Designing Assessments to Judge Attainment of High Level Learning Goals: One Step in Propagating System Dynamics in Education (Stage 1: Background)

Diana M. Fisher, Edward Gallaher, Martin Schaffernicht, Juliette Rooney-Varga

> Every other Jewish mother in Brooklyn would ask her child after school, 'So? Did you learn anything today?' But not my mother, 'Izzy,' she would say, 'did you ask a good question today?' That difference – asking good questions – made me become a scientist. Isador Issac Rabi (Nobel Laureate Scientist)

Abstract
The goal of this study is to design sample assessments that are useful to researchers as they conduct studies that illuminate the difference in student learning that System Dynamics (SD) infused lessons provide.

Stage 1 of this study prepared the background for our assessment efforts. In this paper we will identify some of the higher level educational goals used in several curricular disciplines in US secondary schools. We will identify core concepts embedded in an SD approach to analyzing a systems problem. We will describe some learning outcomes that have been identified by six secondary school teachers that differentiate their students that have used SD modeling techniques from students who have not used SD modeling in their classes over more than 20 years. Finally, we will set the stage for part 2 of this assessment study that will identify a generic assessment template designed to capture this differentiated SD learning. In stage 2 we will provide some example assessment questions applied to two specific disciplines, and report on the results of tests that were performed in several classrooms.

Introduction

The US Educational System
What do we expect our children to learn by attending school? Do we want to prepare them to get a job that will help them support themselves and, eventually, their family? Do we want them to learn from people who have lived before them so they do not make the same mistakes made by past generations? Do we want them to develop skill at those tasks that past generations found useful (like solving algebraic equations, learning to write coherent sentences, playing a musical instrument)? Do we want them to learn to work cooperatively with people who are different from themselves? Do we want them to make a contribution to the society in which they will live and work?

The answer to these questions is, for most people, probably yes. To operationalize these 'goals' educational leaders develop standards to guide pre-college instruction. Some of the standards for the traditional disciplines that are represented in a school setting deal with detail at a skill-based level. For example, in algebra, a standard might indicate that students should be able to "solve quadratic equations in one variable" (Common Core Standards Math Practice) and go on to explain the preferred method for obtaining such a solution. Or, in physics, a student should be able to "plan and conduct an investigation to provide evidence that an electric current can

produce a magnetic field and that a changing magnetic field can produce an electric current" (NGSS Lead States, 2013).

Statewide exams often base questions on these low-level standards, as they are easier to assess using computer generated questions. Sample questions in algebra (national Common Core State Standards-Mathematics) and physics (state of Oregon) include:

1. Simplify $(x^3+3x^2-1) + (-2x^2-x+4)$. (There are 4 multiple choice responses.)(Common Core Standards Math Practice)
2. A surfer paddles out from shore in search of the perfect wave. The surfer has a weight of 500 N and the surfboard weighs 100 N. When the surfer is on a surfboard floating on calm water, what is the buoyant force pushing up on the board? (There are 4 multiple choice responses.) (Oregon Department of Education, 2011)

Schools in the US are rated based on student performance on statewide tests, and such ratings put pressure on school administrators to try to make sure their students perform respectably on these tests. Pressure is then brought to bear on teachers to shape their instructional content to make their students proficient test takers (Kaestle, 2013).

Organizations that develop national standards for school disciplines usually preface their standards guidelines by including a more broad based set of goals to try to guide instruction. The individual standards are aligned with the broader goals for each discipline.

Some examples of the goal categories/statements for several disciplines (middle and secondary school level in the U.S.) that might best relate to System Dynamics activities include:

1. **Math (M):** Mathematical practices include (National Governors Association Center for Best Practice, 2010):
   1.1. Make sense of the problem,
   1.2. Reason quantitatively,
   1.3. Construct viable arguments and critique the reasoning of others,
   1.4. Use appropriate tools,
   1.5. Model mathematically,
   1.6. Look for and make use of structure
2. **Science (S):**
   2.1. Cross-Cutting Concepts (NGSS Lead States, 2013) include the categories of:
      2.1.1. Cause and effect,
      2.1.2. Systems and systems modeling,
      2.1.3. Structure and function,
      2.1.4. Stability and change.
   2.2. Scientific and Engineering Practices (NGSS Lead States, 2013) include:
      2.2.1. Asking questions and defining problems,
      2.2.2. Analyzing and interpreting data,
      2.2.3. Developing and using models,
      2.2.4. Constructing Explanations and Designing Solutions,
      2.2.5. Engaging in Argument from Evidence,
      2.2.6. Using mathematical and computational thinking, and

