Developing and Field-testing a Rubric for Evaluating Students' Causal Loop Diagrams

Kim A. Kastens (kastens@ldeo.columbia.edu)
Lamont-Doherty Earth Observatory of Columbia University

Wayne Wakeland
Complex Systems Program, Portland State University

Thomas F. Shipley
Department of Psychology, Temple University

Submitted for 2024 International Systems Dynamics Conference, Teaching & Learning Theme
Revised: June 27, 2024

Abstract:

We have developed and field tested an educational rubric for instructors who are using causal loop diagrams (CLD) in undergraduate science and social science courses. Given a student's CLD and accompanying explanatory narrative, the rubric describes three levels of performance quality on the following dimensions: (1a) Nodes on the CLD, (2a) Links/arrows on the CLD, (2b) Narrative description of causal links, (2c) Connection of links to the real-world system, (3a) Diagram depiction of feedback loop, and (3b) Understanding of impact of this loop on the broader system within which it is embedded. The field test involved four geoscientists and 17 undergraduate psychology students, who watched an instructional video about how to make CLDs, read loop-containing narratives from the popular media, sketched a CLD, and wrote an accompanying explanatory narrative. Among the scientist-participants the weakest performance was on rubric element (3b). Among the undergraduate participants, notably poor performance was on rubric element (2c). Note that both these areas of poor performances lay in making the analogical mapping between modelled system and real-world system. In response to field test findings, we clarified and strengthened some elements of the rubric and accompanying instructional video.

Problem Statement:

Colleges and universities are filled with science and social science faculty who teach courses in which at least one important phenomenon is driven or constrained by a positive or negative feedback loop, and yet they do not seize the opportunity to foster broader feedback loop thinking in their students. One reason is that they feel they cannot spare the time to detour through an introduction to systems thinking on their way to their discipline-specific learning goals. Another is that they do not have in their instructional toolkit a way to accurately and transparently judge whether students are developing the ability to detect feedback loops in the content domain of the course and then explicate the system of causal influences that underlies the observed behavior. If college faculty ask undergraduates to produce drawings or diagrams as products for assessment, they may encounter pushback from students who say "This isn't fair; I'm not an artist," and they may experience self-doubt about their ability to score such visual representations in a way that is fair and defensible if challenged. We are tackling this problem by developing a rubric with which instructors and education researchers can score a student-produced causal loop diagram and accompanying explanatory narrative and thus gauge students' understanding of a discipline-relevant feedback loop.

Context & Learning Goals:

Our rubric development and testing effort is taking place in the context of a larger project called *Supporting Feedback Loop Learning in Natural and Social Science Courses,* funded by the U.S. National Science Foundation program for Improving Undergraduate STEM Education (IUSE). Earlier insights from this project can be found in last year's ISDC conference paper on *How the Human Mind thinks about Feedback Loops* (Kastens & Shipley, 2023). The target audience for the project are faculty members teaching an undergraduate course in natural or social sciences in which at least one phenomenon that is central to the course content is driven or constrained by a feedback loop. Our collaborating faculty teach courses in psychology, neuroscience, environmental science, and race & gender studies. In these fields, feedback loops are usually referred to as "positive" or "negative," and so we use those terms throughout this paper, rather than "reinforcing" and "balancing."

The IUSE feedback loop project is working to provide evidence-based tools and instructional strategies that will help faculty guide their students towards the following learning goals:

1) *Appreciate the breadth of feedback loops (FLs) in the world:* Look at the world with eyes open to the likely presence of positive and negative feedback loops as controllers or drivers of important processes.

2) *Detect FLs:* Be able to detect the behavior patterns that signal the possible presence of a feedback loop in a system within which one is living, or from a narrative description of a situation (Kenzie et al., 2024), or in a data set in one's discipline.

3) *Unravel and articulate causal structure:* Having detected the probable presence of a feedback loop, be able to construct, explain and illustrate a logically consistent and empirically supported chain of influences that form a closed loop (Senge, 2006, Ch 5) and that, upon repeated passages around the loop, tend to push the system farther away from its starting configuration (positive loop) or tend to pull the system back towards its starting configuration or towards an equilibrium value or goal state (negative loop).

4) *Situate the loop in a broader system:* Be able to explain how the existence and activation of the feedback loop impacts the larger system within which the FL is embedded.

5) *Solve problems:* Use FL mental models and FL representations to formulate and critique potential solutions to problems that are underlain by a feedback loop.

These five learning goals comprise what we are calling "feedback loop thinking." "Feedback loop thinking" as used here can be thought of as subset of the system dynamicist's "endogenous perspective" or "endogenous point of view" (e.g. Richardson, 2011; 2020).

The rubric project aims to develop tools to help faculty help students advance towards learning goals 2, 3 and 4, and to evaluate to what extent they have done so. This project explores and leverages the systems thinking capabilities of the human mind, unaided by computational modelling.

Methodology and Rationale:

Decision to develop a rubric: A rubric is an education and education research tool that articulates the criteria for a student performance or product along with gradations of quality for each criterion (Andrade, 2001). Typically, the criteria are laid out in the rows of a matrix while the gradations of quality form the columns; that is the format we have used in this paper. The existence of the quality levels distinguishes a rubric from a checklist. The high-performance level communicates the desirable qualities in student work, while the low-performance level can point out common pitfalls or failure modes. The intermediate level(s) describe student work that includes a messy mixture of both strong and weak aspects--which in many instructional contexts is most students as they move haltingly towards mastery of a challenging new skill or understanding. Education scholars assert that use of rubrics enables instructors to set -- and students to meet -- more ambitous learning goals (Andrade, 2005; Reddy and Andrade, 2010; Stevens & Levi, 2012).

