Pre-College Student Understanding of Accumulations:
An Experiment at a WPI Summer Workshop for Students

by Diana Fisher
with Chris DiCarlo, Alan Ticotsky, and Rob Quaden

Abstract

This paper will present the results of a pre- and post-test assessment given to 18 students ages 14 to 17 years who participated in a WPI Sustainability Workshop in the summer of 2014, where systems thinking and system dynamics modeling were the primary tools used to study environmental issues. The pre- and post-test were designed to assess the students' ability to determine simple dynamic behavior of phenomenon when the rate of change of the phenomenon was given in text, pictorial, or graphical form. The assessment determined that students have a reasonably robust intuitive ability to determine simple dynamic behavior from text and pictorial descriptions but not from graphical descriptions.

Introduction

Understanding the relationship between patterns of accumulation and rates of change is a fundamental concept not only in System Dynamics modeling but other areas of quantitative analysis, as well.

> "The idea of accumulation both grows out of and contributes to a coherent understanding of rate of change (Carlson, Smith, & Persson, 2003, in Thompson & Silverman, 2008). When something changes, something accumulates. When something accumulates, it accumulates at some rate. To understand rate of change well, then, means that one sees accumulation and its rate of change as two sides of a coin" (Thompson & Silverman, 2008, p. 11).

In 1999 Linda Booth-Sweeney and John Sterman conducted an experiment with Massachusetts Institute of Technology (MIT) graduate students to determine how well they understood the pattern of growth/decline of a quantity that accumulated when given the pattern of the rate at which that quantity was changing over time. They were surprised to find poor performance on their experimental scenarios by a group of well-educated, mathematically sophisticated MIT graduate students (Booth-Sweeney & Sterman, 2000). This study has come to be known as the 'Bathtub Experiment' and has raised concern about whether educated people can be taught to think systemically if they are not able to perform well on activities considered precursor basic skills necessary for improving systemic thinking. (Booth-Sweeney & Sterman, 2000)

Cronin, Gonzalez, & Sterman (2009) suggest that research should explore whether this poor performance is reinforced by the US educational system. For those K-12 teachers who have been using systems thinking and System Dynamics modeling with their students for many years there is a relatively strong belief[1] that young children are more systemic thinkers, and that, as

[1] Informal conversations between teacher leaders in the Waters Foundation, the Creative Learning Exchange, NSF CC-SUSTAIN project, to name a few.

they progress through the educational system in the US, that systemic thinking is not only **not** fostered, but that more linear thinking is strongly reinforced. If this belief is accurate, it would seem reasonable, then, that students at MIT, a university that attracts students who have excelled in the current US educational system, might be exhibiting a strong propensity for the thinking that Cronin, et al. identify as stock/flow failure.

Some of the K-12 teachers who are and have been introducing System Dynamics modeling into the precollege curriculum made attempts to show that pre-college students were able to perform reasonably well on two of the experiments Booth-Sweeney and Sterman conducted in their 1999 study, the bathtub experiment and the cash flow experiment. Those results were presented at the International System Dynamics Conference held in New York City in 2003 (Sterman & Stuntz, 2003). The 2003 paper indicated that the pre-college students did, in fact, perform better than the MIT students. Yet, due to the fact that the teachers who conducted the experiments and analyzed the results were not researchers, there was a question about the validity of the results. Consequently, the precollege student results received little, if any, attention.

There had been some informal discussion among SD society members that assessing accumulation pattern understanding using only graphical representations of the rate of change patterns could be confounding the results of Booth-Sweeney's and Sterman's Bathtub Experiment. It is well documented that reading graphs, an essential skill for understanding issues in the world, is not an intuitive skill for many students and should be an explicit skill taught in classes requiring graphical interpretation (Glazer, 2011). Problems such as pattern matching, viewing the graph as a picture or a map, or prior knowledge anticipation about the content displayed in a graph, among others, are especially troublesome (Beichner, 1994; Clement, 1985; Glazer, 2011). Some of these are issues the MIT students had with their bathtub and cash flow analyses, mentioned by Booth-Sweeney and Sterman. Cronin & Gonzalez (2007) suggest future research should investigate whether the issue of stock/flow failure might be due to the use of only graphical representations for presenting the problems experienced by the MIT students. In point of fact, their 2007 study indicated "…the visual representation [graphs] of the dynamic system is the critical source of difficulty for understanding the relationship between flows and the stock." The experiment described in this paper will add some data addressing this specific concern.

