Thinking about systems: student and teacher conceptions of natural and social systems

L. Booth Sweeney[a]* and John D. Sterman[a]

Linda Booth Sweeney received her doctorate in education from Harvard University and is currently the director of SoL's Education Partnership. Dr. Booth Sweeney lectures widely, translating the conceptual ideas of systems theory into programs, tools and materials for curricular, school and organizational transformation. She has written several books and articles designed to simplify understanding of systems theory and its related tools and has developed activities, on-line simulations and animations to provide visual and experiential examples of living systems-in action.

John Sterman is the Jay W. Forrester Professor of Management at the MIT Sloan School of Management and Director of the MIT System Dynamics Group.

Abstract

Many in the system dynamics community argue that children are natural systems thinkers. Here we study how middle school students and teachers think about everyday settings involving feedback, stocks and flows, time delays and nonlinearities, prior to any formal training in these concepts. We develop instruments to elicit understanding of systems concepts and test them with students and teachers from two middle schools in the U.S.A. We find, with some exceptions, generally limited intuitive systems thinking abilities. "Open-loop" or one-way causal thinking is common. Explanations lack references to time horizons and time delays. Significant misconceptions of stock and flow structures appear regardless of age. Teachers generally outperformed students, although one-quarter of the students performed at the median level for teachers. We discuss the nature of students' and teachers' intuitive models of dynamic systems, explore potential barriers to understanding dynamic systems, and discuss implications for effective teaching of systems concepts. Copyright © 2007 John Wiley & Sons, Ltd.

Syst. Dyn. Rev. **23**, 285–312, (2007)

Introduction

Systems thinking is widely considered essential in the effective management of complex dynamic systems at the core of problems such as poverty, environmental degradation, and climate change. For more than 10 years, K-12 educators in the U.S.A. and around the world have been integrating systems thinking and dynamic modeling into the curriculum and aligning systems thinking concepts and tools with state standards.

A large body of research, however, shows that even highly educated adults have poor systems thinking skills (e.g., Booth Sweeney and Sterman, 2000; Sterman, 1994; Dörner, 1980). What is the problem, and what can be done? One possibility is that systems thinking does not develop "naturally" as does, say, language, but must be developed through formal education, similar to calculus; on this view, schools need to introduce systems concepts into their curriculum. Alternatively, many educators argue (Brown and Campione, 1994; Senge *et al.*, 2000), that children are natural systems thinkers who can recognize interdependencies and interrelationships long before they are schooled in these concepts. While the world around them grows increasingly complex and interdependent, schools continue to fragment and compartmentalize, reinforcing

[a] MIT Sloan School of Management, System Dynamics Group, 30 Wadsworth Street, Room E53-351, Cambridge, MA 02142, U.S.A.
* Correspondence to: L. Booth Sweeney. E-mail: Linda_Booth_Sweeney@post.harvard.edu
Accepted July 2007

System Dynamics Review Vol. 23, No. 2/3, (Summer/Fall 2007): 285–312
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr.366

the notion that knowledge is made up of many unrelated parts and providing little opportunity for students to see recurring patterns of behavior across subjects and disciplines. On this view, formal education suppresses children's natural inclination to think about systems and more radical reforms are needed.

Does current schooling fail to develop people's intuitive systems thinking capabilities or does it actively suppress them? What should be done, and what curriculum materials, pedagogical reforms, and other changes can help? Answering these questions requires first that we understand the intuitive systems thinking capabilities of young people. Here we develop instruments to elicit student and teacher understanding of key systems concepts, and test them with a sample of participants from two middle schools in the U.S.A. We examine participants' naïve "systems intelligence", which we define as a layperson's understanding of how dynamic systems function prior to any formal education in such concepts. Systems intelligence combines *conceptual knowledge* (knowledge of system properties, structures and reoccurring patterns of behavior) and *reasoning skills* (the ability to locate situations in wider contexts, see multiple levels of perspective within a system, trace complex interrelationships, look for endogenous or "within system" influences, be aware of changing behavior over time, and recognize "homologies"—recurring patterns that exist within a wide variety of systems (von Bertalanffy, 1968). Many definitions of systems thinking and lists of systems thinking skills have been offered, e.g., Capra (1999), Espejo (1994), Hogan (2000), Meadows (1991, 2001), and Hämäläinen and Saarinen (2004). Here we assess participants' abilities to (i) recognize recurrent patterns of behavior in different domains, (ii) distinguish types of system structures, and (iii) make relevant policy recommendations.

The results contribute to the ongoing dialogue about effective means of developing systems intelligence and related reasoning skills (e.g., Doyle *et al.*, 1998; Booth Sweeney and Sterman, 2000; Ossimitz, 2001; Pala and Vennix, 2005). More important, the assessment tools and protocols developed here can be used by others to document the intuitive systems thinking capabilities of young people and adults in other settings, including other cultures and nations.

Our objective is not to determine whether study participants use the "right" terms to describe systems, but rather to understand the degree to which they appreciate the systemic features of the phenomena they are asked to consider. As the authors of *Benchmarks for Scientific Literacy* (American Association for the Advancement of Science, 1993) suggest: "The main goal of having students learn about systems is not to have them talk about systems in abstract terms, but to enhance their ability (and inclination) to attend to various aspects of particular systems in attempting to understand or deal with the whole" (p. 262). For example, we focused more on whether a student or teacher could describe the feedbacks involved in the growth and regulation of animal populations, and less on whether they used specialized terminology such as feedback, cycles or mutual interdependence to do so.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

We find that a significant number of participants, regardless of age, exhibit limited understanding of complex natural and social systems, for example, assuming open-loop causality where feedback exists, low awareness of the temporal dimension of dynamic systems and pervasive violations of fundamental relationships between stocks and flows. At the same time, we found a surprisingly rich variety of terms used to describe systemic phenomena, most not found in system dynamics texts. In addition, a number of students and teachers demonstrated proficient homological reasoning skills, e.g., the ability to recognize common patterns across natural and social settings.

The paper is organized as follows: we (1) review the literature on children's and adults' understanding of complex systems; (2) describe the development and testing of the "Systems-Based Inquiry (S-BI) method and protocol; (3) describe the schools where the S-BI was tested and the results for both teachers and students; and (4) discuss the results, consider limitations, and suggest extensions and refinements. In closing we raise a number of issues for scholars seeking to enhance the effectiveness of systems thinking education for children and adults. The online appendices (available at http://www.interscience.wiley.com/jpages/0883-7066/suppmat/sdr.366.html) provide full documentation of the S-BI protocol, coding criteria, scoring rubric, and other information needed to replicate, extend, and improve the tools and study.

Children's and adults' conceptions of complex systems

Scholars have examined how children and adults think about specific complex systems such as ecosystems, climate change, and business organizations. Others examine how people understand scientific concepts or manage particular (simulated) systems. Results show that students and adults find it difficult to trace causality beyond simple one-way connections (Grotzer and Bell Basca, 2000; Hogan, 2000) and most do not spontaneously close the loop when feedback exists (Green, 1997; White, 1997). It is common, for youths and adults, to consider close entities before distant ones (Doyle *et al.*, 1998; T. Grotzer, unpublished doctoral dissertation, Harvard University, 1993; Marek, 1986; Moxnes, 1998; Vennix, 1990; White, 1997), focusing on "immediate effects" and ignoring or not noticing "extended, indirect effects" (Grotzer, 1997; Perkins and Grotzer, 2000; Grotzer and Sudbury, 1998). People (both youths and adults) tend to assume centralized control must be at work to explain regular, apparently purposeful behavior in decentralized, self-organized systems such as a flock of birds or the Internet (Resnick, 1995; Resnick and Wilensky, 1998), and students tend to confuse levels of description when describing a system with many interacting parts (Wilensky and Resnick, 1999). People do not perform well in dynamically complex simulations of markets (Bakken *et al.*, 1992; Sterman and Meadows, 1985), management systems (Diehl and Sterman, 1995; Sterman, 1989), educational systems (Hirsch,

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

1998), and ecosystems (Jensen, 2003; Wagenaar and Timmers, 1979; White, 1997). Adults tend to assume cause–effect relations are linear (Koplowitz, 1984), do not readily recognize feedback (Sterman, 1989) and stock and flow structures (Booth Sweeney and Sterman, 2000; Sterman and Booth Sweeney 2002, 2007), and have difficulty when cause and effect are distant in time and space (Doyle *et al.*, 1996; Vennix, 2000; Dörner, 1980).