Earlier CLD rubrics: We know of two prior efforts to develop and disseminate a rubric for causal loop diagrams. iRubric is a web-based assessment system in which educators can create rubrics for any type of performance or assignment, share the rubric publicly, integrate the rubrics into their campus' Learning Management System, score student work, and generate aggregated reports on student learning across a program or institution. Among its gallery of shared rubrics, iRubric includes a rubric for a causal loop diagram (stardiverdwr, n.d.), which includes dimensions for Variables, Causal Relationships, and Analysis. However, this rubric is tied to a specific assignment, and lacks the generalizability that we are trying to achieve. A group of K-12 educators, working as the DynamiQUEST 2000 Committee, developed a suite of rubrics for assessing the understanding generated by the use of system dynamic tools (Stuntz, et al, 2000). Among these is a Rubric for Causal Loop Diagrams (Lyneis, 2001), with dimensions for Variables (The words), The Arrows (Causal connections between variables), and Usefulness. Although designed for K-12 use, the types of mastery called for in the DynamiQUEST rubrics are also highly applicable for undergraduate instruction. The DynamicQUEST rubrics include only a single performance level, called "Standard," in a checklist format. The checklist format cannot capture or convey insights about emerging understanding or partial mastery.

Decision to require a causal loop diagram: Kastens and Shipley have a long history of researching the power of visualizations in science education, as a means of conveying information about the world, as a professional practice of science, and as a research tool for revealing the mental models of science students (Kastens, et al, 2016; Resnick, et al., 2018). So there was never any doubt that our feedback loop instructional sequence would feature instructor and student use of visual representations. Although system dynamicists generally favor stock-and-flow diagrams, we decided to require students to construct causal loop diagrams (CLD) (Sterman 2000 Ch 5, Ford 2010 Ch.9), for the following reasons: When science and social science textbooks and instructors do present a feedback loop as an explanation for a discipline-relevant phenomenon, they usually use a single, simple loop, much closer to a CLD than a stock-and-flow diagram. In part because of prior familiarity from textbook examples, CLDs take less time to introduce and require less practice to get to the stage where insights about the represented system can be generated. Finally, the visual presentation of a well-crafted CLD foregrounds the closed-circuit character of the chain of influences.

Decision to require an accompanying narrative: From a CLD alone, an instructor can judge whether the student has mastered the mechanics of making a CLD. But it is often difficult to gauge how well the student understands the real-world system of which the CLD is a representation, especially when the student's understanding is incomplete or emergent. An ideal instructional sequence would be for the student to draw their CLD, and then talk their way around the diagram explaining to a knowledgeable instructor their reasoning for including each link in the causal chain, and then step back and reflect on the impact of the loop in its entirety on the larger system within which the loop is embedded, all the while being gently questioned and encouraged by said knowledgeable instructor. Such a Socratic teaching methodology is impractical for whole-class instruction in most university settings. We decided to require a narrative to accompany the CLD to capture some of the information that would have emerged in an oral walk-around. The rubric draws on both the CLD and accompanying narrative.

Process of constructing the rubric. The rubric was constructed by iterative discussion among the three authors, over the course of seven bi-weekly conversations over zoom. The three co-constructors comprised a geoscience education researcher who has experience developing coding schemes for student products and teaching systems concepts to undergraduate Earth & environmental science students (Kastens), a cognitive scientist who researches students' mental models and has experience teaching systems concepts to undergraduate psychology students (Shipley), and an experienced systems dynamicist who has taught systems dynamics at all levels from novice to advanced (Wakeland). We began with the simplest and most reductionist dimensions and worked our way out to the most integrative dimensions. The decision to require an accompanying narrative came about one third of the way into the process. Within each dimension, we began by articulating the "meets expectations" performance level and then the "unacceptable level." Decisions were reached by consensus. Many iterations were required, as work on the higher-order dimensions often required modifications on the lower-order dimensions to avoid gaps and overlaps and maintain internal consistency. Such an iterative process is the norm for rubric development for a new educational construct.

Decision to prompt the loop construction with a reading passage: When a rubric is used in education or education research – especially a rubric that sets a high bar for student performance – it is essential that the prior instruction and the prompt for the student performance task be well matched with the rubric and the learning goals that informed the rubric. Alignment of learning goals, rubric, instruction and task prompt sets students up for success. We elected to prompt students' loop construction by having them read a passage from the popular media that describes a situation that we understand to be underlain by a feedback loop, but where the journalist or writer has not used the explicit term "loop" or "feedback loop." This task aligns with our learning goal #2: "Detect FLs: Be able to detect the behavior patterns that signal the possible presence of a feedback loop in a system within which one is living, **or from a narrative description of a situation,** or in a data set in one's discipline." The reading passage approach can be easily adopted to any disciplinary domain, takes minimum instructional time, and can be assigned as homework. All students are diagramming and narrating the same phenomenon, thus making student-to-student comparison easier for the instructor and more transparently fair in the eyes of the student. Unlike textbook accounts, journalistic accounts of loop phenomenon typically contain some ambiguity, or gaps or confounding factors, so the task is challenging – and yet students *can* succeed. Finally, detecting loops in newspapers or websites or similar writings for popular audience is a life skill that will serve student well in post-college civic engagement and personal and professional life.

The instructional video: Prior to doing the loop drawing task, participants in this project viewed a short video on *How to make excellent casual loop diagrams,* developed and presented by author Kastens. The video emphasizes elements that are valued in the rubric: put entities in your nodes that can increase or decrease; put a symbol in the middle of your loop to indicate that you have modelled a positive or negative feedback loop system; talk your way around the loop (Sterman 2000 Ch. 5) scrutinizing the empirical evidence or mechanism that justifies each link in the causal chain, and so on. The video was revised and re-recorded after the first field test (see below) to further emphasize some rubric elements on which some of the first field test participants had done poorly. The power point slides from the version 2 video are included in the supplementary documents; feel free to use or modify them, giving appropriate credit. The slides and video use "positive" and "negative" loop terminology because that is the language used in the science departments of our collaborating faculty. The slides and video use "S" (same) and "O" (opposite) to label links to avoid confusion with the + and - symbols used for the loop as a whole.

Results: The Rubric

The rubric has six dimensions and three performance levels. The full rubric is included in the Appendix, and Table 1 shows the "Meets Expectations" level for each of the dimensions.

