# Analysis of Student Performance on The Bathtub and Cash Flow Problems

**Diana M. Fisher**
Wilson High School
1151 SW Vermont St.
Portland, Oregon USA
(503) 916-5280 (phone)/(503) 916-2705 (fax)
dfisher@pps.k12.or.us or dmfisher@aracnet.com

Both the "bathtub" and "cash flow" problems were administered to four classes at Wilson High School in Portland, Oregon in the spring of 2003. Each problem contained the same explanation, diagram, flow graphs, and blank graph pad (for the solution graph) that were written by Linda Booth Sweeney and John Sterman in their study "Bathtub dynamics: initial results of a systems thinking inventory" published in the System Dynamics Review Volume 16 Number 4 Winter 2000.

There was no special preparation given to the students except a verbal explanation of the problem scenario, an explanation of the shape of the flow graphs (but no comment about a comparison of relative position of each graph to the other), and a request to sketch the graph of the accumulation for each task. It was mentioned that the initial value of the accumulation was to be 100. Students were given the tasks and asked to work independently on each task. It took them an average of 10 minutes to complete both tasks. Those who took longer were allowed the extra time, so time to complete the task was not a factor in the results. The students were told that they would not be graded on the tasks, nor would they receive extra credit. They were told it was for a study that the instructor was conducting.

Two of the classes to whom these tasks were administered were advanced algebra classes (students ages 16 to 17). The other two classes were advanced placement (first year) calculus classes (students ages 17 to 18). Both of the advanced algebra classes and one of the AP calculus classes were taught by the author, where system dynamics modeling is used a few times throughout the year. The second calculus class was taught by another instructor. An attempt was made to get a third instructor to administer the tasks to his advanced algebra class, but the administration of the tasks was given to a substitute teacher who did not follow the administration guidelines, so those papers had to be discarded.

The students were also asked to indicate which category below best described their previous experience with STELLA modeling (the synonym the students use to describe system dynamics modeling at Wilson High School). The categories were:

1. I have had no previous experience with STELLA modeling.

2.  I have had one class where STELLA modeling was used a few (2 or 3) times.
3.  I have had a class were STELLA modeling was used more than 5 time during the year.
4.  I have taken a system dynamics modeling class using STELLA.
Comments: ________________________________________________________________

The comments section was used to have the students who selected category 2 or 3 list the classes in which they had used STELLA.  If a student had used the STELLA software a few time, but had more than one class were it was used a few times in each class, they were instructed to circle category 3.

**Numerical Summary:**

The same rubric was used to grade the student papers as was listed in the Booth Sweeny – Sterman article.  An attempt was made to interpret each item in the rubric in a manner similar to that listed in the article.  There could be error introduced in the comparison of the high school results and the MIT results due to potential differences in interpretation of specific rubric descriptions.  In a few places a score of 0.5 or 0.75 was used if the student graph indicated an understanding of at least half or more of the concept described but was not fully correct.  The high school results were compiled by one person so there should be minimal error introduced in comparing the results of one high school class to another high school class, in this study.  The results are summarized in the table below.  (Note: AA tot represents the total results for both advanced algebra classes.  c1, c2, c3, c4 represent the four categories of previous modeling experience as described in the previous section.  Cal A tot represents the total results of a first year calculus class taught by the same instructor as the two advanced algebra classes.  Cal B tot represents the total results of a first year calculus class taught by an instructor who does not use SD modeling in the instruction process.)