Despite these generally discouraging results, some people are able to think systemically. Developmental theory and empirical research offer some evidence that the development of systems intelligence is not age-related. For example, some research (Brown and Campione, 1994; Kates and Katz, 1977) and anecdotal evidence (Senge *et al.*, 2000) shows that children as young as kindergarten age are able to grasp, at a rudimentary level, systems concepts and tools. Studies with students between the ages of 7 and 12 (RG Boutilier, unpublished PhD disserta-tion, University of British Columbia, 1981; Chandler and Boutelier, 1992; Grotzer, 1993; Wylie *et al.*, 1998), middle school students (Roberts, 1975/1978), high school students (Jackson *et al.*, 1996; Ossimitz, 1996), and college students (d'Apollonia *et al.*, 2002) reveal students' nascent comprehension of the behaviors and characteristics of dynamic systems. Yet our intuitive models of the structure and behavior of natural and social systems remain poorly understood. As Resnick and Wilensky (1998) observed: "There has been very little research in the developmental and cognitive communities on how people make sense of complexity" (p. 170).

Method

The instrument

The Systems-Based Inquiry (S-BI) protocol was developed to surface participants' intuitive models of complex system dynamics. The S-BI protocol consists of 17 probes organized into four major sections and designed to surface participants' intuitive models of feedback structures, nonlinearities, time horizons, time delays, and stock and flow structures (see online Appendix A for the complete S-BI protocol):

- *Part I: Systemic scenarios:* Participants consider system dynamics in six simple scenarios. These scenarios emphasize feedback dynamics.
- *Part II: Homology challenges:* Participants imagine six related but different systemic scenarios. This requires participants to use homological reasoning.
- *Part III: Loop comparison:* Participants compare the six scenarios in terms of underlying dynamics. They are asked to consider and describe similarities and differences between feedback loops.
- *Part IV: Policy thinking:* Participants consider the consequences of adding an additional lane to a highway to relieve traffic congestion and make policy recommendations.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

A coding guide was developed and tested. Student and teacher responses were coded and assigned to one of five levels on a scale of systems intelligence, ranging from Level 0 (*no response*) to Level 4 (*fuller utilization of systemic reasoning*). We refined the protocol through pilot testing with 12 middle school students to eliminate ambiguities, ensure that the questions were age-appropriate, and limit the time required (Sweeney, 2004, provides details).

In Part I, participants are asked to think aloud about six natural and social system scenarios designed to surface conceptions of cause-and-effect relationships and behaviors that can "feed back" to form reinforcing or balancing processes. The scenarios are:

A. *Wolves and rabbits:* Wolves (predators) and rabbits (prey) are linked in a balancing feedback. Understanding of predator–prey interactions has been studied extensively in children (Hogan, 2000; Marek, 1986; Webb and Boltt, 1990) and adults (Green, 1997; Jensen, 2003; White, 1992).
B. *Teacher perception and student achievement and self-esteem:* Participants are asked to talk about the interrelationships between a teacher's perception of a student and the student's perception of him or herself. Reinforcing feedback plays a role in this relationship: e.g., the positive expectations of the teacher may foster positive self-perceptions by the student, which in turn fosters even greater expectations by the teacher; or, visa versa.[1]
C. *Hunger and eating:* These are linked through balancing feedback and exhibit goal-seeking behavior: as hunger increases, one eats, decreasing hunger. This scenario was chosen for its universality and prior successful use with fifth and sixth graders (Roberts, 1978).[2]
D. *Births/population:* Population growth arises from a reinforcing feedback. Specifically, populations grow when the reinforcing birth feedback dominates the balancing deaths feedback; the balance between these loops depends nonlinearly on the population relative to its carrying capacity. Several studies (Griffiths and Grant, 1985; Roberts, 1978; Summers and Summers, 1976) have shown that population dynamics is an accessible topic for middle and high school students.
E. *Practice, performance and enthusiasm:* Achievement and motivation can be linked by reinforcing feedback; for example, the amount one practices a sport or musical instrument affects how well one plays, which affects the enthusiasm one has for that activity, which feeds back to the willingness to practice. This scenario was chosen because the topic is relevant to most middle school students.
F. *Room clean-up and parents' attitude:* Participants are asked to consider the relationship between the state of a their room (clean or dirty) and the attitude of their parent or caregiver (happy or upset), forming a balancing feedback. This scenario was included because of its universal appeal.[3]

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Three scenarios involve positive feedback loops (population–births, teacher expectations–student achievement, practice–performance) and three involve negative feedback loops (wolves–rabbits, hunger–eating, messy room–clean-up). Throughout Part I, each participant was first asked an open-ended question, followed by two structured questions and one "homology" question (Part II of the S-BI protocol). We use the wolf–rabbit scenario to illustrate (Table 1).

Table 1. S-BI protocol Parts I and II: interview questions

S-BI interview prompt, Level 1:
(Open-ended question): Certain animals (predators) prefer to eat other animals (prey). *I'd like you to tell me about the relationship between wolves and rabbits (one of the wolf's favorite foods). For the purposes of this story, the wolves' other favorite foods are not available.*
S-BI interview prompt, Level 2:
How might wolves and rabbits be interrelated or connected? What do you think happens over time?
S-BI interview prompt, Level 3:
What might happen next? What if there are more wolves? If the wolf population decreases, then what happens to the prey population? How long do you think it takes for this increase or decrease in wolves or rabbits to happen?
Part II: S-BI interview, homological reasoning prompt:
How might this scenario be similar to another scenario in a different setting?

In Part III participants consider pictures of two feedback situations they just discussed (including predator–prey, performance–enthusiasm, teacher expectation–student self-perception and room clean-up (see online Appendix B for complete set of pictures). They are then asked the following questions:

a. Interviewer points to the two balancing loop examples, asking "Do you think these are similar or different? How?"
b. Interviewer points to the two reinforcing loop examples, asking "Do you think these are similar or different? How?"
c. Interview points to one reinforcing and one balancing loop example, asking "How about these two, are they similar or different? How?"

Finally, in Part IV participants are presented with the following scenario:

> City Planners in San Francisco, looking to make life easier for city motorists, decide to build more traffic lanes on the Bay Bridge. This bridge is notorious around commuting time for traffic jams and congestion.
>
> How do you think this will affect traffic patterns on the Bay Bridge?
>
> How would you recommend City Planners manage traffic congestion?

For the S-BI scoring rubric, see online Appendix C.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Participants and setting

Student participants were drawn from a previously studied sample of 105 students from two Berkeley, California middle schools. These 105 students were interviewed by the first author in the fall of 2001 as part of a study evaluating a Food Systems (FS) curriculum created by the Center for Ecoliteracy (CEL) (Murphy and Schweers, 2003). The CEL is a non-profit educational organization in Berkeley, California founded by Fritjof Capra, David Orr, and Alice Waters. The CEL works with educators and students to promote sustainable living by fostering understanding of nature through direct experience, for example, through a school garden. The 105 students were chosen based on the following criteria: students who have been in their school for at least one full school year and who are English speaking (or have at least one English-speaking parent). From this population, 30 students were randomly selected from two different middle schools (15 from each); one dropped out. Twelve accredited science teachers and two non-science teachers (garden managers) from both schools were also recruited. Three teachers dropped out due to school obligations, leaving 11 participants. The middle school students ranged in age from 10 to 12 (average 10.8 years) and the teachers ranged in age from 24 to 60 (average 36.5 years).