2.2.7. Obtaining, evaluating, and communicating information.

3. **Social Science (SS):** (National Council for the Social Studies, 2013)
   3.1. Culture (cultures are dynamic and change over time);
   3.2. Time, continuity and change (analyzing the past helps students analyze causes and consequences of events);
   3.3. Geography (migration, economic interdependence, cooperation and conflict, physical systems affect human systems, resource consumption);
   3.4. History (spread of political systems, revolution, peace).
4. **Health (H)** (Joint Committee on National Health Education Standards, 2007)**:**
   4.1. Health promotion and disease prevention,
   4.2. Reducing health risks,
   4.3. Health services,
   4.4. Influence of health on life and work,
   4.5. Influence of health decisions on life and work,
   4.6. Communicating health information.
5. **Economics €** (Siegried, J., et al, 2010)**:**
   5.1. Financial literacy;
   5.2. Students should be able to fully and effectively participate in a complex global economy
6. **Education in General (EG)** (National Research Council, 2012): To achieve their full potential as adults, young people need to develop a range of skills and knowledge that facilitate mastery and application of English, mathematics, and other school subjects. At the same time, business and political leaders are increasingly asking schools to develop "21st century skills"
   6.1. Problem solving,
   6.2. Critical thinking,
   6.3. Communication,
   6.4. Collaboration, and
   6.5. Self-management.

System Dynamics model-building activities tend to more closely align with higher-level goals in the disciplines described above. Many statewide and national assessments tend to address lower level standards such that current assessment instruments are unable to evaluate the core knowledge and skills acquired through system dynamics modeling activities. Moreover, System Dynamics modeling brings another entire dimension to the problems students can address, the ability to analyze complex systems.

A Shift in the Types of Problems to Solve

Complex systems problems have been extant for millennia. Researchers at New England Complex Systems Institute claim that since the 1990s there has been an "explosion in complex systems science" (Bar-Yam, 2012). But only comparatively recently have the techniques and tools become accessible in sufficient secondary school classrooms that studying such problems could become a mainstream occurrence. We will not present arguments in this paper to justify the need to bring the study of complex systems to the secondary school

environment. We feel almost all educators are aware of the complex problems that face local, regional, national, and global communities and understand that addressing those problems is essential for the peace and prosperity, nay even the survival, of the human population.

Before educational decision makers are willing to make changes in traditional curricula it is necessary to provide compelling evidence that a new learning approach is worth the enormous effort needed to infuse it into the school curriculum. Therein lies the problem we try to address in this paper.

System dynamics, one of the effective tools used to address complex systemic problems, supports a holistic, dynamic, cross-disciplinary, hypothesis testing approach that uses a visual mathematical representation that is out of the mainstream experience of most secondary school educators. Not only is the representation unfamiliar, the approach (focusing on non-linear dynamics and feedback control, fostering trans-disciplinary analysis, incorporating delays in process structures) is also foreign to their academic preparation.

> Complex systems analysis changes what the equations are written about, shifting them from properties of things to properties of dependencies or relationships and collective behavior. Calculus and statistics do not have the tools to describe what we want to study. … It's not just about the math. We have to develop an entirely new set of concepts that enable us to think about the behavior of the system. Different languages. Really, most importantly, it changes the nature of the questions that we ask about the behavior of the system. It changes how we think. (Bar-Yam 2015)

Few, we feel, would question the need to have adults who are more capable of analyzing complex system behavior. Learning activities that introduce and develop student skill in working with complex systems have been developed and used with some pre-college students for over two decades (Fisher, 2011; Grotzer & Mittlefehdlt, 2012). But allocating the time in the current tightly packed curriculum will not have an easy solution.

The Goal of this Study

As members of the educational research community we need to design assessments that will demonstrate the level/degree to which students have achieved the desired goals identified in the previous section in mathematics, science, health, economics, social science, or education in general, while also allowing students to display their budding understanding of some of the behavior of complex systems.