**Analysis of Bathtub Problem**

| Class | AA tot | AA c1 | AA c2 | AA c3 | AA c4 | Cal A tot | C A c1 | C A c2 | C A c3 | C A c4 | Cal B tot | C B c1 | C B c2 | C B c3 | C B c4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of students | 45 | 2 | 42 | 1 | 0 | 30 | 0 | 19 | 9 | 2 | 16 | 10 | 4 | 0 | 2 |
| Criterion | | | | | | | | | | | | | | | |
| 1.  When the inflow exceeds the outflow, the stock is rising | .71 | 0 | .74 | 1.00 | | .93 | | .89 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | | 1.00 |
| 2.  When the outflow exceeds the inflow, the stock is falling. | .73 | 0 | .76 | 1.00 | | .93 | | .89 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | | 1.00 |
| 3.  The peaks and troughs of the stock occur when the net flow crosses zero. (ie t=4,8,12,16) | .69 | 0 | .71 | 1.00 | | .93 | | .89 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | | 1.00 |
| 4.  The stock should not show any discontinuous jumps (it is | .91 | 0 | .95 | 1.00 | | 1.00 | | 1.00 | 1.00 | 1.00 | .97 | .95 | 1.00 | | 1.00 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| continuous) | | | | | | | | | | | | | | | |
| 5. During each segment the net flow is constant so the stock must be rising(falling) linearly. | .61 | 0 | .63 | 1.00 | | .97 | | .95 | 1.00 | 1.00 | .94 | .90 | 1.00 | | 1.00 |
| 6. The slope of the stock during each segment is the net rate (ie ± 25 units/time period) | .57 | 0 | .58 | 1.00 | | .83 | | .89 | .67 | 1.00 | .69 | .65 | .63 | | 1.00 |
| 7. The quantity added to (removed from) the stock during each segment is the area enclosed by the net rate (ie 25 units/time period*4 time periods = 100 units, so the stock peaks at 200 units and falls to a minimum of 100 units) | .59 | 0 | .61 | 1.00 | | .83 | | .89 | .67 | 1.00 | .69 | .65 | .63 | | 1.00 |

The vast majority of advanced algebra students created an M-shaped graph, correctly. Five students did not create any graph. Three students who entered the class half-way through the year did not draw M-shaped graphs. Their graphs were a mimic of the inflow graph. The majority of mistakes made in the rests of the graphs were that the M was drawn with a slightly curved M shape than an M drawn with straight lines and sharp peaks and valleys. Other mistakes involved drawing a more sinusoidal curve than even M-shaped. Finally, there were errors in drawing the correct height between the peaks and valleys.

Almost all the calculus students in group A drew the correct M-shaped graph. Three students drew and M-shaped graph that was not tall enough. One student did not draw a graph. One student drew two M-shaped graphs instead of one, both of which were too short. Finally one student drew a graph that just repeated the inflow graph.

Most of the mistakes made in the calculus group B papers involved M-shaped graphs that were too short.

**Analysis of Cash Flow Problem**

| Class | AA tot | AA c1 | AA c2 | AA c3 | AA c4 | Cal A tot | C A c1 | C A c2 | C A c3 | C A c4 | Cal B tot | C B c1 | C B c2 | C B c3 | C B c4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of students | 45 | 2 | 42 | 1 | 0 | 30 | 0 | 19 | 9 | 2 | 16 | 10 | 4 | 0 | 2 |
| Criterion | | | | | | | | | | | | | | | |
| 1. When the inflow exceeds the outflow, the stock is rising | .32 | 0 | .34 | 0 | | .70 | | .63 | .78 | 1.00 | .56 | .70 | 0 | | 1.00 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2. When the outflow exceeds the inflow, the stock is falling. | .33 | 0 | .35 | 0 | | .69 | | .63 | .78 | .88 | .56 | .70 | 0 | | 1.00 |
| 3. The peaks and troughs of the stock occur when the net flow crosses zero. (ie t=2,6,10,14) | .29 | 0 | .31 | 0 | | .68 | | .62 | .78 | .75 | .52 | .65 | 0 | | .88 |
| 4. The stock should not show any discontinuous jumps (it is continuous) | .84 | .50 | .86 | 1.00 | | .90 | | .84 | 1.00 | 1.00 | .94 | 1.00 | .75 | | 1.00 |
| 5. The slope of the stock at any time is the net rate. | .36 | 0 | .38 | 0 | | .59 | | .51 | .67 | 1.00 | .38 | .50 | 0 | | .50 |
| 6. The slope of the stock when the net rate is at its maximum is 50 units/period (t = 0, 8, 16) | .34 | | .37 | 0 | | .63 | | .51 | .78 | 1.00 | .56 | .70 | 0 | | 1.00 |
| 7. The slope of the stock when the net rate is at its minimum is -50 units/period (t = 4, 12) | .34 | 0 | .37 | 0 | | .63 | | .53 | .78 | .88 | .56 | .70 | 0 | | 1.00 |
| 8. The quantity added to (removed from) the stock during each segment of 2 periods is the area enclosed by the net rate (ie a triangle with area ± (1/2)*50 units/perios*2 periods = ±50 units.) The stock therefore peaks at 150 units and reaches a minimum of 50 units. | .28 | 0 | .30 | 0 | | .50 | | .42 | .61 | .75 | .41 | .50 | 0 | | .75 |

The majority of advanced algebra students who did not draw correct graphs drew graphs that either reflected the input graph or were slightly truncated sinusoidal graphs. There were half a dozen graphs that were not strictly periodic, where the peaks on the second half of the graph were higher or lower than the peaks on the first half of the graph. There were six students who chose not to draw any graph at all.