Descriptive data collected for a related longitudinal study (Murphy and Schweers, 2003) show the two schools to be similar in terms of socioeconomic status, racial mix, and academic achievement. Prior to the study, students from the FS school participated in an experiential ecology curriculum, while students from the other school participated in one semester of a standard natural sciences curriculum. The inclusion of two schools allowed us to ask whether a curriculum focused explicitly on ecosystems (as experienced by the FS students) improved students' understanding of and reasoning about natural and social systems. Including teachers allowed for a comparison of teacher and student responses.

Interview process and data analysis

The S-BI protocol was administered through one-on-one, semi-structured interviews roughly 45 minutes in length. All interviews were audiotaped, transcribed, and coded. Various analyses were performed to assess instrument validity, including classical item analysis, corrected item–total correlation, Cronbach alpha, factor analysis, and Rasch analysis. These tests show that the S-BI protocol possesses good psychometric properties. For example, the protocol has a high degree of reliability, as exhibited by the Cronbach alpha measure of 0.94. To establish content validity, transcript excerpts were reviewed by experts including the second author, George Richardson (SUNY, Albany), David Lane (London School of Economics), Mitchell Resnick (MIT, Media Lab), and Uri Wilensky (Northwestern University).

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Several background variables were collected for each student. These included the students' GPA (grade point average) and scores on an "Eco IQ" test. The Eco IQ test, administered the previous semester to the same sample of students as part of an evaluation of the FS program, assesses students' knowledge of and attitudes about ecosystems such as the dynamics of watersheds and why food webs are important. For more information on the Eco IQ test see Murphy and Schweers (2003).

Results

A number of common themes emerged from the 40 interview sessions with students and teachers. We report the most significant findings on the following themes: intuitive models of feedback structures, time, stock/flow structures, recognition of feedback structures, homologous reasoning, policy thinking. We close with a summary of results related to quantitative analyses.

Intuitive models of feedback structures

Students and teachers generally do not describe feedback processes in situations where feedback is present. Spontaneous recognition of feedback structures was relatively rare; only 15% of the students and 32% of the teachers spontaneously "closed the loop" when presented with scenarios involving feedback structures. The students and teachers who recognized feedback structures frequently used similar words and phrases such as "spiraling", "chain effect", "zig zag", "up and down", "like a cycle", and "self-fulfilling prophecy" to describe the behaviors within the S-BI scenarios (Table 2).

Cycle-related phrases were most commonly used to describe feedback for students (n = 13) and teachers (n = 10). Middle school students learn about water and weather cycles in natural science classes. Students and teachers both named various types of cycles; for example, students made reference to the *life cycle*, *food cycles*, *water cycle*, and "*up and down*" *cycles*. While teachers referred to some of the same cycles, they tended to refer to more complex and abstract cycles such as *vicious cycles*, *the nitrogen cycle*, *self-imploding cycles*, *self-perpetuating cycles*, *a cycle of self-defeat*, *cycle of plant diversity*, *sleep cycles* and *the poverty cycle*.

Participants' use of "cycle" fell primarily into two categories (Table 3). The most common usage of "cycle" was to describe a repeated sequence of events (e.g., "it happens over and over again"). For example, "Cycle means . . . in a circle. Like when you're hungry you eat, and it goes in a circle. It keeps on going forever." The problem arises when "cycle" is used to describe balancing or reinforcing feedback. To describe a predator–prey relationship, as "a cycle that keeps going around" as one might describe the lunar cycle, confounds the description of behavior (rising and falling predator and prey populations) and

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

system structure (the reinforcing and balancing feedbacks that generate that behavior).

Three students and six teachers revealed an intuitive understanding of the behaviors of balancing and reinforcing feedback structures (level 3 on the S-BI rubric).

Table 2. Partial list of phrases used to describe feedback structures

	Reinforcing feedback	Balancing feedback
Students	• "it just keeps getting worse" • "one thing pushes up the other. It's a chain effect" • "it keeps going up and up" • "negative emotions . . . lead to a downfall" • "and the grades get lower and lower . . ." • "it just gets better and better and better" • "it zig zags down"	• "they bounce off each other" • "it's a balance"; "it balances out" • "they level each other out" • "the population would just keep on getting lower, until it leveled at the number of births" • "it goes up and down and up and down" • "it's an up and down cycle"
Teachers	• "spiral" and "spiraling" • "reinforcing"; "like a reinforcing loop" • "an ever growing cycle" • "that is like exponential growth" • "it's a vicious cycle" • "they build on each other" • "they feed off each other"	• variations on "balance": balancing, balanced out, a balancing act • "regulating", "self-regulating" • "it's a closed loop" (suggesting the amount stays relatively the same) • "it's self-maintaining" • "that is like supply and demand" • "it's in a flux"

Table 3. Participant use of term "cycle"

Cycle as repeated sequence	Cycle as feedback loop
"A cycle *is like when it goes around, so this thing turns into this thing,* which turns into this thing and then it ends up being back at the thing that you started with." (Student) "[A cycle] is a sequence of events that happens over and over again." (Student) "Something that tends to go over and over again, is sort of cyclical." (Teacher) "Well the *cycle* just keeps on repeating. Death, birth, death, birth. Which allows a normal population level." (Student)	"*That would be like a positive cycle* where one thing reinforces another." (Teacher) "Like the wolves would make sure that there aren't too many rabbits. WHY? That's like *how the whole cycle* works. WHAT IS A CYCLE? They're animals that are predators and they eat animals that are prey and sometimes prey eat other animals. And so it . . . *keeps everything in balance*." (Student) "These two seem like a cycle where it's reinforcing. A change here causes a *positive change* here, so it's not up and down. And the same goes for populations. The more births, the more population, the more births." (Teacher)

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Participants who did not attend to feedback structures focused instead on power dynamics within the scenarios ("They're the same because they both have someone telling you what to do"), emotional issues ("I eat even when I'm not hungry" or "sadness eating"), and taxonomic similarities ("They both have fur" and "They are both mammals").

The ability to recognize feedback loops ranged from recognition only of open-loop causality to coupled feedback loops. Table 4 shows four levels of open-loop to closed-loop responses to the wolf–rabbit scenario (S-BI protocol, Part I).

A key skill in systems thinking is the ability to expand the boundary of one's mental model to increase the range of feedbacks and factors considered (Sterman, 2002). However, students and teachers alike tended not to describe factors outside the immediate boundary described by the elements of the scenario. Only the four participants who provided the highest-level responses also noted the role of additional influences such as weather patterns, availability of food sources, the impact of pollution and human interventions.

Summary of intuitive models of feedback structures

Table 5 summarizes the participants' intuitive models of feedback dynamics.

Two students and nearly half the teachers ($n = 5$) attempted to describe the nature of growth in the population–birth scenario. Their ability to identify the positive feedback loop driving population growth, and the ability to describe the resulting behavior, particularly the exponential character of population growth, was poor. Of course we do not expect middle school students to know what exponential growth is mathematically or to be familiar with such terms. The question is whether they appreciate the core idea of positive feedback: that the greater the quantity of something (for example, the amount of money in a bank account) the greater the rate of growth of that quantity, so that, absent other feedbacks, the state of the system grows at an accelerating rate. Most students did not show such understanding. The most common type of student response is illustrated by response A:

A. "There'll be more people, cause there'll be more people being born and it'll make the population more bigger."

A few showed far greater sophistication:

B. "The main thing is, counting births isn't like counting bricks. Let's say this person has twins. That's two. That makes two more people in the world, plus the other two parents. Now these two may go on to have another two more each. It keeps splitting up farther and farther making more and more."