The specific goal of this study is to design sample assessments that can guide researchers as they conduct control-experimental group studies that illuminate the difference in student learning that SD infused lessons provide. To achieve this goal the assessments must be designed to capture the deeper learning we believe is happening when SD is used. We do not believe these assessments exist but their development can be guided by other

assessment research that is occurring within the NGSS and "Deeper Learning" educational communities.

The secondary, but probably more overarching goal of this study is to (eventually) use assessment results (which we believe will show the value of learning using SD) to help promote the propagation of SD to broader K-16 educational communities.

Background Information

In preparation for designing new assessments, it is useful to take two slight detours. One detour is needed to make sure we can agree on those concepts that we consider central to learning when using System Dynamics modeling activities. The authors of this paper devised the following list of important concepts we felt were central to an understanding of System Dynamics. It is, no doubt, incomplete, but will suffice for this initial stage of our assessment study.

What are the core concepts imbedded in the SD analytical approach?

1. Assumptions about the nature of the world (systems) (Forrester, 1994)
   1.1. The world is *causal* (not correlational)
   1.2. Observable *behaviors* are *caused by* causal *structure*
   1.3. Endogenous view: systemic *behavior* is produced *endogenously*
2. Regarding behavior, the following patterns between stocks and flows always exist:
   2.1. *Bathtub dynamics*: for a stock (accumulator variable) to increase the total inflow values must be larger than the total outflow values.
   2.2. *Linear behavior* for the stock is produced if all its flows are constant in value,
   2.3. *Exponential behavior* for the stock is produced if its flow has values proportional to the current value of the stock,
   2.4. *Parabolic behavior* for the stock is produced if its flow changes in a linear (non-constant) pattern,
   2.5. *Logistic* (S-shaped behavior) for a stock requires at least one reinforcing feedback and one balancing feedback with its flow(s),
   2.6. To produce *oscillation*, one needs two stocks connected by a balancing feedback loop with a delay in it.
3. Regarding structure
   3.1. Cause and effect are not closely related in time or space.
   3.2. Feedback loops are pervasive and control everything that changes through time (Forrester, 2009).
   3.3. Students should be able to:
      3.3.1. Assign link polarity to designate the cause-effect relationship between two components in the system/model. Determining each link polarity in a feedback loop is valuable in determining loop polarity for feedback loops containing 3 or more components.
      3.3.2. Recognize/explain that transfer of (feedback) loop dominance is a demonstration of nonlinear (complex) system behavior.
   3.4. Any feedback loop contains at least one stock

3.5. Stock variables can only change due to flow variables
3.6. Students should be able to differentiate the components of the system correctly into stocks and flows as a necessary first step
3.7. Nonlinear relationships are fundamental in the dynamics of systems of all types. (Sterman, 2000)
3.8. Students should understand the need for dimensionless multipliers within a models and understand that dimensionless multipliers (sometimes called table functions, graphically defined functions) often capture the nonlinear relationships in SD models.

4. Consequences for a systemic world and our possibilities to influence systems
   4.1. Static equilibrium is the exception.
   4.2. Students should appreciate that "many industries are characterized by an unstable, nonlinear, self-limiting behavior" (Forrester, 1987).
   4.3. Low-leverage policies are often ineffective.
   4.4. High-leverage policies are often wrongly applied.
   4.5. We cause our own problems.
   4.6. A fundamental conflict exists between short-term and long-term goals.
   4.7. Consequently, students should be able to surface examples of systems/problems that demonstrate one or more of these characteristics.
5. Aspects of human understanding and modeling
   5.1. The process of building SD models is more important than the models arising from it. (Forrester, 1985)
   5.2. Students should understand that:
      5.2.1. All decisions are made on the basis of models. Most models are in our heads. (Forrester, 2009)
      5.2.2. Mental models control nearly all social and economic activities. (Forrester, 1994)
      5.2.3. Quantifying SD models make the reasons for change visible and testable.
      5.2.4. Inferences are not reliable.
   5.3. Model validation is a gradual process of building confidence in the usefulness of a model; validity cannot reveal itself mechanically as a result of some formal algorithms. (Barlas and Carpenter in Lane, 1996)
   5.4. Students should understand that:
      5.4.1. SD model building is an iterative process of conceptualization, formulation, evaluation, and policy recommendation. Each stage has recommended procedures for providing effective implementation.
      5.4.2. The visual nature of the model representation allows stakeholders to be part of the model-building experience, providing important input at each stage of the process and significantly enhancing the validity of the model and policy generated.
   5.5. It is essential to communicate SD modeling insights in a way that is clear and understandable by the general public.
   5.6. Students should appreciate that:

5.6.1. Including stakeholders in the modeling process is a significant advantage - improving the chance that policy decisions will actually be accepted and implemented.
5.6.2. Explaining model design and insights is an integral part of the modeling process.

