow, the amount of water in the pool increases. As soon as the outflow is greater than the inflow however, the amount of water will decrease and continue to go down.

			day 21 and then decreases while the pollutants added continues to decrease.	
100	0	0	The total amount of pollutants changes over thirty days because people burn fossil fuels very often and people ruin the ecosystem on a daily basis.	I drew the line like this because the inflow and outflow start of steady with the inflow greater than the outflow. This means the amount of water is increasing. Then the inflow and outflow meet and start to trace places. This means the water is decreasing.
100	0	0	no idea	At the beginning the inflow is greater than the outflow, so the amount of water keep raising, until it reach 120 mintues. The flow drop and out flow raise, which make the amount of water drop. After 180 (?) the outflow constantly greater than the inflow, so the amount of water keep dropping.
100	0	0	At first, the added pollutant was increasing, then it started decreasing. Even though it increased a few times, it still less than the peak (day 8). I think it changed because they became aware of the pollutants added and started working on it	From [0,120] the inflow is greater than the outflow so that pool is increasing. From [120,180] the inflow is decreasing as the outflow increases. This results in a concave down graph. The maximum point (peak) is where the inflow and outflow are equal and so there is no change. After the maximum, the graph decreases because outflow is greater than inflow. Even when the rates are not changing, the graph still decreases because outflow > inflow.
100	0	0	i dont know	As long as the inflow was greater you had a slight increase in the amount of water in the pool. Where the two lines meet, is the last point where water levels would stay even. After this the water level slowly decreases up to 180minutes and after this the rate continues to drop dramatically since you have a loss of 105 gallons a minute. .5gal/minute increase first 120 minutes, 1.5gal/minute loss from 180 minutes (?)
100	0	0	The amounts closely mirror each other.	At first the inflow amount is .5 gal/min more than outflow. This would make the total amount increase. Overtime the amounts change and outflow becomes 1.5 gal/min more than inflow. This could cause the total amount of water in the pool to decrease at a constant rate.
100	0	0	As the added went up the natural way it was removed was boosted and eventually overpowered the added.	I drew the graph like this because intifially, there was more water coming in than going out. So the levels rose. Then around 130 minutes this changed dramatically and far less was inflowing than out flowing so the water decreased dramatically and this remained the case as far as the graph tells past 360 minutes. So the water will continue to drop.
100	0	0	the people finally realized that what they were doing was wrong.	When the flow is greater than the outflow the pool will fill up. When the two lines approach each other the line will begin to even out. When the outflow is greater than the inflow the pool will derease in water.
100	0	0	Don't know	Up until 120 minutes the line moves up because the inflow of water is larger than the outflow. The amount of water slowly increases. After 120 minutes, the outflow exceeds the inflow by 1.5 gallons/min. This means that water will decrease and keep decreasing more and more water is being removed.

100	0	0	It decreases over the 30 days. I don't know why this happened though.	In the beginning, since the inflow was greater than the outflow, the total water was increasing. When the outflow became higher than the inflow, the total amount of water began to decrease.

Detail of results: baseline graphing and systems performance

Respondents demonstrated poor systems understanding on the baseline assessment, with an average **Systems Score** of 2.04 (out of 7) on the multiple choice questions (Table 2). This was not explained by poor graph skills, however. Students did surprisingly well on the graphing questions, given they were mostly not in STEM majors. The mean **Graphing Score** was 4.09 (out of 6) (Table 3). Figures 8 and 9 show two of the graphing questions, with results.