A multi-representational approach to assessing students' understanding of the relationship between accumulation patterns over time, when a description of the entity's rate of change was provided, was undertaken in the summer of 2014. An opportunity presented itself through a workshop for precollege students, Environmental and Sustainability Studies, offered through Worcester Polytechnic Institute (WPI) and taught by Chris DiCarlo, Alan Ticotsky, and Rob Quaden, all teachers quite knowledgeable about System Dynamics modeling. These three teachers designed a one-week, intensive set of activities and discussions for the students using a systems thinking and System Dynamics modeling approach. A description of the lessons developed is provided later in this paper.

[The author] was tasked with designing a pre-test and a post-test assessment to try to determine whether the students' level of understanding accumulation patterns over time, given their rate of change patterns, improved after the week long workshop. The pre-test and post-test were designed to be of equivalent difficulty. George Richardson suggested the assessment questions should contain items in different formats, text descriptions, pictorial representations, and graphical representations.

What follows is a description of the students who were involved in the week-long workshop at WPI, the sequence of activities and lessons they experienced, and the results of the pre- and post-assessments about accumulations that were administered.

A Description of the WPI Workshop Student Population

There were 18 students who participated in the workshop, ranging in grade from 8th (age ≈ 14 years) to 11th (age ≈ 17 years). There were 6-8th graders, 7-9th graders, 4-10th graders, and 1-11th grader.

Students responded as follows to the question about their comfort level with math skills (i.e., solving equations, drawing graphs, reading graphs, writing equations from word problems, etc.): not very comfortable-0, average comfort-1, pretty comfortable-5, very comfortable-12.

Students responded as follows to the question about their comfort level with science skills (i.e., drawing diagrams, doing experiments, etc.) not very comfortable-0, average comfort-4, pretty comfortable-3, very comfortable-11.

None of the students had previous exposure to or experience with systems thinking (causal loops, feedback thinking), although one student said he may have heard something about this in his earth science class in 7th grade. Another student said he thought he may have heard about this (causal loops and/or feedback) in robotics class. A third student said he may have heard about this in his engineering class at school. None indicated they could draw causal loops.

When asked why they (the students) took this workshop, 8 students said it sounded interesting/fun, 7 wanted to learn more about the earth or the environment or science, one had taken a WPI workshop the year before and liked it so signed up for another, one said it was a spur of the moment decision, and one said his parents didn't want him to watch TV all summer so signed him up for the workshop.

The WPI Workshop Student Lessons

The WPI workshop was advertised as an introductory sustainability class dealing with complex systems where students would be using hands-on activities, designing causal loops, manipulating simulations, and building small computer models to help them learn to 'bring their use of natural resources into balance.' The course lasted one week (in early July) and the lessons ran from 9 am until 4 pm each day. The topics covered each day are listed below.

Day 1: After a brief introduction the pre-assessment was administered. Lessons began with the Friendship Game (Quaden, R., Ticotsky, A., & Lyneis, D.,2008) physical activity where students walk into a large stock/flow diagram that is constructed on the floor using masking tape, adhering to certain rules for entering the stock. The activity demonstrated linear and exponential change over time, followed by building the corresponding STELLA models. Then the students played the Mammoth (population) Game (Quaden, et al., 2008) and built the population model.