Table 1

		Meets Expectations
Nodes	(1a) Nodes on the CLD	• All nodes that are part of a loop depict something that can potentially increase or decrease, such as a quantity (e.g. # of people), state (e.g. anger, temperature), or attribute (e.g. strength), *and* • Nodes do not state or imply the direction of change.
Causal Relationships/Connections	(2a) Links/arrows on the CLD	• All arrows have either a +/- or an S/O *and* • The polarity of every link matches the narrative *and* • The number of links and nodes is sufficient to convey all of the relationships that are important for the intended use of the CLD.
	(2b) Narrative description of causal links	The narrative progresses methodically around the loop, describing how a change in each node results in a change in the next node downstream. The narrative may start at any point in the loop, but must unambiguously return to the starting node. Explicitly causal language is used for at least some links.
	(2c) Connection of depicted links to real-world system	For every link in the CLD, the narrative expresses a causal relationship that does exist in the real-world system or that could plausibly exist in a hypothesized system, as judged by a coder/ instructor with knowledge of the system.
Net Effect of Loop as a Whole	(3a) Diagram depiction of feedback loop	• Diagram includes text nodes and arrows which form a closed loop; arrows all point in the same direction (CW or CCW around the loop) *and* • There is a R/B or +/- label for the net outcome of the loop in the center of the diagram *and* • The type of loop indicated in the center of the diagram matches the cumulative effect of stepping around the loop.
	(3b) Understanding of impact of this loop on the broader system within which it is embedded.	• The narrative clearly describes the net effect or impact of the loop taken in its entirety. • The impact of the loop on a broader system within which the loop is embedded is explicitly discussed, and is an accurate representation as judged by a coder/instructor with knowledge of the real-world system.

Building up from dimension 1a to 3b, the rubric probes for a more and more integrative understanding of the system under consideration: Dimension 1a starts with nodes. Dimensions 2a, 2b, and 2c deal with relations between pairs of nodes, i.e. with links. And dimensions 3a and 3b deal with the loop in its entirety, as embedded in a larger system.

Dimensions (1a), (2a), (2b), and (3a) deal with the mechanics of following best practices in assembling the elements of a CLD (Sterman, n.d) and writing a narrative that progresses methodically around the loop. These dimensions can be coded/scored by a person with knowledge of systems thinking, but minimal knowledge of the real-world system being depicted.

Dimensions (2c) and (3b) require the student to combine their knowledge of systems thinking with their knowledge of the workings of the system being represented. Coding these dimensions requires the coder to have substantial content knowledge in the domain of the represented system. This is realistic for the intended context of disciplinary science or social science courses. If the depicted system is a real-world system of the present or the past, the coder is asked whether the student's analogic mapping between the real world and the representation is accurate, as understood by the relevant disciplinary community. If the depicted system is a hypothesis about how things might work under differing circumstances (for example, in the future), the coder is asked whether the system of causal relationships mapped by the student is plausible.

The "Meets Expectations" performance level of the rubric describes a student product that exhibits best practices for making CLDs coupled with a narrative that exhibits understanding of the workings of the loop in the real world and how the loop impacts the larger system. "Best practices" for making CLDs are those that have been found by systems thinking practitioners to support insightful and flexible thinking about the real-world system. For example, the rubric rewards CLDs in which the loop nodes all depict something that can potentially increase or decrease but that do not state or imply the direction of change (dimension 1a). In a depiction of the climate ice-albedo positive feedback loop, the rubric would look favorably on "air temperature" as a node and would frown upon "air temperature rises." By specifying the direction of change, "air temperature rises" unnecessarily limits the explanatory power of the model. "Air temperature rises" can depict the current global warming scenario of rising air temperature and shrinking Arctic ice coverage. But a more flexible and thus more powerful model would also accommodate a scenario of cooling atmosphere and expanding ice coverage, as characterized the Snowball Earth episodes of Earth history when ice spread almost all the way to the equator (e.g. Schirber, 2015). By encouraging students to use these CLD best practices, faculty can position students to think more insightfully about the real-world systems.

In evaluating the narratives, the Meeting Expectations performance level looks for a methodical and thorough explication of the workings of the loop, as well as some insight about the role, function, or impact of the loop within a larger system (rubric level 3b). It's not enough to make a generalized statement that this is a positive feedback loop. A Meets Expectation performance must also say something like: "Because of the existence of this positive feedback loop, the global air temperature has been rising faster than it would otherwise have done," or "The net effect of this feedback loop is to keep the organism's body temperature within the range that supports life," or "A beneficial outcome of this feedback loop is that neither the predator species nor the prey species comes to overdominate the ecosystem," or "In societies where this positive feedback loop is active, the effect is that over time the rich get richer and the poor get poorer."

The "Unacceptable" performance level notes representational error types that we have found to occur in student products, for example using lines instead of arrows, or arrows with two heads, or arrows that merge or bifurcate. This performance level also calls out symptoms of lack of understanding of the real-world system, at both the node-link-node level (dimension 2c) and at the whole-loop level (dimension 3b). The current version of the rubric also gives an Unacceptable score on dimension (2b) for a narrative that does not "close the loop" (return to the starting point). Our team debated this point: If the narrative spells out clearly how a change in A influences B and how B influences C and then simply fails to mention how C influences A, how bad a failure is that? Could it be mere carelessness? In the end, we decided that such a narrative misses the essential aspect of a feedback loop – which is the closed nature of the circuit of influences – and we structured the rubric to push strongly on students to attend closely to the loop-closing step in their logical exposition of the workings of the loop.

The "Room for Improvement" performance level describes a student product with mixed or uneven quality indicators. A CLD which shows nodes and arrows in a closed loop, but lacks a +/- or R/B indicator in the center of the diagram would score at this level on dimension (3a). A narrative that mentions some but not all of the relationships in the diagram and where the logic is hard to follow would score at this level on dimension (2b).

In addition to the qualitative descriptions of performances levels, the rubric also offers quantitative levels, ranging from 1 for Unacceptable, through 3 for Room for Improvement, to 5 for Meets Expectations. In an education research context, the quantitative scores also allow the coder to give intermediate scores of 2 or 4 for a student product in between the described levels. In an education context, a score of zero would be reserved for a student who didn't do the assignment at all. For a few dimensions, there is also an Exceeds Expectation score of 6 or 7 available for indications of exceptional insight: for example, a narrative that indicates that a loop can be driven in either direction depending on the nature of the initial nudge.