For the calculus students in group A the mistakes centered around graphs that were straight lines with sharp peaks and valleys rather than sinusoidal. There were two papers that had graphs that were shaped as two U’s next to each other. Three students chose not to draw graphs. One student drew two M-shaped graphs next to each other in the upper half of the graphing rectangle.

For the calculus students in group B about a third of the papers showed graphs that were composed of M-shaped lines with sharp points at the peaks and valleys. Some of these had the correct height, others did not. Some of the sinusoidal graphs were not of the correct height. One student did not draw a graph. Another student drew a graph that was a series of horizontal segments, connected to each other, that rose and fell in a shape that could roughly mimic the inflow rise and fall without regard to its relative position to the outflow graph.

**Comparison of High School Results with MIT Student Results**

This table compares the total advanced algebra students, the total first year calculus students in group A and the total first year calculus students in group B with the MIT students. An additional column is introduced in this chart to include those students who had taken a system dynamics modeling class (listed as c4 students above) and including another small group of four independent study system modeling students whom the author taught this year. (Note: One student is in both a c4 category and a calculus A class, but was counted only once.)

**Bathtub Problem**

| Class | AA tot | Cal A tot | Cal B tot | SD tot | | MIT |
|---|---|---|---|---|---|---|
| Number of students | 45 | 30 | 16 | 7 | | 182 |
| Criterion | | | | | | |
| 1. When the inflow exceeds the outflow, the stock is rising | .71 | .93 | 1.00 | 1.00 | | .80 |
| 2. When the outflow exceeds the inflow, the stock is falling. | .73 | .93 | 1.00 | 1.00 | | .80 |
| 3. The peaks and troughs of the stock occur when the net flow crosses zero. (ie t=4,8,12,16) | .69 | .93 | 1.00 | 1.00 | | .86 |
| 4. The stock should not show any discontinuous jumps (it is continuous) | .91 | 1.00 | .97 | 1.00 | | .89 |
| 5. During each segment the net flow is constant so the stock must be rising(falling) linearly. | .61 | .97 | .94 | 1.00 | | .78 |
| 6. The slope of the stock during each segment is the net rate (ie ± 25 units/time period) | .57 | .83 | .69 | .93 | | .66 |
| 7. The quantity added to (removed from) the stock during each segment is the area enclosed by the net rate (ie 25 units/time period*4 time periods = 100 units, so the stock peaks at 200 units and falls to a minimum of 100 units) | .59 | .83 | .69 | .93 | | .63 |
| Mean for all items | .69 | .92 | .90 | .98 | | .77 |

**Cash Flow Problem**

| Class | AA tot | Cal A tot | Cal B tot | SD tot | | MIT |
|---|---|---|---|---|---|---|
| Number of students | 45 | 30 | 16 | 7 | | 150 |
| Criterion | | | | | | |
| 1. When the inflow exceeds the outflow, the stock is rising | .32 | .70 | .56 | 1.00 | | .47 |
| 2. When the outflow exceeds the inflow, the stock is falling. | .33 | .69 | .56 | .96 | | .44 |
| 3. The peaks and troughs of the stock occur when the net flow crosses zero. (ie t=2,6,10,14) | .29 | .68 | .52 | .89 | | .40 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4. The stock should not show any discontinuous jumps (it is continuous) | .84 | .90 | .94 | 1.00 | | .99 |
| 5. The slope of the stock at any time is the net rate. | .36 | .59 | .38 | .86 | | .28 |
| 6. The slope of the stock when the net rate is at its maximum is 50 units/period (t = 0, 8, 16) | .34 | .63 | .56 | 1.00 | | .47 |
| 7. The slope of the stock when the net rate is at its minimum is -50 units/period (t = 4, 12) | .34 | .63 | .56 | .96 | | .45 |
| 8. The quantity added to (removed from) the stock during each segment of 2 periods is the area enclosed by the net rate (ie a triangle with area ± (1/2)*50 units/perios*2 periods = ±50 units.) The stock therefore peaks at 150 units and reaches a minimum of 50 units. | .28 | .50 | .41 | .57 | | .37 |
| Mean for all items | .39 | .67 | .56 | .91 | | .48 |

**General Results:**

One would not expect high school students to be more successful than MIT students at analyzing change behavior for each task. Since the results seem to indicate that most of the high school students, the calculus ones especially, performed comparatively well on these tasks it suggests that perhaps more than one factor might be involved in generating these results. Some of the factors to consider are presented in the conclusion section of this paper.