6. Assumptions about the depth structure of phenomena
   6.1. Many SD structures are found repeatedly across disciplines. This supports transferability of structure for SD modeling.
   6.2. Students should recognize:
      6.2.1. Typical (simple) model structures that transfer:
      6.2.2. Exponential, S-shaped, Overshoot & Collapse, Oscillation, Spread of Epidemic.
      6.2.3. Archetypes: Fixes that fail, Shifting the burden, Escalation, Tragedy of the commons, Eroding goals, Limits to growth/success, Success to the successful, Growth and underinvestment.

An understanding of these core concepts could only occur if systems thinking (ST) and System Dynamics (SD) were infused throughout the K-12 curriculum. A subset of these core understandings could occur from less comprehensive ST/SD learning experiences.

The second detour is to look historically at those qualities that pre-college teachers who have incorporated/infused SD modeling into their curriculum have observed in their SD educated students.

What is different about precollege students who use SD as a tool in their learning?

The authors contacted some middle school and secondary school educators in the United States who have taught their students (for more than five years) to build system dynamics models to study simple and/or complex problems in their classes and asked them what differences in learning they (the teacher) observed in their students who studied using SD and those who did not use SD. The results are summarized below:

Table 1: Linking teacher observation of students who have used SD with educational learning goals.

Observed Effects	**Educational standards and learning goals**
Easier translation of real-world dynamic problems into more understandable mathematical representation: Students have less trouble translating a dynamic problem into stock/flow representation than they do into closed-form equations (J. Darkow, personal communication, August 22, 2017).	M1.2, M1.5, M1.6. S2.1.2, S.2.2.1, S2.2.2, S2.2.3, S2.2.6, EG1.1
Stronger quantitative reasoning: Models must meet certain criteria or they won't work. Models are built in context (A. Ticotsky, personal communication, August 15, 2017).	M1.2. S2.1.2, S2.2.1, S2.2.5

Begin to understand complexity: Students build models involving nonlinear dynamics, a characteristic of complex systems. Counterintuitive model behavior and transfer of loop dominance arise out of model students can build. Students can build and study problems that were out of their scope using traditional closed-form equation representation (A. Ticotsky, personal communication, August 15, 2017).	M1.1, M1.3, M1.6. S2.1.1, S2.1.2, S2.1.3, S2.1.4, S2.2.2
Builds understanding of dynamic behavior: Students gain an understanding of the fundamentals of dynamic behavior. They understand better what is needed to maintain a stable system. They understand how certain flow behaviors affect stock dynamics (Personal communications, J. Darkow, August 22, 2017, T. Joy, October 27, 2017, R. Quaden, October 23, 2017).	M1.1, M1.3, M1.6. S2.1.1, S2.1.3, S2.1.4, S2.2.2
Model diagrams help identify student misconceptions: When students display their stock/flow diagrams they are surfacing their mental models about how the system problem works. This allows other students (and the teacher) a chance to give feedback and helpful suggestions on the model design. The focus is on improving the model design. The modeling process allows the computer to become the arbiter of what works. Students can test their mental models (Personal communications, J. Darkow, August 22, 2017, C. DeCarlo, November 2, 2017).	M1.1, M1.4, M1.6. S2.1.1, S2.1.2, S2.2.7, EG1.2
Reinforces concept that structure determines behavior: Students connect structure and behavior of systems (R. Quaden, personal communication, October 23, 2017).	M1.1, M1.6. S2.1.1, S2.1.3, S2.2.5
Students practice scientific method: Building models allow students to continually apply the scientific method – starting small, testing, revising, adding more detail, testing, revising- in a process that supports deeper understanding of the problem studied (R. Quaden, personal communication, October 23, 2017).	S2.1, S2.2
Learn to identify leverage points in a system: Students learn to analyze the feedback structure of a system, how to determine whether it is reinforcing or balancing, and how to test a model to determine leverage points. This helps them identify policies that might improve the behavior of a troublesome system (J. Darkow, personal communication, August 22, 2017).	S2.2.4, EG1.1
Strengthens interpretation of graphical results: Students are continually evaluating their models based on its graphical output. Moreover, students are using covariational thinking by looking at multiple time-series graphs to determine the effect of different parts of the system on the main variables of interest. They also start to pay attention to scaling of graphs to improve interpretation of results and communicating model behavior to an audience (Personal communication, J. Darkow, August 22, 2017, C. DeCarlo, November 2, 2017, R. Quaden, October 23, 2017).	S2.2.2, S2.2.5, S2.2.7, EG1.2
Students want to know "Why" it happens: Students are more likely to ask why they are getting different results, which allows teachers to dig deeper into the subtleties of the problem. With paper/pencil analysis students more accepting of an incorrect result. With modeling students initiate conversations about why their graph is different. This helps the	M1.1, M1.6, S2.1.1, S2.2.1, S2.2.7, EG1.2