Results: Field Test #1

Our first field test of the rubric involved four geoscientists at the level of PhD candidate or university research staff, who volunteered to participate in a systems thinking curriculum development project. As geoscientists, all engage with systems thinking in their work, but they vary in their degree of experience with formal or computational systems modelling. This group of participants was expected to produce CLDs and narratives that would fall towards the higher-performance end of the rubric.

The geoscientist-participants worked with two reading passages: one close to their general area of expertise, dealing with air conditioners and global warming, and one far from their area of expertise, dealing with exercise, muscle mass and aging. Both passages depict positive feedback loops.

The geoscientist-participants completed the activity in their own office or in the office of one of the authors. They watched version 1 of the *How to Make Excellent Causal Loop Diagrams* video; then read the first reading passage, constructed a CLD and wrote an accompanying narrative; then read a second reading passage and constructed a CLD and narrative for the second passage. The prompt for one task is shown in figure 1. The order of presentation of the readings was counterbalanced; half saw air conditioners first and half saw muscle mass first. The CLD and narrative were both done with pen or pencil on paper. The entire sequence took between half an hour and an hour.

Two authors independently scored all eight diagram/narrative products. We then methodically compared, discussed, and sometimes adjusted one or both scores for each rubric element for each participant, until we reached consensus or near consensus. "Near consensus" was operationalized as within one point out of five for a rubric element. To reach near consensus, we sometimes clarified or amplified the wording of the rubric. The rubric presented in Table 1 and in the Appendix is the consensus version after this interrater-reliability process was completed. One scorer tended to score participants more generously than the other scorer. However, the order of scores from best to worst was nearly identical for the two scorers, bolstering our confidence that the rubric would be fair and informative if used by a course instructor or used in a research context where the scoring was done by the same personnel throughout.

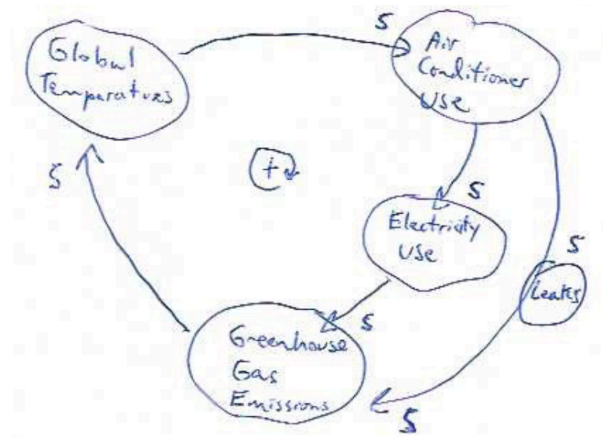
Figure 1. The prompt used for the air-conditioning task. The muscle mass prompt is similar.
Attached is an excerpt from a recent article written for a popular audience: *Air-conditioning use will surge in a warming world,* from The New York Times, December 5, 2023, by Hiroko Tabuchi. This article includes a description of a real-world phenomenon that can be explained in terms of a feedback loop. Drawing on your knowledge of environmental systems, the article itself, and your systems thinking skills: 1. Mark the place in the article that you think conveys a feedback loop by highlighting or underlining. 2. Sketch a causal loop diagram (CLD) that depicts the feedback loop, using best practices for CLD construction as shown in the video. 3. Write a narrative that describes how the loop works. Your narrative should start at one node and progress all the way around the loop, and should discuss the net effect of the loop as a whole.

Across all participants and both readings, the performance level was better on the more reductive dimensions of the rubric and worse on the more integrative dimensions, ranging from high average score of 4.6/5 on rubric dimension 1a (Nodes) to a low average score of 2.8/5 on dimension 3b (Understanding of impact...).

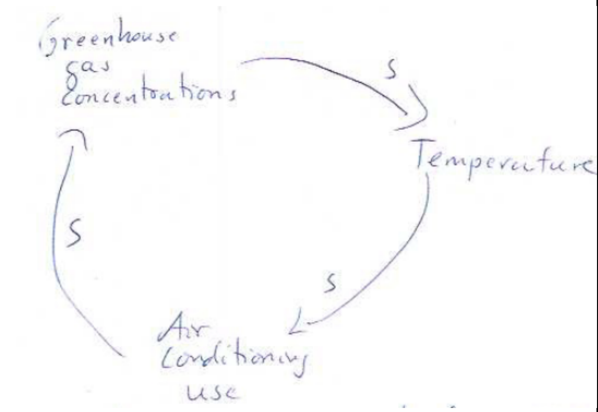
Performance on the air-conditioner task was stronger overall than on the muscle mass task. This may have been because the reading was more straightforward, or because the domain was closer to the participants' own expertise, or likely both. Key portions of the air conditioning reading are shown in figure 2.

Figure 2: Key phrases from the air-conditioning reading. The entire reading passage comprised seven short paragraphs, or 327 words, extracted from a longer article (Tabuchi, 2024). The phrases shown here were underlined or highlighted by at least one participant in response to the prompt to "mark the places in the article that you think conveys a feedback loop."
... As global temperatures, rise, more people will turn to air-conditioners to ward off the heat. ... But additional air-conditioning in buildings and other spaces, which is also driven by rising incomes, population growth and urbanization, means that the world could use more than double the electricity it does now to stay cool, leading to more planet-warming emissions, according to ... ... The surge in electricity use in turn threatens to drive up the very greenhouse gas emissions that cause global warming, heating the planet to even more dangerous extremes. Special refrigerant gases used in air-conditioners and refrigerators, when leaked into the atmosphere, are also potent greenhouse gases...

Figure 3 shows a stronger and a weaker performance on the air-conditioning task.



This is a positive feedback loop. As global temperatures increase, people will tend to use air conditioning more. This increases electricity use, which in turn leads to greenhouse gas emissions, which raise global temperatures, accelerating the cycle. Increased use of air conditioners also directly increases emissions by increasing the leakage of greenhouse gases used in the air conditioners, which also enhances the cycle.