Day 2: The second day began with an introduction of feedback and causal loop diagramming, followed by an introduction to designing a STELLA interface layer to ease model testing. Next there was an explanation of the design of a graphical function and a demonstration of the use of a graphical function to incorporate the idea of carrying capacity into the Mammoth population model. Students were then introduced to the Connection Game (Quaden, et al., 2008) and how the connection circle can be used to help identify feedback loops from a story description. Students read The Lorax story and used a connection circle to determine the causal loops in the story. Finally, students built a model to depict the population dynamics associated with Easter Island.

Day 3: The students played Fishbanks and discussed the issues surrounding regulating the fishing industry. Students designed a paper/pencil model for Fishbanks, then created policy recommendations to try to prevent the crash of the fishery. Students presented their recommendations using the actual Fishbanks stock/flow model to back up their recommendations.

Day 4: Day 4 started with the Tree Game (Quaden, et al., 2008) activity, where students simulate planting and harvesting trees in a forest, according to certain rules. The next activity had students pour sand into containers of different shapes, record and graph the number of scoops of sand needed over time to fill the container. Students posted the 'sand' graphs and other students tried to figure out which container each graph matched. Students then built the Tree Game model, which contained an aging chain. Finally, the concept of delays was introduced.

Day 5: Students started the final day playing the It's Cool (Quaden, et al., 2008) activity, where they took temperature readings of a hot liquid as it cooled over time. The activity was followed by creating a STELLA model to capture the cooling liquid dynamics. Next students were introduced to the C-Learn Global Climate Change Simulation, talked about pertinent statistics associated with certain countries and participated in a climate summit. Finally, the post-assessment was administered, and the teacher leaders wrapped up the discussion for the day and the week.

The Shape of Change book referenced in the previous paragraphs provides significant teacher support for those interested in conducting the activities described. The Lorax lesson is described in detail on The Creative Learning Exchange (www.clexchange.org) website.

The next section will describe the types of questions used on the pre- and post-assessments, and will describe the results of the assessments.

The Workshop Assessment Questions and Results

The pre-test assessment contained 5 multiple-choice questions. The questions specified flow information for a given scenario in one of three formats, text, pictorial, or graphical. Three questions were text descriptions, one was a pictorial description, and one contained a graphical description. All questions involved determining, from the flows presented, whether the pattern of accumulation over time a. increased, b. decreased, c. stayed the same, or was d. not able to be determined.

Question 1 involved a text description asking how a population would change over time given that deaths exceeded births over the entire time. Question 2 showed a large water tank with a very large inflow pipe and a very small outflow pipe, and asked how the level of water in the tank would be changing over time. Question 3 provided a text description indicating the number of people entering the mall was growing, but one did not know anything about the number of people leaving the mall. The student was to determine the pattern for the total number of people in the mall over time. Question 4 was graphical, showing a graph of a linearly decreasing inflow of deposits over time, a graph of constant outflow of money spent but the inflow graph was always above the outflow graph. Question 5 was a text description about harmful gasses sent into the atmosphere decreasing over time, but always greater than the removal of those harmful gasses in the atmosphere. The question asked about the pattern of the total amount of harmful gases in the atmosphere over time.

For the 18 students who took the pre-test assessments, the results are shown below.

Table 1: Results of pre-test assessment regarding understanding the pattern of accumulation over time, given a description of the inflow and outflow pattern over time.

Question #	Type of question	Num. Correct Responses, n = 18
1	Text: population	16 (89%)
2	Pictorial: water tank	15 (83%)
3	Text: people in mall	16 (89%)
4	Graph: bank account	1 (6%)
5	Text: gasses in atmosphere	12 (67%)

The post-assessment contained 5 multiple-choice questions, of which there were three text descriptions, one pictorial description, and one graphical description. There was also one short answer question. The five multiple-choice questions involved determining whether the pattern of accumulation increased, decreased, stayed the same, or was not able to be determined, based on information about the rate of inflow and the rate of outflow presented in the problem. An additional consideration taken into account, on the 5 multiple-choice questions for the post-test, was that they contained no scenarios about sustainability that might have been studied in the workshop lessons. This effort was made to try to prevent the possibility that an increase in correct responses could be due to learning more about sustainability, studied in the course.