teacher to address individual misconceptions (C. DeCarlo, personal communication, November 2, 2017).	
Students demonstrate deeper understanding of core concepts: A significant amount of learning goes into the creation of models. The hands on trial-and-error, peer communication, attention to communicating how their model works reinforce and strengthen understanding of core concepts triggered by the problem under study (Personal communication, C. DeCarlo, November 2, 2017, T. Joy, October 27, 2017).	M1.3, S2.2.1, S2.2.4, EG1.2
Support transferability of structure: Students begin to see generic structures as part of other models they build. Transferability of model design supports cross-discipline applications (Personal communication, J. Darkow, August 22, 2017, T. Joy, October 27, 2017, A. Ticotsky, August 15, 2017).	M1.3, S2.1.3
Builds increased intuition about limits: Running models with different DT values shows students the effect that calculating in smaller and smaller time intervals can have on model behavior (R. Quaden, personal communication, October 23, 2017).	
Strengthens teamwork skills: Leaders in the SD field recommend that models should be team efforts. Communication between team members focus on model design. Lots of peer coaching goes on (A. Ticotsky, personal communication, August 15, 2017).	EG1.4, EG1.2
Students are engaged: Usually, when students are working on models, almost everyone in the classroom will be on task, thinking hard and working on solutions. The model-building activities are hands-on. Learning is student centered. Students felt they were learning something important for their lives. Students were willing to work on problems for longer periods. Students seemed more engaged in school (Personal communication, T. Joy, October 27, 2017, A. Ticotsky, August 15, 2017).	EG1.1, EG1.2, EG1.5
Communicating more succinctly: The model diagram needs to be drawn neatly and with an eye for communicating the structure. The structure must be as simple as possible but sufficient to capture the behavior of interest. Students are expected to explain how and why their model has the structure they designed AND how the structure produces the model behavior. Explanations are both broad and deep and include references to: o Feedback and multiple contributing factors o Language about change and degrees of change o Contribute to looking at long term behavior o Used language of growth/decay, causal relationships, equilibrium, sustainability, linear/exponential change over time, reinforcing/balancing feedback, etc. (Personal communication, T. Joy, October 27, 2017, A. Ticotsky, August 15, 2017	S2.2.7, EG1.2, EG1.3

The list above provides a useful set of characteristics that calls for research to support the claims of these teachers. Some claims would not be difficult to justify. Building a set of assessment instruments to bring more statistical scrutiny to the claims is part of what our research project is attempting to accomplish. It is our hope that other researchers will

develop methods to validate some of these claims as soon as possible, adding to the body of evidence to support SD as a valuable educational experience for students.

Current evaluation tools (state or national exams) cannot assess the characteristics described in this section. It is necessary to focus on the broader goals mentioned in the previous section to guide our evaluation efforts.

Other Considerations for Developing Assessments

Not only do we need to design an assessment instrument that can provide evidence that students are moving toward the goals desired, but we need an assessment template that can:

- Be used across disciplines,
- Help generate assessment items that are relatively easy to administer,
- Help generate assessment items that can be scored in a relatively short period of time so a large number of students in a wide variety of classrooms can provide the statistical data needed to validate the new (SD) learning paradigm,
- Allow teachers who are not SD experts a relatively simple task of understanding how to score the assessment items.