As greenhouse gas concentrations rise the temperature also rises. Warmer temperatures lead to more use of air conditioning, which emits more greenhouse gases. This is a positive feedback loop.

Figure 3: Two participants' responses to the air-conditioning task. The narrative below each loop has been transcribed from the participant's hand-written version.

Both figure 3 participants constructed loops with nodes and links that correspond to relationships conveyed in the reading passage. Both used arrows and nodes appropriately, and correctly depicted the relationships as "S" or "Same" links. Notably, the right-hand loop of figure 3 lacks the "+" symbol in the center of the loop, thus coding lower on rubric dimension 3a.

The left-hand narrative conveys more information about the real-world mechanisms by which node influences node, thus scoring higher on rubric dimension 2c. In support of the link from air temperature to air conditioning use, the left-hand narrative introduces the decision-making agent: "people will tend to use air conditioning more." In support of the causal chain from air-conditioner use to greenhouse gas emissions, the left-hand response inserts an additional node, "electricity use," in both the diagram and narrative. In contrast, the right-hand narrative just leaps from "more use of air conditioning" to "emits more greenhouse gases" with no clue as to mechanism.

The current rubric does not do a very good job of capturing the important distinction that the left-hand diagram and narrative include an entire extra chain of influence that is missing from the right-hand diagram: the pathway from air conditioner use to greenhouse gas emissions via leaks. Rubric dimension 2a calls for a "number of links and nodes ... sufficient to convey all the relationships that are important for the intended use of the CLD," and the left-hand participant did score higher than the right-hand participant on that dimension. However, that "sufficient to convey" wording is vague and will be hard for coders or instructors to apply consistently. And the narrative dimensions of the rubric do not give any credit to the additional causal pathway in the left-hand response.

Both narratives state that this is a positive feedback loop, which is one way of complying with the prompt to "discuss the net effect of the loop as a whole." However, these answers fall short of meeting our project learning goal that learners should "Be able to explain how the existence and activation of the feedback loop impacts the larger system within which the FL is embedded." One more sentence could have provided strong evidence of meeting this learning goal, something along the lines of "because of this positive feedback loop, air temperatures around the globe are rising a bit faster than they would have risen in the absence of this loop." The left-hand narrative hints at this understanding, by use of the words "accelerating" and "enhances."

Figure 4 shows portions of the muscle mass reading. In response to the prompt to "Find the loop description and mark it by underlining or highlighting," the most commonly marked phrase (3 of 4 participants) was "people are less active because they are weaker, and they are weaker because they are less active." This phrase evokes the reciprocal nature of a positive feedback loop.[*] The individual who did not attend to this phrase scored most poorly on the rubric, with weakness in both diagram and narrative. This association reminds us that sorting out what to attend to amid the firehose of information available in considering a real-world phenomenon is an important subskill within our learning goal (2): Detect FLs.

Notably, no participant marked the sentence: "To paraphrase Hemingway, this process begins in two ways: gradually and then suddenly." Hemingway's oft-quoted phrase comes from *The Sun Also Rises,* where the character Mike uses it to describe going bankrupt. "[Something bad] happens gradually and then suddenly" is a vivid verbal description of the behavior over time characteristic of a positive feedback loop with undesirable outcome. That none of our participants called out this phrase suggests that learners could benefit from instruction and practice in spotting linguistic clues that may signal the presence of a feedback loop such as "[things got worse] gradually and then suddenly." Other such linguistic clues include "vicious cycle," "virtuous cycle," "downward spiral," and "chicken and egg situation" (Kastens & Shipley, 2021; Kenzie, et al., 2024).

---

[*] Interestingly, the author of the book from which the reading passage was taken may not have fully understood this phenomenon as a feedback loop. In its entirety, the paragraph reads: "By age eighty, the average person will have lost eight kilograms of muscle or about eighteen pounds, from their peak. But people who maintain higher activity levels lose much less muscle, more like three to four kilograms on average. *While it's not clear which way the causation flows here, I suspect it's probably both ways*: people are less active because they are weaker, and they are weaker because they are less active" [emphasis added]. In medicine, where randomized controlled trials are the "gold standard" for researching causality, the default assumption tends to be that causality can flow in only one direction, and that rigorous experimental design and large enough participant pools will enable the researcher to sort out which direction. The author is departing from the norms of his discipline by confessing his suspicion that in this particular situation, causation might flow in both directions.

Figure 4: Key phrases from the muscle-mass reading. The entire reading passage comprised four paragraphs, or 354 words, extracted from a book (Attia & Gifford, 2023).

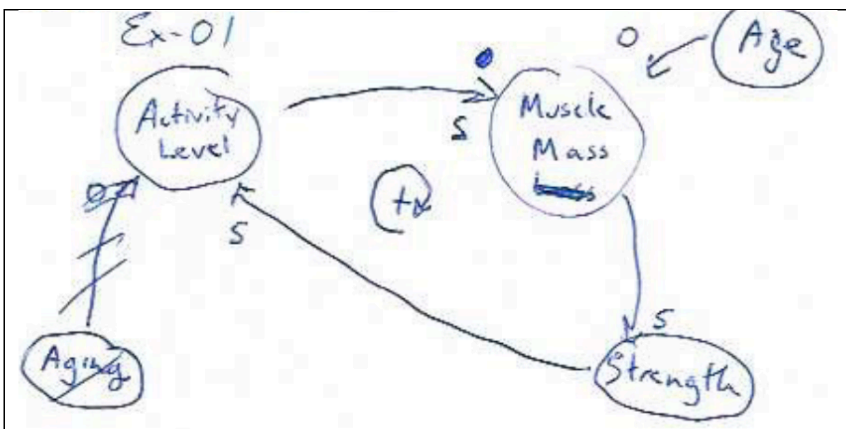
These phrases were underlined or highlighted by at least one participant in response to the prompt to "mark the places in the article that you think conveys a feedback loop."