In response A, the student recognizes the reinforcing feedback linking population and the number of births. The student does not refer (at any level of

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Table 4. Levels of feedback thinking

S-BI feedback Level 1: open loop *When asked an open-ended question about the relationship between a predator–prey pair, this student notes a one-way connection between wolves and rabbits. She is aware that there may be a reciprocal connection between rabbits and wolves, but cannot name it. At the second prompt, this student was able to close the loop*	"The wolf needs the rabbit to eat and I don't know how the rabbit needs the wolf though."
S-BI feedback Level 2: closed loop *This student "closes the loop" by describing the mutual relationship between rabbits and wolves. She does not, however, describe the behavior of this feedback structure over time, for example, what would happen next when the rabbit population decreases*	"The rabbits are the wolves' only prey and so the wolves are probably going to be the rabbits' main predators and if there's a lot of rabbits, then there'll be more wolves. And if there are a lot of wolves, then there probably wouldn't be as many rabbits."
S-BI feedback Level 3: behavior of closed loop over time *This student closes the loop, continues to trace causal relationships around the loop and describes the behavior of the feedback loop, noting that the oscillating behavior of the predator–prey relationship continues to "bounce off each other" over time*	"The number of wolves increases and then the rabbits decrease. And then the wolves decrease and the rabbits increase. It sort of bounces off each other."
S-BI feedback Level 4: multiple closed loops *Student describes balancing feedback (fewer wolves, higher rabbit population, greater wolf population), then observes that feedback structure underlying wolf population growth hits a limit as the rabbit population decreases. He implies a related balancing loop (wolves increase, rabbits decline, wolves eventually decline and rabbit population "grows more again")*	"If there are just a few wolves and lots of grass, then the rabbits get a lot to eat and there is a much higher rabbit population. Then the wolves get more to eat and then they get a much higher population. The rabbits are dying off, the wolves' population drops, then the rabbits grow more again. Pretty simple cycle."

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Table 5. Summary of intuitive models of feedback structures

Model of feedback dynamics	Illustrative quote
One-way link between elements or agents	"The wolves eat the rabbits." (Student)
Two-way (or mutual) causality between elements or agents	The wolves eat the rabbits and if the rabbit population goes down, that hurts the wolves." (Student)
Chain effects: suggests a chain reaction between elements or agents but does not suggest a change in levels or amounts	"If there's not enough wolves or rabbits or there's too much of each*, it affects the whole chain of things.*" (Teacher)
Participants also spoke of "chain effects" that were amplifying or increasing	"One thing pushes up the other. It's a chain effect." (Student) "The more you perform there's *a chain reaction* that contributes to increasing your enthusiasm. So . . . *they feed off each other.*" (Teacher)
Growth in terms of levels or steps	One teacher suggested that her musical performance improved in "multi steps"
Wave-like models	Participants who used phrases such as "it's a balance" or "it balances out" often waved their hands up and down to suggest oscillation
Feedback loop models emerged from students and teachers who described this type of relationship: *x* causes more/less of *y*, which causes more/less of *z*, which causes more/less of *x*	"If there are less wolves then the rabbits can grow and then eventually there will be more wolves, but more wolves will eventually mean less rabbits and then less wolves. It goes on like that." (Teacher)

prompt) to the rate at which the population grows or the idea that the growth would accelerate over time. In response B, the student described the doubling characteristics of exponential growth by drawing a branching diagram that looked like a family tree.

Growth of real quantities cannot go on indefinitely but encounters limits. Systems practitioners have long extolled the virtues of recognizing "limits to growth" patterns. For example, Meadows (1982) has argued that real insight into complex system behaviors comes not only from recognizing the factors that produce growth, but also through shifting one's attention to the "next potential limiting" factors. This, according to Meadows, is "to gain real understanding of and control over the growth process" (p. 24). Intuitive models of nonlinear structures (including limits to growth and shifting dominance) were described by nearly one-third of the students (n = 8) and one-third of the teachers (n = 4). The following quotes shows students' intuitive understanding of a reinforcing process hitting a limit:

> If you practice a lot, usually you get better. But for some people if you practice too much you get burned out on it and you don't want to do it anymore. (Student)

> If you're very enthusiastic, you tend to want to practice more, and if you practice more your performance goes up, then your enthusiasm would go up. But in the end,

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

> your enthusiasm would probably die down. Then the amount of practice goes down, level of performance would go down, after a while, you just keep getting better only sometimes you get a little worse. (Student)

Some accounts were more sophisticated. This teacher uses her intuitive understanding of "limits to growth" as a guiding heuristic (scored as Level 4) in reasoning through Part I of the S-BI protocol:

> I was thinking of something that's successful. Say you have a school. If it becomes a very successful school, people really want to send their kids to that school. Then all of a sudden if they don't increase their resource base, they're not providing the thing that attracted everybody in the beginning. So you either have to check that growth, prepare for that growth, or you have to spread out that growth.

Intuitive models of time

Most students and teachers did not spontaneously describe specific time horizons or the impact of time delays. Teachers were twice as likely to describe time-related dimensions as students. Table 6 shows the different types of references to time in participant accounts.

In the Time Level 3 quote, the teacher characterizes the nature of a real-life problem (e.g., the need to "be instantly something") and calls into question society's epistemological assumptions about time. In doing so, the teacher demonstrates the highest S-BI level. Studies have shown (see Fischer and Pruyne, 2002; Kitchener and King, 1990) that this type of higher-order reasoning skill begins to emerge in early adulthood (i.e., 19–20 years of age). No students in this study spontaneously demonstrated Time Level 3 awareness.

Intuitive models of stock and flow structures

Understanding the relationships between stocks and flows is fundamental in dynamics. However, people's ability to recognize and understand stock–flow structures is poor. For example, highly educated adults at elite universities have great difficulty inferring the behavior of a stock from knowledge of its flows (Booth Sweeney and Sterman, 2000; Sterman and Booth Sweeney, 2007; Cronin and Gonzalez, 2007). Two common misconceptions of stock and flow structures emerged in the S-BI:

1. *Assuming stocks can be rising (or falling) without explicitly considering the outflow rate*. Meadows (1991) refers to this conception as an "inflow focus". For example, as Meadows explains, most people assume that to increase a stock you have to increase the flow into that stock (e.g., to accumulate more money in the bank, you have to increase the rate at which you deposit money). However, you can also increase a stock by reducing the outflow (reducing expenditures by saving more). The Inflow-Focus model was exhibited by nearly

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Table 6. Four levels of time dimension awareness

Time Level 0: No reference or discrete reference to time At the open-ended prompt, this student explains the relationship between hunger and eating, but does so in a way that does not refer to a time dimension	"If he likes it he'll probably eat a lot and if he didn't he'd probably eat a little. The hungry person probably would keep eating, cause he's hungry."
Time Level 1: Non-specific references Non-specific references to behavior over time (time delays, rates of change). When prompted, the same student suggests that after some time delay, the boy will be hungry again, "later on". This is an example of a non-specific reference to the change of hunger levels over time	"He'll probably eat again and he's probably going to be hungry again, cause he'll be hungry later on."
Time Level 2: Specific time references References to specific intervals of time (e.g., "it takes a year") over which a system's behavior unfolds, a delay occurs, rates of change ("50 births/day") or the impact of a decision is seen. At the open-ended prompt, this teacher brings a specific time dimension (i.e., "20 minutes") to her explanation of the time delay inherent in hunger and eating	"There's a biological response to eating . . . it takes about 20 minutes for your body to know that you really are not hungry. So unless you eat incredibly slowly, which most people don't, or you wait, you actually eat more than your body needs."
Time Level 3: Demonstration of fuller time dimension awareness References to the significance of considering time-related dimensions. Here, a teacher describes the value of and obstacles to taking a long-term view (Time Level 3 awareness) when learning how to play a musical instrument	". . . it's hard in this society that we live in, where people want to be instantly something. They want to be instantly rich or instantly good at something. Instantly feel better. I think there's something about looking over time . . ."

half of the students ($n = 18$) and teachers ($n = 5$) at the open-end prompt. These participants did not attend to the relationship between the rate of flow into or out of the stock and the rate of change of the stock. For example, in response to the birth/population scenario, this student describes the interrelationship between the number of babies born in Berkeley, California and population size:

> There are thousands of people that live in Berkeley, because Berkeley is a big city. And I think the number of births . . . there are at least 50 people having babies every day. So the population gets bigger.

Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/sdr

This student (responding at the open-ended prompt) describes only the inflow to the stock of population, and does not mention the outflow from the stock (deaths), so the inference drawn by the student that births cause "the population gets bigger" does not follow: it is necessary to know whether births exceed deaths (or, more precisely, whether births + in-migration > deaths + out-migration) to determine whether population is rising or falling. Perhaps, however, the student understands stocks and flows but makes the implicit and reasonable assumption that births are greater than deaths. At the second-level prompt, however, the same student again focuses primarily on the inflow:

> *What if the number of babies being born went down*?
> The population would get a little bit smaller.

This response clearly violates the basic principles of accumulation. Population falls only if deaths exceed births. Declining births can still lead to rising population (if births exceed deaths).

2. *Assuming inflow equals outflow*. An illustrative example of this model is: "The population will stay the same because each day a baby is born and each day somebody dies." In the following quote, we see a more accurate conception of population in equilibrium. This student describes the impact of inflows and outflows on a stock (population) and appears to grasp the concept of feedback from the stock (the population) to the inflow (number of births). She also grasps the concept of net inflow (births minus deaths) as an influence on the level of the population stock:

> The more fully-grown people there are, the more births there's going to be. But then fully-grown people die. So basically, there's someone dying every like three point fifteen seconds. And there's someone being born around that too, but with the human population actually more people are being born.

This student's observation that "with human population, actually more people are being born" reveals an understanding that an accumulation of people occurs if the inflow is greater than the outflow. This student also demonstrates some understanding of rates (e.g., "someone dying every like three point fifteen seconds"), maturation time, and time delays caused by the aging process. More than half of the students ($n = 18$) and 45% of the teachers ($n = 5$) exhibited inflow-focus, and 24% students ($n = 7$) and 18% teachers ($n = 2$) assumed inflow = outflow, respectively.

Recognizing feedback structures

In Part III, participants are presented with causal loop diagrams of the simple feedback loops operating in the six scenarios described above and are asked

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

"Which loops are similar?" "Which are different?" and "Why?" To what degree were students and teachers able to distinguish between the different feedback structures? Without any formal "systems thinking" education, approximately one-quarter of the students and half the teachers demonstrated a high-level ability (i.e., S-BI Level 2) to recognize and distinguish between balancing and reinforcing feedback (Table 7). Four of the eight

Table 7. Student and teacher examples of feedback recognition

Feedback Recognition Level 0 This student identifies hunger-eating (balancing) as similar to practice-performance (reinforcing). This was scored at level "0" because: a) the student picked two structurally dissimilar scenarios and b) the explanation refers to surface features of the two scenarios and not to their underlying structures.	"If you're hungry and you're on your way to Sizzler [a restaurant], you feel very enthusiastic that you're on the way to Sizzler. And you'll try to perform in a manner that . . . it's like you won't burst out and say I'm very hungry and you're already on your way to Sizzler. So you wait."
Feedback Recognition Level 1 In this quote, the student gives a coherent explanation for the similarity between the room clean-up loop (balancing) and performance–enthusiasm loop (reinforcing); i.e., they have the same underlying cyclic structure ("they both go in a circle"). He/she does not appear to appreciate the different feedback behaviors of the two loops	"With this one (room clean-up) it's like you do something that leads to another thing that leads to another thing and it kind of goes into a circle. And it's the same with that one (performance–enthusiasm)."
Feedback Recognition Level 2 *A. Recognition of structural similarity between two feedback loops* Here a student accurately describes the underlying feedback pattern between the practice–enthusiasm–performance loop and the number-of-babies–population loop	(Refers to practice loop): ". . . when you practice a lot and then you start to get good at it, that makes you want to play it more and you'll keep doing that over and over. That is similar to this one [refers to birth–population loop], when more babies come, more people will be created and when more people are created, they're having more babies and that'll keep on going around and around."
B. Recognition of structural similarity between two balancing loops Scored as Level 2 because he/she accurately describes the "bigger then smaller then bigger" dynamic created by balancing feedbacks with delays, and recognizes that the "better and better and better" dynamic of the reinforcing loops is different	(Compares wolves–rabbits with room clean-up): "The wolves eat the rabbits and the rabbits get bigger . . . [here] someone's telling you to go clean up and then you get it cleaned up but then it gets messy again, so it's like a loop. Like that one [refers to clean-up]. It keeps going around like a circle. These two things get bigger, then smaller and then bigger. Those [points to r-loops] just get like better and better."

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

participants who scored high on "feedback recognition" also scored at the highest level on Part I (assessing closed-loop thinking). The remaining four participants (those who did not spontaneously describe feedback when presented systemic scenarios in Part I) were able to distinguish between different feedback structures when given a visual depiction. This finding suggests that the interview probes provided in Part I as well as the visual depiction of the feedback structures in each scenario helped the remaining four students to recognize and distinguish between feedback structures.

The remaining students and teachers tended to focus on the surface features of the scenarios (Table 8). "Population" was the most frequently used cover story. For example, students (n = 9) and teachers (n = 4) concluded that the wolves–rabbits scenario (exemplifying balancing feedback) and the births–population scenario (exhibiting reinforcing feedback) had the same underlying structure because they were both about "population".

Table 8. Comparison of homologies and cover story explanations

Focus on structure	Focus on cover story
Student: Recognizing structural differences	**Student: Focusing on cover story**
"These are different, because the number of births, it just keeps on adding to the population. And it doesn't really drop down that much. But with the number of wolves, the rabbits drop down and then the wolves drop down and then the rabbits drop down and it keeps on going up and down."	"The wolves and the population are similar because it's about population and about how many wolves there are in relation to how many rabbits there are . . . If there are more births, then the more people there are. Then say the more rabbits there are, the more wolves there are.
Teacher: Recognizing structural differences	**Teacher: Focusing on cover story**
"The wolves–rabbits and the birth–population. They are different because wolves and rabbits are an up and down cycle. So more rabbits equals more wolves equals less rabbits equals less wolves. Whereas births and population is a reinforcing loop. A spiral up . . . where there is more births there is more population and then more births. So this feedback loop continues and this one regulates in an up and down cycle."	"*Rabbits–wolves and population* . . . are related. We're talking about the effects of abundance on a population. If there was a huge boom in the rabbit population, that would give the wolf population a leg up, they would be healthier, they'd be a stronger population, their population would grow. Then all of a sudden, if you have so many wolves that the environment doesn't support, then it's the same thing . . . the erosion of the foundation of the whole thing. You have your boom of births, but after a while Berkeley just gets so overcrowded that there is not enough to sustain that population, the bottom falls out of everything."

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Homologous reasoning

To what extent are students and teachers able to spontaneously conceive of homologous structures (those with different surface features but the same underlying feedback structure)? In Part II of the S-BI interview protocol, participants were asked to imagine a situation involving a similar dynamic to those described in each of the six scenarios (e.g., room clean-up), but from a different domain. For example, after describing the relationship between wolves and rabbits, participants were then asked to describe another situation that seemed similar, but was preferably not about other animals. It is this ability to recognize deep structural similarities in diverse domains that many psychologists (Doyle, 1997; Nisbett *et al.*, 1987), social theorists (Capra, 1999; Odum, 1977), system dynamicists (Forrester, 1993; Sterman, 2000), and educational researchers (Brown and Campione, 1994; Bruner, 1960/1977; Grotzer and Bell Basca, 2003) argue can facilitate the transfer of insights from one context to another.

We found that 33% of student and 77% of teacher responses demonstrated the ability to recognize homologies within and across domains. Fewer than one-third of the students (21%) and more than half of the teachers (53%) described homologies that were distinct in terms of the domain. Student-generated homologies include over-fishing and coral reefs as a homologous structure for predator–prey relationships, and video games and amount of play as a homologous structure to hunger and eating (Table 9).

Policy thinking

At the close of the S-BI interview, participants were asked what might occur if a lane were added to a highway as a means of reducing traffic congestion. The majority of teachers and nearly half of the students anticipated some unintended consequences of a road-building program. 43% of the students and 81% of the teachers concluded that an additional lane would not improve the traffic congestion problem, while 31% of students and 18% of teachers offered a "more lanes means better traffic flow" response.