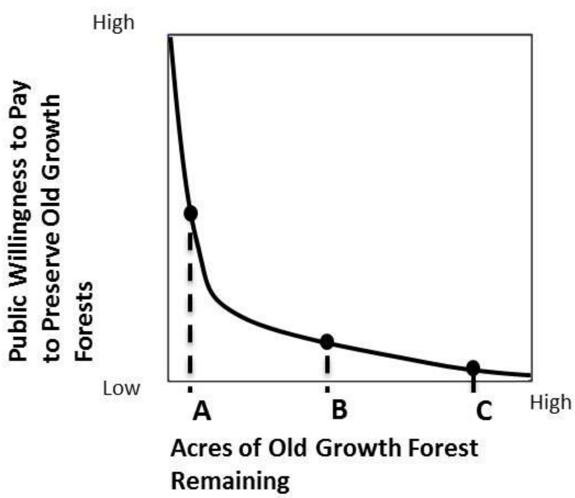


Figure 8 Graphing Question 1

GQ1: Based on the graph, at which point would the public be LEAST willing to help protect old-growth forests?

Row Labels	N	%
a.Point A	31	21.99%
b.Point B	7	4.96%
c.Point C	**85**	**60.28%**
d.All three points indicate equal willingness to pay to protect old-growth forests.	4	2.84%
e.I don't know	14	9.93%
Grand Total	**141**	**100.00%**

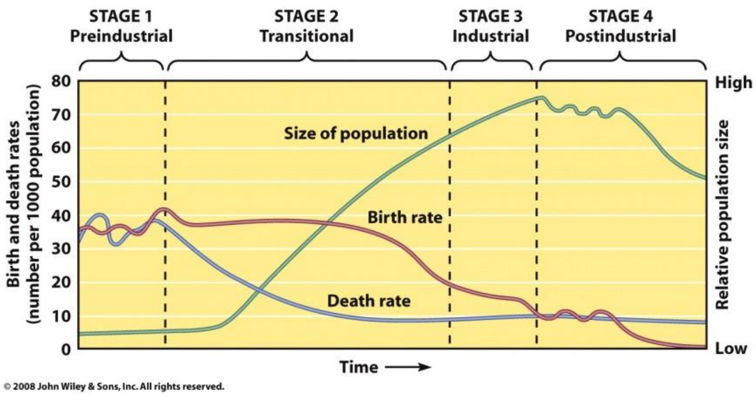


Figure 9 Graphing Question 3

GQ3: In which section of the graph is birth rate consistently below death rate?

Row Labels	N	%
a.STAGE 1	8	5.67%
b.STAGE 2	6	4.26%
c.STAGE 3	8	5.67%
d.1st half of STAGE 4	4	2.84%
e.2nd half of STAGE 4	**107**	**75.89%**
f.I don't know	6	4.26%
not answered	2	1.42%
Grand Total	**141**	**100.00%**

Table 2 Graphing Score

Graphing score (max=6)

Row Labels	N	%
0	5	3.55%
1	11	7.80%
2	14	9.93%
3	17	12.06%
4	26	18.44%
5	35	24.82%
6	33	23.40%
Total	**141**	**100.00%**

Mean 4.02
Std dev 1.73

Table 3 Systems Score

Systems score with PINP (max = 7)

Systems Questions Correct	N	%
0	25	17.73%
1	29	20.57%
2	34	24.11%
3	29	20.57%
4	19	13.48%
5	4	2.84%
7	1	0.71%
Total	**141**	**100.00%**

Mean 2.04
Std dev 1.45

Student performance on the modified PINS problem (Pollution in the Pond) mirrored the pattern seen by others, although the percentages were even lower than for STEM-oriented students at more rigorous schools. Table 4 shows the results from the multiple choice questions.

Table 4 Results from the Pollutant in the Pond (modified PINS) problem

	Most Added		Most Removed		Most in Pond		Least in Pond	
	N	%	N	%	N	%	N	%
a.1	0	0.00%	0	0.00%	1	0.66%	4	2.65%
b.5	0	0.00%	0	0.00%	0	0.00%	2	1.32%
c.8	**109**	**72.19%**	4	2.65%	89	58.94%	4	2.65%
d.13	1	0.66%	1	0.66%	**9**	**5.96%**	10	6.62%
e.17	3	1.99%	1	0.66%	1	0.66%	3	1.99%
f.21	10	6.62%	**107**	**70.86%**	13	8.61%	50	33.11%
g.23	2	1.32%	12	7.95%	4	2.65%	39	25.83%
h.25	2	1.32%	2	1.32%	3	1.99%	1	0.66%
i. 27	1	0.66%	3	1.99%	2	1.32%	0	0.00%
j. 30	6	3.97%	1	0.66%	3	1.99%	**11**	**7.28%**
k. Can't be determined	4	2.65%	6	3.97%	8	5.30%	7	4.64%
l. I don't know	12	7.95%	12	7.95%	17	11.26%	17	11.26%
not answered	1	0.66%	2	1.32%	1	0.66%	3	1.99%
Total	151	1	151	1	151	1	151	1