In the past, SD teachers have used student-constructed models and student technical papers as evidence that students actually understood the more complex nature of the new problems they were addressing. But this approach is very time consuming and requires a great deal of SD skill on the part of the teacher. Other teachers have used observation techniques, student presentations of their (student) work, teacher generated tests/quizzes, scoring rubrics, student written narratives, among others (Fisher, 2017). But these approaches do not provide the rigor necessary to convince decision-makers that the SD learning approach is valid.

Another question that arises is how can one create a fair assessment instrument, for control group studies, when the new learning paradigm is so different from traditional learning strategies? It is like trying to assess the new dimension "up" in Flatland, while almost all of the control group of students would have no sense of what "up" means.

New Assessment Protocols?

One statement that has often been expressed by SD instructors is that their students "read the paper differently" or that they "ask better questions." That gave us an idea that we could focus on the higher level learning goals that we want for all students by determining if students really do ask better questions using SD analytical reasoning. Asking better questions seems to be a reasonable, fair approach to assess if the new learning paradigm is an improvement. And "asking better questions" is a desirable learning goal for all students in all disciplines.

Other promising research into science assessment talks about three-dimensional assessment that is being used to support the Next Generation Science Standards. In these assessments each assessment task has three categories of standards that are addressed: Science practices, cross-cutting concepts, and core ideas from the unit under study (NGSS Lead States, 2013). Learning activities are built to blend these three views of "doing" science (National Academy of Sciences, Engineering, and Medicine, 2017). All writing assessments use rubrics to score the student responses. Three-dimensional assessments drive the format and content of multiple learning activities used in the classroom to address certain goals.

Assessments surrounding the "Deeper Learning" strategies, funded by the Hewlett Foundation, are designed to "require that all students develop a broader set of competencies" (Rothman, R., 2011, p. 1) that encompass concerns about an increasing complex society in which they live. The deeper learning structure is based upon having students learn core academic content, engage in critical thinking about complex problems, work together in team, express their ideas well, and be able to focus on their own learning but be open to feedback from others (Rothman, R., 2011). One of the assessment instruments used in conjunction with deeper learning is the PISA (Programme for International Student Assessment) test, which started in 2000 and is used to measure the scholastic performance in mathematics, science, and reading of 15 year-old students, globally. It is designed to test application of knowledge rather than retention of specific content.

Next Steps

Generating a new assessment protocol is not an easy task. Some efforts in the previous section have merit and would be useful to consider for our task. Our research group does not, at this time, contain an assessment expert. We feel we need guidance from a researcher who specializes in assessment to help us wade through the new assessment strategies that seem relevant to our SD focus. We are currently trying to find funding to hire such an assessment expert and also to support the classroom testing that would need to occur using any new assessment items we might construct. These tasks will be left for the second half of this research project.

Conclusion

In this paper we presented some higher level US educational goals in several content disciplines that could apply more directly to the use of System Dynamics. We have also enumerated some of the core concepts that are at the heart of a System Dynamics approach to understanding complex problems. Finally, we have summarized those qualities that differentiate secondary school students who have built SD models as part of their instruction (in at least one subject) from those students who have not built SD models. These qualities were submitted by six secondary school (US) teachers who taught their students to build SD models in their courses for at least 5 years. An alignment of the (SD) student qualities with the higher level US educational goals was displayed in Table 1.

Stage two of this research project will focus on designing a discipline-independent assessment template designed to illuminate the differences between learning that occurs when students build SD models to study a specific problem versus the learning that occurs when students use more traditional instructional activities (not involving building SD models). Once the template is developed examples of more specific assessment items will be developed for 2 content areas and those assessment items will be tested in a few classrooms. We are in the process of trying to find funding to support the addition of two assessment experts to guide the committee in its work for stage two.

It is anticipated, due to the large scope of this research question, that an additional stage will be required. Stage three would involve developing specific assessment items for one or more additional subject areas and completing more extensive classroom testing. Funding, again, on a larger scale than stage 2 will be needed, perhaps from NSF.

Those of us who have taught students at the secondary and tertiary educational levels know that having our students build SD models enhances their learning about specific problems. At this time there are no broad-based assessments that capture the "deeper learning" that support our experience. Documenting the differentiated learning that occurs in SD infused instruction via research studies is one method that we believe will strengthen conversations among educational decision makers about the value of infusing SD broadly in education.

References

Bar-Yam, Y. (2015, September 26). Complex systems science: Where does it come from and where is it going to? *Conference on Complex Systems*. Tempe, Arizona. Video viewed at: www.necsi.edu/research/overview/ccs15.html.