The modeling group performed extremely well, but with only 7 students in that category one must be cautious about drawing too many conclusions. Certainly it can be said that SD modeling appears to aid the analysis of behavior over time graphs, since the course has as one of its instructional strategies the analysis of flow and accumulation graphs. Even the more difficult task (cash flow) was not too challenging for the SD students.

The calculus students from both classes performed well on the bathtub task. It has become part of the calculus curriculum to analyze flow and accumulation graphs in relation to one another, so this task was well within the experience of most of the calculus students, even though they had not seen a separate inflow/outflow scenario before. The cash flow task was more difficult for the calculus students, although many created graphs that were somewhat sinusoidal. There were ten students who specifically drew the accumulation graph as connected parabolas, as they should be drawn. Too many students did not pay close enough attention to the location of the peaks and troughs of the stock. This should have been easy for them. The study of when an accumulation increases and decreases is a fundamental concept in calculus.

The advanced algebra students had the most difficulty, as might be expected. However, they performed well on bathtub task. In class there were motion detector activities throughout the year that required them to study the relationship between distance and velocity curves for linear and quadratic functions. This may have helped their performance on this task. The cash flow problem was another situation altogether. Six students didn't even draw a graph. Still there were a surprising number of graphs (22) that were drawn that were sinusoidal in nature, even though some of them were not the correct height nor had the correct locations for the peaks and valleys.

**Implications/Conclusions:**

Before discussing the overall results it is important for the reader to understand the change in mathematics instruction that has been undertaken in the United States over the past five years. Due to changes initially adopted in calculus classes, generated by the calculus reform movement of the past 15 years, mathematics instruction now is supposed to provide multiple representations of function behavior for analysis. The success of the calculus reforms prompted the inclusion of many of the multiple representation instructional strategies in all high school algebra and pre-calculus classes. So students today are much more accustomed to analyzing a variety of graphs.

It appears that the new methods of instruction in mathematics in the United States provide a compatible environment for inclusion of tasks such as those represented by the bathtub and cash flow problems. Also, it is evident that instructional strategies used in the lower grades in mathematics, especially in middle school (with 12 and 13 year old students), is supporting the newer (and broader) approach to function analysis, including graphical analysis, at the high school level (14 to 18 year old students). Although the students, especially the advanced algebra students, had more difficulty with the cash flow problem, their work on the bathtub problem seems to indicate that they don't fall too short of MIT students in their performance. While it would be nice to think that high school students are getting smarter, it is more realistic to look at the learning environment and how it might be different today from the high school environment even three or four years ago for the MIT students. It may be that many MIT students studied mathematics in a more traditional instructional environment. This could be due to the fact that some of the MIT participants were older students, hence a significant delay time is present in the curriculum to which they were exposed compared to the curriculum for the current high school student. Additionally, many instructors who teach upper level mathematics in high school are the more senior staff who may or may not have been willing to change a method of instruction that had been successful for many years. So the MIT students may have had a more traditional calculus class, focusing on proofs rather than more conceptual interpretations of representing change behavior. Also, it is necessary to determine whether the interpretation of the high school student papers was done in a manner similar enough to the MIT papers that the numbers do represent a true comparison of the results. This alone could introduce significant error, negating the results.

That said, it is heartening to think that high school students might be gaining skill in the areas that we hope to see develop in future citizens. The results above, as much as can be concluded based on seven students, do indicate that students with significant exposure to SD modeling are adept at analyzing problems similar to the two represented in the tasks given. The bathtub and cash flow tasks are wonderful examples of problems that students should be encouraged to analyze as part of a normal mathematics course of study. It is hoped that more problems of this type can be designed and included in math curriculum. The change needs to occur in the teacher's minds more than the student's minds. The students are capable of much more sophisticated thinking than we have given them opportunity to exercise in the past. But the process must start in elementary school, and be developed throughout each succeeding year of study.