Discussion and Secondary Analysis

Since systems understanding increased, and the only systems instruction students received was through the simulation exercises, the simulations appear to be the reason for the increase. Looking more closely at the relationship between simulation use and systems understanding, however, raises more questions.

We expected to see a strong positive correlation between the **Run Count** (number of times they ran the simulation) and performance on the final synthesis question. However, Figure 10 shows a slight trend in the opposite direction.

There is a significant correlation between run counts in the simulations and the grade on that assignment as shown in Figures 11 and 12 and Table 5.

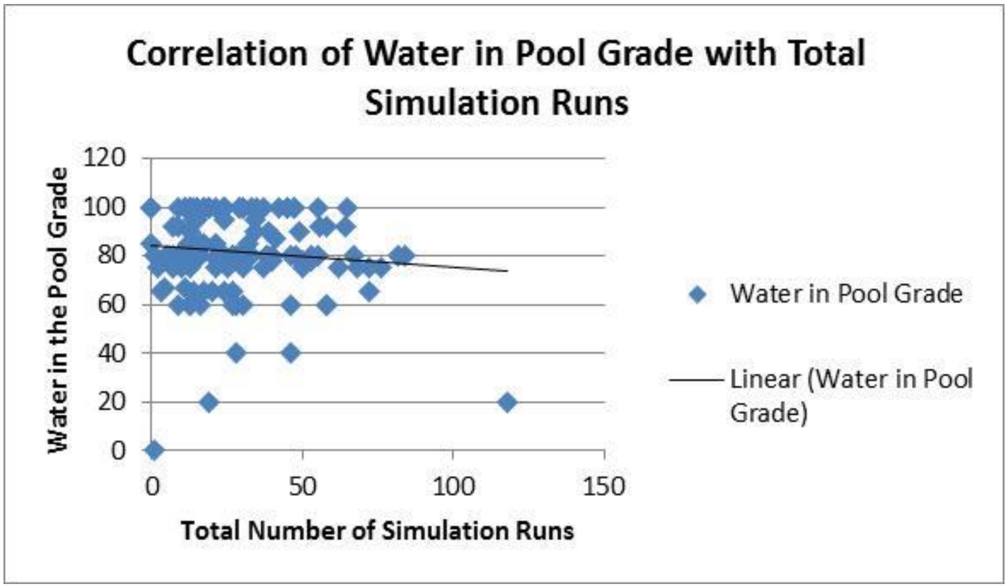


Figure 10

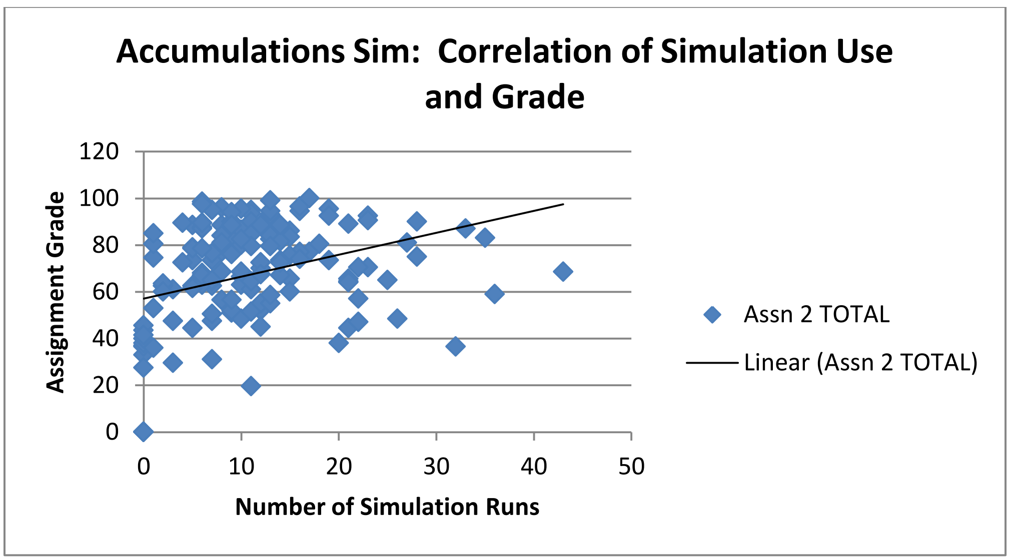


Figure 11

Table 5 Correlation analysis between Assignment 2 grade and number of simulation runs

Correlations

		Assn2Total	Assn 2 Run Count
Assn2Total	Pearson Correlation	1	.417**