> ...hallmarks of aging is that our physical capacity erodes. ... We lose strength and muscle mass with each passing decade, our bones grow fragile and our joints stiffen...
>
> ... Longitudinal and cross-sectional studies find that [muscle mass] and activity levels remain relatively consistent as people age from their twenties and thirties into middle age. But both physical activity levels *and* muscle mass decline steeply after about age sixty-five, and then even more steeply after about seventy-five. It's as if people...
>
> ...By age eighty, the average person will have lost eight kilograms of muscle or about eighteen pounds, from their peak. But people who maintain higher activity levels lose much less muscle, more like three to four kilograms on average. While it's not clear which way the causation flows here, I suspect it's probably both ways: people are less active because they are weaker, and they are weaker because they are less active.

The following phrases were not marked by any participant:

> "To paraphrase Hemingway, this process begins in two ways: gradually and then suddenly." ["this process" refers to a cluster of phenomena described in the previous paragraph: "hallmarks of aging ... physical capacity erodes... lose strength and mass... bones grow fragile ... joints stiffen."]
>
> "Continued muscle loss and inactivity literally puts our lives at risk. Seniors with the least muscle mass... are at the greatest risk of dying from all causes. One Chilean study..."

Figure 5 shows a strong response to the muscle mass reading task, illustrating several interesting features. This participant discerned that although "Age" was central to the reading passage, it was not a part of the loop. Instead, they modeled "Age" as an exogenous influence connected to muscle mass by an "O" link, expressed in the narrative as "when people age, their muscle mass declines." In their narrative, "Age" serves as the nudge that gets the loop going: "When people age, their muscle mass declines, which negatively impacts strength, activity level, and muscle mass, a positive feedback." None of the other eight responses to either reading were so explicit, in either words or diagram, about the nature of the nudge that starts or accelerates a feedback loop. This facet of feedback loop thinking was not emphasized in the instructional video, and may be a leverage point for improved instruction.



Narrative: The feedback loop represents a positive feedback. As activity level increases, muscles become stronger. Having more strength enables people to be more active. When people age, their muscle mass declines, which negatively impacts strength, activity level, and muscle mass, a positive feedback.

Figure 5: CLD and accompanying narrative created by a participant asked to find a feedback loop in a reading passage from a book on aging, diagram the found loop, and write an accompanying narrative. This is a strong performance, which scored high on our rubric.

As shown by the crossed-out node and link, the figure 5 participant initially envisioned "Aging" as an exogenous influence on "Activity Level." In their final answer, "aging" did not appear in either the diagram or narrative, although we think the instinct was correct that "aging" should be included somehow. In our interpretation, the constellation of phenomena that the English language calls "aging" is best viewed as an emergent outcome of the action of the figure 5 feedback loop (along with other downward spiral type feedback loops impacting other components of human physiology). Thus "aging" could have been part of a strong answer to rubric dimension 3b (Understanding of impact on broader system). Considering the human body as the "larger system," a strong answer might say "the net effect of this feedback loop is that the body's physical capacity declines more rapidly than would happen were the loop not active, accelerating the effects of aging and risk of death." As with the air-conditioner task, none of the four scientist-participants did a good job of connecting the structure of the loop with its impact on the larger system. The diagram of figure 5 is sufficiently robust that it could also have been used to narrate a net outcome in the stronger-stronger direction: "If a person begins to work out or otherwise purposefully increase their activity level, that will increase their muscle mass, which will in turn increase their strength. As their strength grows, they will be capable of sustaining higher activity level, thus completing a positive feedback loop. The overall action and outcome of this loop are called 'conditioning'."

One participant diagrammed and described a negative loop rather than a positive loop, introducing an "O" link from "Stress put on bones and joints" to "Muscle mass." Stress on bones and joints was not mentioned at all in the reading passage, and the evidence or mechanism for this claimed link was not elucidated by the participant. The rubric is currently silent on what the coder should do if a response features phenomena that were not contained in the reading passage. This same participant seems to have fallen into the trap of confounding negative structure with undesirable outcome, writing "Net effect: this loop is always negative, in that mobility always decreases to some extent over time, but the extent to which it's negative seems to differ from person to person..."

After reviewing the products from the first field test, we took the following steps:

- We revised and re-recorded the instructional video. The new version places both verbal and visual emphasis on the importance of putting a + or - symbol in the center of the diagram. In the old video, the process of walking and talking one's way around the loop was presented primarily as a way to check whether the diagrammed loop was positive or negative, whereas in the new video this talking around the loop process is presented as a way to understand the system as a whole and how the pieces fit together. The new video also explicitly shows, in both written and spoken word, the process of articulating the impact of the loop on the larger system within which it is embedded. The powerpoint slides in this paper's supplementary materials correspond to the new video.
- We revised the rubric, especially dimension 3b. A narrative that has a generalized statement of net effect that could apply to any + or - loop was downgraded to performance level 3, Room for Improvement. To achieve level 5, Meets Expectations, the rubric now requires that the narrative explicitly discusses how the action of the loop impacts a larger system within which the loop is embedded.

With respect to participants' difficulty with rubric element 3b, we acknowledge that we could have been more emphatic on this point in the instructional video and in the task prompt – and we will do so in the future. However, we also hypothesize that this difficulty may be rooted in the nature of the human mind. Remember that trick picture, which can appear to be either a vase or two profiles facing each other. Your percept can go from vase to profile and back to vase, but your mind is not capable of seeing both at the same time; you never see two people kissing a vase. In a somewhat similar way, your mind cannot think about parts and whole at the same time. You can go back and forth from parts to whole. With practice, you can get better at going back and forth – but it remains a cognitive challenge. And this cognitive challenge is exactly what we have asked participants to do in this task. The reading passage presents an aspect of reality as an integrated whole. We ask participants to break up this whole into parts (nodes and links), and reassemble the parts into a different kind of whole: a feedback loop. They invest considerable time and cognitive effort into this disassembly and reassembly process. But then we suddenly ask them to make an about face, and take this whole, this loop, which they have been painstakingly assembling, and stop thinking of it as a *whole* and start thinking of it as a *part* of a larger system. Of course, it is possible to learn to do this. However, it may be a harder cognitive task than we had fully appreciated.