The sixth question (short answer) asked the student to describe how to tell another student the method that could be used to answer the types of questions presented in the previous five scenarios. That is, the attempt was to determine whether the student had developed an overall strategy for answering these types of accumulation questions and whether they could put that strategy into words.

On the post-test assessment question 1 contained a (text) question about the air in a rubber raft, where the description indicated that a person was pumping more and more air into the raft, but did not realize there was a hole in the raft, so some air was leaking out. Question 2 concerned the total number of stars in a box. It involved a picture of a box of stars with a conveyor belt showing 3 stars flowing in and 3 stars flowing out. Question 3 was concerned about the total number of active smartphones. It involved a text description indicating that the number of people buying smartphones was decreasing over time but always higher than the number who returned smartphones to the store. Question 4 concerned a generic situation involving the amount of money in a box. The flow information was displayed graphically. The outflow graph was constant but always above a linearly decreasing inflow graph. Question 5 contained a text description with regard to the changing temperature inside a house. The outside temperature was cooler than the inside house temperature. The thermostat is heating the house less and less but a window in the house is partially open and warm air is leaving the house.

There were 16 students who took the post-test. Their results are shown below.

Table 2: Results of post-test assessment regarding understanding the pattern of accumulation over time, given a description of the inflow and outflow pattern over time.

Question #	Type of question	Num. Correct Responses, n = 16
1	Text: rubber raft	11 (69%)
2	Pictorial: stars in box	16 (100%)
3	Text: active smartphones	14 (88%)
4	Graph: stock of money	15 (94%)
5	Text: temperature in house	2 (13%)

Initially it was thought the results of the pre- and post-test showed that representing flow patterns graphically was exceptionally hard for students, possibly due to the fact that reading and interpreting graphs is a skill that needs quite a bit of practice, which these students would have had in the workshop. But upon further analysis it was determined that the post-test graphical question could have had more correct responses due to pattern-matching, or what Cronin, Gonzales, and Sterman (2009) refer to as the correlation heuristic, since the correct answer for the change in accumulation and the change in inflow graph both moved in the same direction (which was not the case in the pre-test). This question will be modified appropriately and the same pre-test and post-test will be administered to the second group of students who take the WPI workshop in the summer of 2015.

As for question 6, the results were mixed. Of the 16 students, 7 responses did not give enough detail to determine whether the students had determined a strategy. Another student had an answer that was on the right track but not detailed. That 8^{th} student indicated that one had 'to determine what is going in and out." None of these 8 responses is considered to have determined a correct strategy. Of the remaining 8 students, all indicated that an accumulation grew if the inflow was greater than the outflow, declined if the inflow was less than the outflow, and stayed the same if the inflow was equal to the outflow. Of these 8, two students also indicated that one should choose 'it is not possible to tell' if 'one or more of the amounts [inflow or outflow] is not given.'

Interpretation

An experiment was conducted with a group of pre-college students, ages 14 to 17 years, who had no previous experience with systems thinking or system dynamics modeling. The significant number of correct responses on the pre-test text and pictorial questions seem to indicate that determining the dynamic behavior of a phenomenon when the rate of change was presented in text descriptions and/or pictorial descriptions did not present problems for most of these students. The very poor success rate with the graphical question, however, suggests that the graphical type of representation adds significant cognitive difficulty to the problem analysis.

Results on the post-test seem to support the analysis about the general ease of interpreting the dynamic behavior of a phenomenon when the pertinent information about its rate of change is presented using either text or pictures. The poor response on multiple choice question 5, a question using a text description, could indicate the question was not as clearly presented, or that contained a topic that was less familiar for students. The correct answer to question 5 was that there was insufficient information given to determine the dynamic behavior of the house temperature. One could surmise that such a response (pattern not possible to determine from the given information) is more difficult for students to assess. But question 3 on the pre-test and question 1 on the post-test required a similar conclusion and students performed reasonably well on those questions.