Some students (n = 12) and teachers (n = 10) provided thoughtful policy-related solutions, such as: reduce public transportation fares (suggested by three teachers); extend the hours of operations (suggested by two students and four teacher); add car pool lanes (suggested by six students); raise tolls (suggested by five students and three teachers); and build another mode of transportation across the bay (suggested by one student).

Summary of quantitative analyses

Students' mean scores were higher at the comparison school than the (FS) school, but the differences in all cases are quite small and not statistically

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Table 9. Examples of student (s)- and teacher (t)-generated cross-domain homologies

S-BI scenario	Examples of participant homologies
Predator/prey	Environmentalists and crisis attitude (s) Coyotes/farmers/gophers (t) Supply/demand—job market (t) Over-fishing and coral reefs (s) Pre-teen leaders and followers (t)
Teacher/student expectations	Fish farming/amount of ocean fish (s) Plant cultivation/weeding (t) “Response to September 11—escalation (t) Person and their dog (s) Sibling/attitude (s) Poverty/availability of resources (t)
Hunger/eating	Growing cycle (life and death) in garden (t) Mixing colors (s) “Life’s ups and downs” (t) Teaching kindergarten/burn-out (t) Fatigue and sleep (s and t) People and merchandise (s) Stress and exercise (t) Video games/amount of play (s) 2 Stuff/want more stuff have/stuff burn-out (s)
Population/births	Concerts and concert goers (s) Antibiotics/antibodies (t) Teachers’ pay/demand for pay increase (s) Plant diversity/insect and bird diversity/garden health (t) Dot-commers and pyramid schemes (t)
Practice–enthusiasm	Report cards/homework (s) Professional development and enthusiasm (t) Competency and risk-taking (t) George Bush/public perception (s)
Room clean-up	People running for political office (s) Class room behavior/teacher attitude (t) Student government/friends approval (s) Beaches/trash/attractiveness of beach (s)

significant (see Table 10). Bivariate correlation analysis shows overall S-BI performance scores were significantly correlated ($r = 0.525$, $p = 0.004$) with students’ grade point average, but not with parents’ educational background. Given that the S-BI is not correlated with parents’ educational background, one might also conclude that the S-BI is not correlated with one’s advantages in life but with how well students do in school specifically, and with students’ ability to deal with complex novel problems in general. It is important to remember that these are preliminary analyses and are subject to reconsideration with further use of the instrument, larger samples, and greater variance in socioeconomic status.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Table 10. Mean performance on targeted measures group and school

Scores	FS Students	Comparison Students	FS Teachers	Comparison Teachers
System scenarios (max. score: 24)	9.39	9.76	14.32	12.71
Homology (max. score: 24)	6.21	6.64	14.33	12.20
Feedback pattern recognition (max. score: 6)	2.57	2.57	4.5	3.60
Policy thinking (max. score: 2)	0.78	0.78	1.5	1.2

Table 11. Percentage of student Level 3 responses at three prompts

S-BI question	Uncued response	1st level prompt	2nd level prompt
Predator–prey	14%	18%	46%
Teach–student	0%	11%	18%
Hunger–eating	3%	14%	29%
Population–birth	18%	46%	68%
Practice–perform	29%	32%	39%
Clean-up	25%	29%	36%

Students' and teachers' performance both improved with the assistance of interviewer questions. Table 11 shows how students' performance is helped in S-BI Part I by the interviewer's prompting questions.

Gains with prompts may simply indicate a demand effect; students and teachers learned what they were supposed to say by the third prompt, and were better at saying it. However, overall performance on S-BI Part I does not improve consistently. For example, those who did well on the second scenario do not necessarily do better on the fifth.

Students made the biggest gains in the predator–prey and population–birth scenarios (both reinforcing feedback scenarios). This may indicate that students had the most latent or "inert knowledge" (Perkins, 1992) in these two domains. It may also suggest that reinforcing feedback is more intuitively understood than balancing feedback. The 0–18% span for the teacher–student scenario may suggest that participants are generally not sensitized to the existence of a closed-loop dynamic between teacher expectations and student performance, that some participants held alternative mental models of student–teacher relationships, or that the scenario may be too abstract for this age group.

Overall S-BI performance scores revealed no significant gender differences among students or teachers. Teachers in the Food Systems school showed significantly higher systems intelligence scores than the teachers from the comparison school, and teachers as a group (across both schools) showed higher levels of systems intelligence than students.

Discussion

Why don't more participants naturally incorporate feedback processes into their reasoning? Through this research, we have shown, as Brazelton (1992), Grotzer (2003), Lackoff and Johnson (1980), and others have found, that people tend to focus on one-way causal structures when more complex "interaction patterns" exist. Without systems-specific content knowledge, individuals appear to default to descriptive, surface features. For example, participants who demonstrated low-level feedback recognition provided simple, descriptive responses such as "the student does what the teacher tells him" and "I play a lot of sports". Chi (2000) offers a theoretical explanation for why students misunderstand complex dynamic concepts such as natural selection: "students focus on an object's actions rather than its interactions, or they focus on the class of the object rather than seeing the object as a collection" (p. 19).

Another possible explanation for the lack of feedback recognition may be found in participants' use of terms such as "cycle" and "chain". The multiple and ambiguous meanings of "cycle" show it is a poor substitute for the term "feedback". People use the term "cycle" to refer to system structure (a feedback loop), as in "vicious cycle" and to refer to system behavior, as in a cycle in population abundance. The term "cycle" as used in school texts often means a closed loop of material or energy flow, as in the hydrological, carbon, or Krebs cycles, obscuring the notion of closed information feedback loops. The term "feedback" itself also carries multiple meanings, and in common usage means criticism or praise, as in "I received negative feedback from my boss." Such multiple meanings create difficulties when a more precise technical meaning is called for. The caution here is that various "cycle models" may persist in the minds of students and may be misapplied. For example, recognizing a cyclic behavior (e.g., hungry–eat–hungry–eat), a learner may stop his or her inquiry, and conclude simply that the "pattern repeats itself" rather than attending to the feedback structure. By stopping at the conclusion that a predator–prey relationship is a cycle, the student may not then consider the impact of accumulations or other feedbacks operating in the system (e.g., the positive feedbacks driving population growth for both species). Students may have used cycle-related phrases because they are more likely to encounter text and graphics related to "cycles", "food chains", and "chain reactions" (in school textbooks, for example) than feedback process-related text and visuals. In our review of middle-school and high school text books (informed by Barman and Mayer's (1994) textbook review) we found that energy transfer in ecosystems is most often represented in terms of one-way causal chains and food pyramids (Sussman, 2000). It is less frequently represented as food webs (Capra, 1999) or

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

causal loops. It is essential for educators to look closely at current curricular materials to identify existing topics that may act as building blocks for learning systems concepts (e.g., cycle to feedback loops), and to ensure that they are not counter-productive in developing systems thinking.

It is tempting to conclude that students' ability to recognize feedback and other system structures will develop through education and normal processes of maturation. Such a conclusion may be premature, however. Teachers may have outperformed students for several reasons unrelated to their understanding of systems or the formal education they have received. First, teachers were likely more comfortable with and more adept at the interview process. Even so, six students outperformed the teachers on various S-BI measures (including the overall S-BI score and time-related references such as time delays and time horizons). One explanation is that these students had experiences beyond the classroom that were relevant to the systems scenarios. For example, one of the six high-performing students had been exposed to the discourse of environmental activism though his parents.