Results: Field Test #2

The second field test involved 17 undergraduate seniors, psychology majors enrolled in a capstone course on "Psychology of food" taught by author Shipley. These participants were expected to fall lower on the rubric performance scale than the scientist-participants in Field Test #1.

The instructor introduced positive feedback loops in a brief lecture, and the students viewed version 2 of the CLD video as homework. Negative feedback loops were not covered in this class. Over the course of one 90-minute class session, the students read and analyzed three reading passages. The instructor led the first analysis as a whole class activity, collecting suggestions from students, gradually building the loop on a white board, and demonstrating the process of talking one's way methodically around the loop link by link. That first reading (from Mason, 2022) touched on abundance of stores selling affordable fresh nutritious food, rates of obesity and diabetes, ability of people to work and generate wealth, and implicitly the ability of the neighborhood to support local shops. The reading referred to the situation as a "downward spiral" and called the outcome a "food desert." For the second reading, students worked in small groups. This assignment included three short descriptions of feedback loops, all associated with the percent of a population who are vegan; these readings came from a reddit thread entitled "Veganism is in the Early Stages of a Positive Feedback Loop." As in Field Test #1, the student groups were prompted to create a causal loop diagram to depict the feedback loop they found in each reading, and to write a narrative describing how the loop works. The third reading returned to the food desert theme, and described how the food desert downward spiral could be reversed by introducing small scale shops, such as a bakery, to generate an outcome referred to as a "food oasis" (Mason, 2022). For this third analysis, students worked individually, following the same prompt as for the small-group activity.

Figure 6 shows the reading passage used for the individual activity.

Figure 6: The reading passage, in its entirety, used for the individual activity in the Psychology of Food course. Extracted from Mason (2022).

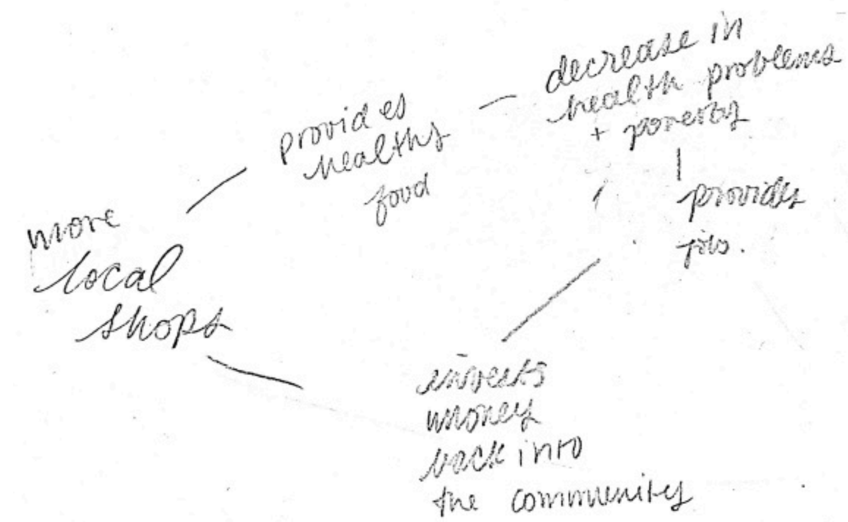
Sowing seeds of hope in food deserts: Virtuous Bread and Bread Angels founder Jane Mason considers the role microbakers can play in feeding and strengthening a community.

I am a baker who set up Bread Angels, one of the largest networks of microbakers in the world. We bake, teach people how to bake, and teach people how to set up microbakeries. By baking and selling Real Bread to people in our local communities, and by teaching others, microbakers provide valuable products and services, generate an income for ourselves, and can potentially employ others.

The presence of a microbakery might also help to inspire and enable other small locally-owned businesses to open and thrive. In a food desert, a microbakery can nurture the first shoots of a 'food oasis'. Having had a fantastic experience buying and consuming Real Bread from their local microbaker, many customers will now never buy a loaf anywhere else. Some people have expanded the range of what they buy from local, small producers, which in turn has encouraged more to set up shop. These changes are not limited to middle class neighbourhoods. With support, anyone can learn to bake and set up a business that serves their local community with fresh, handmade, Real Bread.

Figure 7 shows an example of a student response to the individual activity. This response shows considerable systems thinking insight about the system depicted in the reading. Each node in the diagram relates in a logical way to its adjacent nodes both upstream and downstream, and the narrative progresses methodically around the loop explaining each link in turn. Narrative and CLD both identify the broader system on which the loop is impacting ("the community.") Unlike many of their classmates, this student realized that what is increasing as result of the Bread Angels' intervention is abundance of "local shops" (rather than abundance of microbakeries.) And yet the response scores relatively poorly on our rubric because of weakness in the mechanics of CLD construction: the diagram has lines instead of arrows, has no plus or minus sign in its center, and some nodes indicate the direction of change (e.g. "more local shops.")

Our interpretation of figure 7 suggests that for use by science and social science disciplinary faculty, the rubric may need to deemphasize the mechanics of CLD creation (dimensions 1a, 2a, and 3a in the current rubric) and foreground real-world insights (dimensions 2c and 3b in the current rubric). A useful analogy might be rubrics for English Language Arts: a rubric for teaching English as a second language might include separate dimensions for spelling, vocabulary, subject-verb agreement, and sentence completeness, whereas a rubric for a college-level literature course might collapse all those elements into a single dimension of English usage & grammar. Also of value could be instructional materials to facilitate and incentivize student use of CLD best practices, such as a check-your-work checklist to complete before submitting.



Narrative: With the onset of microbakeries and small local shops, more healthy food will be available to the people living in the community. An abundance of healthier food will benefit the community by decreasing health issues and poverty linked to food deserts. These small shops will also provide jobs for community members, further increasing money back into the community. The more money that goes back into the community, the more availability there is for more local shops to open, and a decrease in the reliance on shops that do not sell fresh/healthy goods.