	Sig. (2-tailed)		.000
	N	151	151
Assn 2 Run Count	Pearson Correlation	.417**	1
	Sig. (2-tailed)	.000	
	N	151	153

**. Correlation is significant at the 0.01 level (2-tailed).

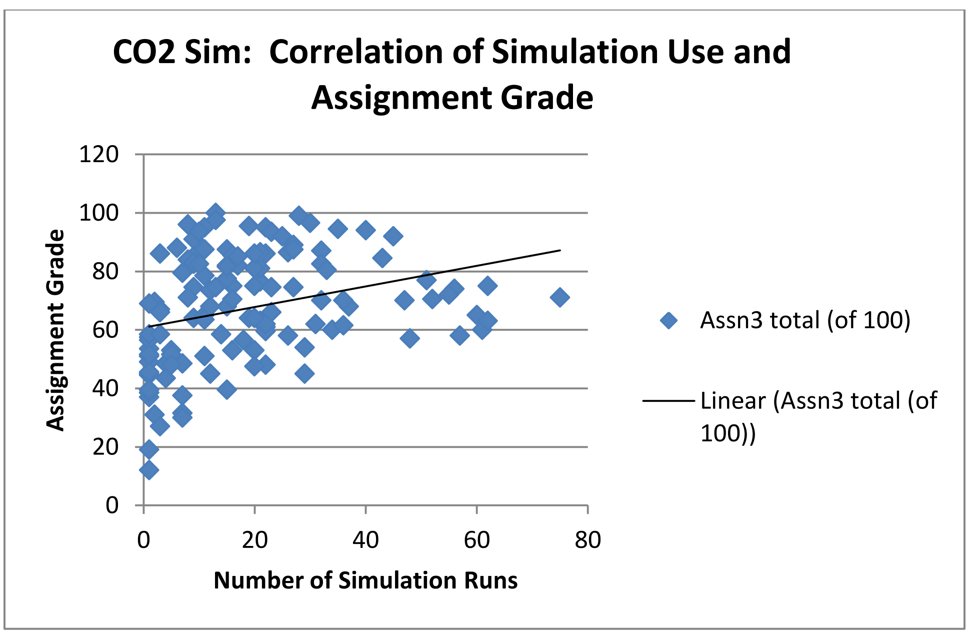


Figure 12 Assignment 3 Grade vs. Simulation Use

So what can we say about the effect of the simulations? The extent of the simulation use has an effect on the immediate grade, but total use does not seem to affect the final systems grade. One explanation may be that there is some threshold level of simulation use by which a user either will or will not “get it”. We can examine this further by stratifying the subjects further to see if we can better understand which users benefit from the simulation.

However, another, potentially more useful interpretation is that the relationship between simulation runs and learning is non-linear. We looked at the data by categories of runs: low (1-9 total runs), mid-range (10-19 runs), and high (20+ runs).