Are students' intuitive models of complex systems simply less differentiated than adults' or do some students show a more intuitive, natural affinity towards systems than some of their teachers? Our results suggest that participants' systems intelligence may not develop in an ordered and organized sequence. Middle-school students, particularly when given structured prompts such as "what happens next?" demonstrated an understanding of interactions among objects in the Gestalt sense of seeing how one thing influences another and how those interrelationships hold together. These students provided explicit theories of how dynamics in a system work, through phrases like "they bounce off each other" or "they maintain a balance". Insight into this debate may be found in neo-developmental frameworks that apply the concept of dynamics to psychology, human development, and education such as Fischer's dynamic skill theory (Fischer and Rose, 1994). These researchers have shown that development is not as unified as Piaget (1954) had once suggested. For example, Case (1992) focuses on "central conceptual structures" such as narrative ability and numerical ability. These structures proceed not necessarily in sweeping waves but as areas of development that are loosely coupled with one another. If we think of systems intelligence as a "central conceptual structure" it is plausible that a student, with enough exposure to systems concepts and content could "outperform" teachers with less or no exposure.

At what age are students able to learn and apply systems thinking tools and skills? Efforts to bring system dynamics (Roberts, 1978; Mandinach and Cline, 1994; Quaden *et al.*, 2004; Fisher, 2005) and related system modeling approaches (e.g., Colella *et al.*, 2001) into K-12 education have shown that students as young as kindergarten can understand and utilize feedback and stock and flow structures. A synthesis of this study's findings with what is known about the development of higher-order reasoning skills (Fischer and Pruyne, 2003; Kitchener and King, 1990), along with the action research

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

conducted by system dynamics educators (with the support of the Waters Foundation), suggests that middle-school students can intuitively understand systems concepts. More research is needed to explore the degree to which younger children (those between the ages of 5 and 9) are able to grasp concepts as feedback, time delays, and accumulation.

To test the hypothesis that "children are natural systems thinkers", we developed and tested a protocol and tools to assess the intuitive systems-thinking skills of children and adults. We tested the System-Based Inquiry protocol with students and teachers in two middle schools, showing that it provides a robust and reliable method to elicit people's understanding of key elements of dynamic complexity, including awareness of dynamics, feedback processes, time horizons, and accumulations, and their ability to recognize the underlying feedback structure of a situation and apply it to other, analogous situations in different domains. Results show generally weak understanding of these concepts, but also show considerable variance in the abilities of both students and teachers, with some students outperforming some teachers. We raise concerns about the degree to which ordinary discourse, educational materials, and common teaching methods may encourage and support sloppy and incomplete thinking about complex systems. We hope this work will generate further research and greater collaboration between educators and members of the system dynamics community. We also hope that this research will inspire researchers to adapt the methods used here to further probe and clarify the questions raised in this research. Our common goal is to prepare students to grapple with the challenges of our times, and become, as Barry Richmond envisioned, "systems citizens".

Notes

1. The positive feedback of expectations and performance (the "Pygmalion effect") is well known (Rosenthal and Jacobson, 1968/1992) and is an example of the more general phenomenon of self-fulfilling prophecy (Merton, 1957).
2. There are, of course, many other feedbacks within the relationship between hunger and eating. As discussed in Booth Sweeney (2004), study participants described a wide variety of dynamics, e.g., "sadness eating", "eating at a party when you're not hungry", "the more you eat, the more you want to eat".
3. The drawings for this scenario were adapted from Roberts (1976).

Acknowledgements

We thank George Richardson, David Lane, Mitchell Resnick, Uri Wilensky for expert analysis of interview transcripts and Jeff Potsh. Tina Grotzer, Bob Kegan, David Perkins, Stay Blake-Beard, Kathy Rhoades, David Eddy-Spicer, Christine Kelly, Michael Murphy, Ann Harbison, and anonymous referees for their comments.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

References

American Association for the Advancement of Science. 1993. *Benchmarks for Scientific Literacy*. Oxford University Press: New York.

Bakken B. 1993. *Learning and Transfer of Understanding in Dynamic Decision Environments.* Massachusetts Institute of Technology: Cambridge, MA.

Bakken BE, Gould-Kreutzer JM, Kim DH. 1992. Management flight simulators and organizational learning: some experimental evidence. *European Journal of Operational Research* **59**: 167–182.

Barman CR, Mayer DA. 1994. An analysis of high school students' concepts and textbook presentations of food chains and food webs. *American Biology Teacher* **56**(3): 160–163.

Booth Sweeney L, Sterman J. 2000. Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review* **16**(4): 249–286.

Brazelton TB. 1992. *Touchpoints: Your Child's Emotional and Behavioral Development*. Perseus Books: Reading, MA.

Brown A, Campione J. 1994. Guided discovery in a community of learners. In *Classroom Lessons: Integrating Cognitive Theory and Classroom Practice*, McGilley K (ed.). MIT Press: Cambridge, MA; 229–270.

Bruner J. 1960/1977. *The Process of Education*. Harvard University Press: Cambridge, MA.

Capra F. 1999. Ecoliteracy: the challenge for education in the next century. Paper presented at the Liverpool Schumacher Lectures, Liverpool.

Case R. 1992. *The Mind's Staircase: Stages in the Development of Human Intelligence*. Erlbaum: Hillsdale, NJ.

Chandler M, Boutelier R. 1992. The development of dynamic system reasoning. *Contributions to Human Development* **21**: 121–137.

Chi MTH. 2000. Misunderstanding emergent processes as causal. Paper presented at the *American Educational Research Association*, Atlanta, GA.

Colella V, Klopfer E, Resnick M. 2001. *Adventures in Modeling: Exploring Complex, Dynamic Systems with StarLogo*. Teachers College Press: New York.

Cronin MA, Gonzalez C. 2007. Understanding the building blocks of dynamic systems. *System Dynamics Review* **23**(1): 1–17.

d'Apollonia ST, Charles ES, Boyd GM. 2002. Acquisition of complex systemic thinking and conceptual change: mental models. Paper presented at the *American Education Research Association*, Atlanta, GA.

Diehl E, Sterman J. 1995. Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes* **62**(2): 198–215.

Dörner D. 1980. On the difficulties people have in dealing with complexity. *Simulations and Games* **11**(1): 87–106.

Doyle J. 1997. The cognitive psychology of systems thinking. *System Dynamics Review* **13**(3): 253–265.

Doyle JK, Radzicki M, Trees S. 1996. Measuring the effect of systems thinking interventions on mental models. Paper presented at the *Proceedings of the 1996 International Conference of the System Dynamics Society*, Cambridge, MA.

——, ——, ——. 1998. Measuring changes in mental models of dynamic systems: an exploratory study. Paper presented at the *16th International Conference of the System Dynamics Society*, Quebec, Canada.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Espejo R. 1994. What is systemic thinking? *System Dynamics Review* **10**(2–3): 199–212.

Fischer K, Pruyne E. 2003. Reflective thinking in adulthood: development, variation and consolidation. In *Handbook of Adult Development*, Demick J, Andreoletti C (eds). Kluwer Academic/Plenum: Dordrecht; 169–197.

Fischer K, Rose S. 1994. Dynamic development of coordination of components in brain and behavior: a framework for theory and research. In *Human Behavior and the Developing Brain*, Dawson G, Fischer KW (eds). Guilford Press: New York; 3–66.

Fisher D. 2005. *Modeling Dynamic Systems: Lessons for a First Course*. Isee Systems: Hanover, NH.

Forrester J. 1993. Systems dynamics as an organizing framework for pre-college education. *System Dynamics Review* **9**(2): 183–194.

Green DW. 1997. Explaining and envisaging an ecological phenomenon. *British Journal of Psychology* **88**: 199–217.

Griffiths AK, Grant BA. 1985. High school students' understandings of food webs: identification of a learning hierarchy and related misconceptions. *Journal of Research in Science Teaching* **22**: 421–436.

Grotzer T. 1997. Children's understanding of complex causality in natural systems: summary excerpts from paper presented at the *Annual Meeting of the American Educational Research Association*, Project Zero, Harvard Graduate School of Education, Cambridge, MA.

——. 2003. Learning to understand forms of causality in scientifically accepted explanations. *Studies in Science Education* **39**: 1–74.

Grotzer T, Bell Basca B. 2000. Helping students to grasp the underlying causal structures when learning about ecosystems: how does it impact understanding? Paper presented at the *National Association for Research in Science Teaching Annual Conference*, Atlanta, GA.