Bar-Yam, Y. (2012, December 12). What is complex systems science: Opportunities and insights. *New England Complex Systems Institute online seminar*. Retrieved from: www.necsi.edu/events/vidlib/#Whatis.

Common Core Standards Math Practice. Retrieved from: http://www.free-test-online.com/ccss/hs.html. (March 5, 2019)

Fisher, D. M. (2011). "Everybody thinking differently": K-12 is a leverage point. *System Dynamics Review*. 27, (4): 394-411.

Fisher, D. M. (2017). The next 25 years in pre-college education: A move toward global understanding of complex systems. *Proceedings from the International System Dynamics Conference*. Cambridge, Massachusetts.

Forrester, J. W. (1985). The" model versus a modeling "process. *System Dynamics Review*, *1*(1), 133-134.

Forrester, J. W. (2009). Some basic concepts in system dynamics. *Sloan School of Management, Massachusetts Institute of Technology, Cambridge*.

Forrester, J. W. (1994, June). Learning through system dynamics as preparation for the 21st century. In *Keynote Address for Systems Thinking and Dynamic Modeling Conference for K-12 Education* (pp. 27-29). Concord Academy Concord, MA.

Forrester, J. W. (1987). 14 "obvious truths". *System Dynamics Review*, *3*(2), 156-159.

Grotzer, T.A., & Mittlefehdlt, S. (2012). Students' metacognitive behavior and ability to transfer causal concepts, In A. Zohar & J. Dori (Eds.) Metacognition and science education. Cambridge, MA: Springer

Joint Committee on National Health Education Standards. (2007). *National Health Education Standards, Second Edition: Achieving Excellence*. Washington, D.C.: The American Cancer Society. Retrieved from: https://www.shapeamerica.org/standards/health/ (Jan 26, 2019)

Kaestle, C. (2013). Testing policy in the United States: A historical perspective. *The Gordon Commission*. Retrieved from: https://www.ets.org/Media/Research/pdf/kaestle_testing_policy_us_historical_perspective.pdf. (March 5, 2019)

Lane, D. C., & Smart, C. (1996). Reinterpreting 'generic structure': evolution, application and limitations of a concept. *System Dynamics Review: The Journal of the System Dynamics Society*, *12*(2), 87-120.

National Academies of Sciences, Engineering, and Medicine. (2017). *Seeing students learn science: Integrating assessment and instruction in the classroom*. Washington, DC: The National Academies Press. doi: 10.17226/23548.

National Council for the Social Studies. (2014), *The College, Career, and Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History.* Silver Spring, MD. Retrieved from: https://www.socialstudies.org/sites/default/files/c3/C3-Framework-for-Social-Studies.pdf (Jan 26, 2019)

National Governors Association Center for Best Practices & Council of Chief State School. (2010) *Common Core State Standards (Mathematics)*. National Governors Association Center for Best Practices & Council of Chief State School Officers. Washington D.C. Retrieved from: http://www.corestandards.org/Math/Practice/ (Jan 26, 2019)

National Research Council. (2012) Education for life and work: Developing transferable knowledge and skills in the 21st century. The National Academies Press. Washington, DC. Retrieved from: https://www.nap.edu/resource/13398/dbasse_084153.pdf (Jan 26, 2019)

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press. Retrieved from: https://www.nap.edu/read/18290/chapter/12 (Jan 26, 2019)

Oregon Department of Education. (2011) Retrieved from: https://www.ode.state.or.us/teachlearn/subjects/science/assessment/sci_stands_verif_hs_sampleoib.pdf. (March 5, 2019)

Rothman, R. (2011). Assessing deeper learning. Policy brief. Alliance for Excellent Education. **ERIC** ED524137

Siegried, J., Krueger, A., Collins, S., Frank, R., MacDonald, R., McGoldrick, K. Taylor, J., & Vredeveld. G. (2010). *Voluntary national content standards in economics*. Council for Economic Education. New York, NY. Retrieved from: https://www.councilforeconed.org/wp-content/uploads/2012/03/voluntary-national-content-standards-2010.pdf (Jan 26, 2019)

Sterman, J. D. (2000). *Business dynamics: systems thinking and modeling for a complex world*. Irwin, McGraw-Hill. New York, NY.