The average score for those running simulations ten or more times is significantly greater than students running fewer than ten simulations. The low-run group does not show a linear relationship between run count and the water in pool score. Students running the simulations more than once are not gradually scoring higher as would be expected. However, after about ten runs the relationship becomes closer to what we would expect, then as the number of runs increases further, the relationship is fuzzy again.

These results suggest that running the simulations only becomes effective after a certain number of runs, then the effect increases as expected until a saturation point is reached. However, the upper threshold does not lead to a steep drop in scores. Rather, upon hitting the saturation level, the average score dips then almost plateaus at a level still higher than the low-run average. The regression outputs below demonstrate that the middle-most run counts have the most sizable effect whereas the high run count group is not predicted to score higher than the low-run group. The high-run group is also not predicted to score differently from the mid-run group. The middle chunk of the run count distribution, on average, performs well enough on the systems score final question to be significantly greater than the low-run group. There is a slight curvilinear relationship wherein the downward curve is not as steep as the initial upward trend.

It appears there is either a saturation point, i.e., that the learning that takes place happens with the mid-range usage level, or that the students who are just not learning from the simulation tend to run it more times, possibly out of confusion.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.959	1.142		3.466	.001
	dummy 10-19=1; ref=1-9	2.105	.851	.306	2.473	.015
	dummy 20+=1; ref=1-9	1.360	.747	.224	1.819	.072
	Graphing index Score (max 6)	.181	.170	.104	1.063	.290
	systems score with PINP (max 7)	.644	.212	.317	3.044	.003
	num previous sci courses	-.458	.544	-.076	-.842	.402
	Why take the course?	.164	.564	.028	.291	.772
	Native born?	-1.208	.711	-.156	-1.699	.092

In summary, we expected to find a positive relationship between run count and final score but it turns out the run count variable has a skewed distribution. Since the relationship between run count and score is not monotonic we tried multiple categorical transformations of the run count variable. This categorization into low-, mid-, and high-range run counts shows expected qualitative differences between groups. The systems baseline score is a significant predictor of the final pool question score, but the regression models reveal that the effects of run counts hold regardless of graphing and systems base scores.

Finally, the effect of the accumulations principles information is unclear. We know that some people used the accumulations principles page and some did not. However, we did not track use of the principles pages for the first simulation assignment. Since the principles information is the same in both simulations, it may be that students used the information in the first simulation and not the second. We do know that students, in general, do not read much of the information on the screens. In the development phase of the simulation, we collected feedback about the simulation design including navigation, interface layout and text. We consistently found students read very little of the text. However, students may be more motivated to pay attention to the instructional text in some circumstances, for example, if they did poorly on the first simulation, or if they are having trouble answering a specific question. We plan to track not only how many times users visit the information page or run the simulation, but where they come from to do it. It may be that the reason they take these actions has a different effect on how they process the information or simulation feedback.

References

Cronin, M.A., C. Gonzalez, and J.D. Sterman. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes* 108: 116-130.

Forio Simulate [Computer Software package] Available from: http://forio.com/

Skaza, H.J. and K.A. Stave. (2010). Assessing the effect of systems simulations on systems understanding in undergraduate environmental science courses. *Proceedings of the 28th International Conference of the System Dynamics Society.* Seoul, Korea, July 25-29, 2010.

Stave, K. (2011). Using Simulations for Discovery Learning about Environmental Accumulations. *Proceedings of the 29th International Conference of the System Dynamics Society.* Washington, D.C., July 24-28, 2011. Available at: http://www.systemdynamics.org/conferences/2011/proceed/index.htm.

Stave, K., B. Jurand, and H. Skaza. (2011). Description and demonstration of a simulation learning environment for discovery learning about accumulations. *Proceedings of the 29th International Conference of the System Dynamics Society.* Washington, D.C., July 24-28, 2011. Available at: http://www.systemdynamics.org/conferences/2011/proceed/index.htm.