Grotzer T, Sudbury M. 1998. Where is the language of causality? A survey of science and thinking skills resource books. *THINK Magazine*. December.

Grotzer TA, Bell Basca B. 2003. How does grasping the underlying causal structures of ecosystems impact students' understanding? *Journal of Biological Education* **38**(2): 16–30.

Hämäläinen RP, Saarinen E (eds). 2004. *Systems Intelligence: Discovering a Hidden Competence in Human Action and Organisational Life*. Espoo: Systems Analysis Laboratory, Research Reports A88, Helsinki University of Technology.

Hirsch GB. 1998. *Innovation in Schools: A Model to Help Structure the Discussion and Guide the Search for Strategies*. Report. Creative Learning Exchange: Acton, MA.

Hogan K. 2000. Assessing students' systems reasoning in ecology. *Journal of Biological Education* **35**(1): 22–28.

Jackson S, Stratford SJ, Krajcik JS, Soloway E. 1996. Making system dynamics modeling accessible to pre-college science students. *Interactive Learning Environments* **4**: 233–257.

Jensen E. 2003. *(Mis)understanding and Learning of Feedback Relations in a Simple Dynamic System*. Orebro University, Orebro.

Kates RW, Katz C. 1977. The hydrologic cycle and the wisdom of the child. *Geographical Review* **67**: 51–62.

Kitchener KS, King PM. 1990. The Reflective Judgment Model: Ten years of research. In

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

Adult development 3: Models and methods in the study of adolescent and adult thought, Commons ML, Armon C, Kohlberg L, Richards FA, Grotzer TA, Sinnott JD (eds). New York: Praeger.
Koplowitz H. 1984. A projection beyond Piaget's formal operation stage: a general systems stage and a unitary stage. In *Beyond Formal Operations*, Commons ML, Richards FA, Armon C (eds). Praeger: New York; 383–394.
Lackoff G, Johnson M. 1980. *Metaphors We Live By*. University of Chicago Press: Chicago, IL.
Mandinach EB, Cline HF. 1994. *Classroom Dynamics: Implementing a Technology-Based Learning Environment.* Erlbaum: Hillsdale, NJ.
Marek E. 1986. They misunderstand, but they'll pass. *Science Teacher* **12**: 32–35.
Meadows D. 1982. Whole earth models and systems. *Co-Evolution Quarterly* Summer: 98–108.
——. 1991. *The Global Citizen*. Island Press: Washington, DC.
——. 2001. Dancing with Systems. *Whole Earth* (Winter).
Merton R. 1957. *Social Theory and Social Structure*. The Free Press: Glencoe, Illinois.
Moxnes E. 1998. Not Only the Tragedy of the Commons: Misperceptions of Bioeconomics. *Management Science* **44**(9): 1234–1248.
Murphy MJ, Schweers E. 2003. *Evaluation of the Food Systems Approach to Fostering Ecological Literacy: Final Report*. Massachusetts General Hospital/Harvard Medical School: Boston, MA.
Nisbett RE, Fong GT, Lehman DR, Cheng PW. 1987. Teaching reasoning. *Science* **238**: 625–631.
Odum EP. 1977. The emergence of ecology as a new integrative discipline. *Science* **195**(4284): 1289–1293.
Ossimitz G. 1996. *The Development Of Systems Thinking Skills Using System Dynamics Modeling Tools*. Institute of Educational Research: Bonn, Germany.
——. 2001. Stock-flow thinking and reading stock-flow related graphs: an empirical investigation in dynamic thinking abilities. Paper presented at the *International System Dynamics Conference*, Atlanta, GA.
Pala Ö, Vennix JAM. 2005. Effect of system dynamics education on systems thinking inventory task performance. *System Dynamics Review* **21**(2): 147–172.
Perkins D. 1992. *Smart Schools: Better Thinking and Learning for Every Child*. Free Press: New York.
Perkins DN, Grotzer TA. 2000. April 24–28, 2000. *Models and Moves: Focusing on dimensions of causality complexity to achieve deeper scientific understanding*. Paper presented at the American Educational Research Association. In the Complex Causality and Conceptual Change Symposium, New Orleans, LA.
Piaget J. 1954. *The Construction of Reality in the Child*. MIT Press: Cambridge, MA.
Quaden R, Ticotsky A, Lyneis D. 2004. *The Shape of Change*. Creative Learning Exchange: Acton, MA.
Resnick M. 1995. *Turtles, Termites and Traffic Jams*. MIT Press: Cambridge, MA.
Resnick M, Wilensky U. 1998. Diving into complexity: developing probabilistic decentralized thinking through role-playing activities. *Journal of the Learning Sciences* **7**(2): 153–172.
Richmond B. 2000. *The "Thinking" in Systems Thinking: Seven Essential Skills*. Pegasus Communications: Waltham, MA.
Roberts N. 1976. *Picture Kits*. Lesley University: Cambridge, MA.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr

——. 1978. Teaching dynamic feedback systems thinking: an elementary view. *Management Science* **24**(8): 836–843.
Rosenthal R, Jacobson L. 1968/1992. *Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development*. Irvington Publishers: New York.
Senge P, Cambron-McCabe N, Lucas T, Smith B, Dutton J, Kleiner A. 2000. *Schools that Learn: A Fifth Discipline Fieldbook for Educators, Parents, and Everyone Who Cares About Education*. Doubleday: New York.
Sterman J. 1989. Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* **43**(3): 301–335.
——. 1994. Learning in and about complex systems. *System Dynamics Review* **10**(2–3): 291–230.
——. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGraw-Hill: Boston, MA.
Sterman J, Booth Sweeney L. 2002. Cloudy skies: assessing public understanding of global warming. *System Dynamics Review*, Special Edition **18**(2): 207–240.
——, ——. 2007. Understanding public complacency about climate change: adults' mental models of climate change violate conservation of matter. *Climatic Change* **80**(3–4): 213–238.
Sterman J, Meadows D. 1985. Strategem-2: a microcomputer simulation game of the Kondratiev cycle. *Simulation and Games* **16**(2): 174–202.
Sterman JD. 2002. All models are wrong: reflections on becoming a systems scientist. *System Dynamics Review* **18**: 501–531.
Summers MK, Summers JM. 1976. Analogue simulation in the teaching of population ecology. *Journal of Biological Education* **10**(5): 229–236.
Sussman A. 2000. *Dr. Art's Guide to Planet Earth: For Earthlings ages 12 to 120*. White River Junction, Vt.: Chelsea Green.
Sweeney LB. 2004. Thinking about Everyday Systems: An Empirical Investigation of Middle School Students' and their Teachers' Conceptions of Natural and Social Systems. Doctoral dissertation. Harvard University, Graduate School of Education.
Vennix J. 1990. *Mental Models and Computer Models: Design and Evaluation of a Computer-Based Learning Environment for Policy Making*. CIP-Gegevens Koninklijke Bibliotheek: The Hague.
——. 2000. *The President's Newsletter*. System Dynamics Society: Albany, NY.
von Bertalanffy L. 1968. *General System Theory: Foundations, Development, Applications*. Brazillier: New York.
Wagenaar WA, Timmers H. 1979. The pond-and-duckweed problem: three experiments on the misperception of exponential growth. *Acta Psychologica* **43**: 239–257.
Webb P, Boltt G. 1990. Food chain to food web: a natural progression? *Journal of Biological Education* **24**(3): 187–190.
White P. 1992. The anthropomorphic machine: causal order in nature and the world view of common sense. *British Journal of Psychology* **83**: 61–86.
White PA. 1997. Naive ecology: causal judgements about a simple ecosystem. *British Journal of Psychology* **88**(2): 219–233.
Wilensky U, Resnick M. 1999. Thinking in levels: a dynamic systems perspective to making sense of the world. *Journal of Science Education and Technology* **8**(1): 3–18.
Wylie J, Sheehy N, McGuinness C, Orchard G. 1998. Children's thinking about air pollution: a systems theory analysis. *Environmental Education Research* **4**(2): 117–137.

Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sdr