As mentioned in the previous section, the strong successful showing on the graphical question, number 4, on the post-test cannot provide useful information regarding the value of the workshop instruction in helping students to determine the dynamic behavior of the money in the box. Since it was possible to use pattern matching to obtain the correct answer to this question, the question was removed from consideration in this analysis.

The pre- and post-tests were intended to be of equivalent difficulty. In the future a counter-balancing technique, where half the students are given the pre-test and half the post-test before the workshop, will be used. After the workshop the students will take the other version of the pre- or post-test. This will allow a reasonably effective way to mitigate the issue of including a question that is more difficult (included unintentionally, such as question 5 on the post-test) and a way to help neutralize a question that has an unintended 'clue' to the answer (such as question 4 on the post-test).

Conclusion

The questions that were used on the pre- and post-test for the WPI student sustainability workshop were only testing whether the dynamic behavior of the phenomenon described would increase, decrease, stay the same, or could not be determined. This is the most basic characteristic of dynamic behavior that one would hope students could understand. The assessment indicates that these somewhat young students appear to have a fairly intuitive understanding of this dynamic (if the problem is presented using text or a picture).

The sample size for this study is quite small. The socio-economic background of the students was not assessed. It could be that these students came from family situations that were not only more supportive, since the WPI workshop advertisement may have been more visible to parents who are well educated, but from families whose parents not only support educational summer camps but have the wherewithal to pay for such a camp. But it does give a glimpse at an assessment of dynamic behavior analysis using younger students.

It would be useful to conduct research, starting with dynamic behavior concepts as simple as the one used in this assessment, at different age and socioeconomic levels. It would be interesting to determine whether high school seniors (18 years of age), and then whether college undergraduates would perform as well as the young students who participated in this WPI workshop. It would be even more interesting to determine if adults, who are not students, could do as well on this simple assessment.

If we can determine whether students have and then lose, through the education process, the ability to intuit simple dynamic behavior, we would have an indication that our educational system is working at counter purposes to increasing adult ability to analyze dynamic system behavior.

The concern about whether adults can intuit the general dynamic behavior of a simple phenomenon when presented with information about the rate at which the phenomenon is changing, is an important one. The Bathtub Experiment has shed light on more that just the 'lack' of MIT graduates to analyze dynamic behavior. The use of a potentially confounding variable, the use of graphs for the dynamic analysis, is one issue to consider. Another, more important concern, is whether the MIT students, exceptional products of the current educational system, are actually excelling at skills that are not just irrelevant for, but may actually be interfering with, analyzing complex dynamic systems.

References

Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American journal of Physics*, *62*(8), 750-762.

Booth-Sweeney, L., & Sterman, J. D. (2000). Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review*, *16*(4), 249–286. Retrieved from http://doi.wiley.com/10.1002/sdr.198

Clement, J. (1985, July). Misconceptions in graphing. In *Proceedings of the Ninth International Conference for the Psychology of Mathematics Education* (Vol. 1, pp. 369-375). Utrecht,, The Netherlands: Utrecht University.

Cronin, M. A., & Gonzalez, C. (2007). Understanding the building blocks of dynamic systems. *System Dynamics Review*, *23*(1), 1-17.

Cronin, M. A., Gonzalez, C., & Sterman, J. D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes*, *108*(1), 116-130.

Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in Science Education*, *47*(2), 183-210.

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

Sterman, J., & Stuntz, L. (2003). Learning bathtub dynamics: A follow-up. *Proceedings of the 21st International System Dynamics Conference*, New York, New York.

Thompson, P. W., & Silverman, J. (2008). The concept of accumulation in calculus. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp. 43-52). Washington, DC: Mathematical Association of America. Retrieved from: http://pat-thompson.net/PDFversions/2008MAA Accum.pdf.