Figure 7. Example of student response to the individual activity in the Psychology of Food class, based on a reading passage from Sowing Seeds of Hope in a Food Desert (Mason, 2022). The narrative has been transcribed verbatim from the student handwriting.

Looking across the 17 responses to the food oasis activity suggests that this group is in an interesting liminal area: their answers are far from totally wrong – but not close to totally right. They are deploying systems thinking to generate insights about an important phenomenon in their domain of study. But they seem capable of doing much more. We have identified two potential leverage points (Meadows, 2011, Chapter Six) where a slight instructional nudge could move students towards even deeper understanding of the system in question.

The first potential leverage point is to guide students to a more focused conversation about what entities do and do not belong in the loop itself. The student of figure 7 put "local shops" in the loop, which we find to be insightful. In contrast, 15 of the 17 students put "microbakers," "microbakery," "microbakeries," or "small bakeries" in the loop. In our interpretation of the reading passage, we think it would be more accurate to model the opening of a Bread Angel microbakery as an initial exogenous nudge that kicks off the virtuous cycle, as telegraphed by Mason's assertion that: "The presence of a microbakery might also help to inspire and enable other small locally-owned businesses to open and thrive. In a food desert, a microbakery can nurture the first shoots of a 'food oasis'." Students could be guided towards this conclusion by asking what exactly is getting larger and larger in this community as it transitions from a food desert to a food oasis? It is *not* number of microbakeries. A substantial increase in the number of microbakeries to dozen or two dozen is implausible as they would compete with each other and undercut each other's profitability, and thus would not drive the improvements in jobs, income, and health claimed in the students' diagrams and narratives. When students are struggling with what should and should not be part of the loop, another useful teaching move is to insist that entities placed in the loop must be able to be both actor and acted-upon.

A second leverage point for instruction is to have students more carefully scrutinize and defend why they think it is legitimate to draw each of the causal links in their diagram. At a glance, many of the 17 student responses look like pretty good CLDs, with a circuit of nodes and arrow, S's or occasionally O's on the arrows, and a + in the center. But on closer inspection, many individual A→B links do not stand up to scrutiny. In other words, they are Unacceptable on rubric dimension (2c): Connection of the depicted links to real world system. These include:

- more choices →microbakery
- reduce poverty & create more food options (food oasis) –S→ microbakeries sell bread & generate income
- community access to fresh baked goods –S→ baking/selling real bread
- expanded market –S→ teach people how to bake/set up microbakeries
- microbakeries --> food deserts
- food deserts →jobs
- microbakeries –S→ Food oasis

An instructional intervention that might help with this failure mode would be to ask students as they talk their way around the loop: Do you have in mind a mechanism by which a change in this A would tend to bring about a change in this B? Do you have empirical evidence that when a change in A occurs, a change in B tends to follow?

Next steps:

Kastens, Shipley and colleagues are scheduled to run a workshop on *Wrapping your Head around Environmental Problems by Leveraging Feedback Loop Thinking* at the Earth Educators' Rendezvous in July of 2024, for a target audience of Earth and environmental science faculty (Kastens, et al, 2024). On Day 1, participants will practice strategies for discerning – in data, in language, and in lived experience – that an Earth- or Earth-human system is driven by feedback. On Day 2, participants will use causal language, causal loop diagrams, and behavior over time graphs to depict and explain the system of influences by which positive feedback loops cause growth or collapse and negative feedback loops cause stability or oscillation. On Day 3, participants will use CLDs and causal language to propose and critique potential interventions into FL-driven problems.

In preparation for the workshop, we are compiling a library of reading passages that can be used for the find-the-loop activity, drawing from everyday systems and Earth systems. From the workshop, we hope to recruit collaborating instructors to try the find-the-loop activity with their students and work with us to analyze their students' CLDs and diagrams. Out of that collaboration, we envision two versions of the rubric emerging: one version suitable for educational researchers (probably similar to our existing rubric), and a simplified practitioners' version that foregrounds the features that frontline educators find to be most indicative of student mastery of feedback loop thinking as deployed in a disciplinary context.

As we move towards including the find-the-loop and draw a CLD activity in our instruction, it is our intention to include the requirement that students must write a narrative to accompany their CLD. We encourage readers of this paper to do the same. In addition to its value to the instructor, we think that writing the narrative has pedagogical value for the learner. Across a wide range of task-types and content domains, education researchers have documented the so-called "self-explanation effect," in which learners perform substantially better if they are prompted to speak or write an explanation of the concept they are trying to understand or the problem they are trying to solve (Chi, et al, 1994; Kastens & Liben, 2007). This effect, which is very strong, is attributed to a process in which creating and articulating an explanation helps people find and repair conflicts between their evolving mental model and the information in the text they are trying to make sense of (Chi, 2000).

We would welcome collaboration with SDS members who would like to try our instructional approach and scoring rubric with their own students.

Acknowledgements

Shipley and Kastens acknowledge financial support from the National Science Foundation IUSE (Improving Undergraduate STEM Education) program, through grants DUE21-42010 and DUE21-41939 respectively. We thank the students and scientists who participated in the field-testing phase of this project, our co-PIs on the IUSE project Logan Brenner and Alexandra Davatzes, and systems thinking educator Ezgi Şenyurt, who participated in the conceptualization and articulation of the rubric.

References cited

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad and the ugly. *College Teaching, 53*(1), 27-30. Retrieved from http://www.georgianc.on.ca/staff/ctl/wp-content/uploads/2009/02/rubrics_article_montgomeryk.pdf

Attia, P., with Gifford, B. (2023). *Outlive: The Science and Art of Longevity*: Harmony, 496pp.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in Instructional Psychology: Volume 5 Educational Design & Cognitive Science* (pp. 161-238). Mahwah, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., de Leeuw, N., Chiu, M. H. , & LaVancher, C. (1994) . Eliciting self-explanations improves understanding. *Cognitive Science*, 18,439-477.

Ford, A. (2010). *Modeling the Environment*, Second Edition. Spain: Island Press.

Hemingway, E. (1926). *The Sun Also Rises.* Scribners.

Kastens, K.A., L. Brenner, A. Davatzes, and T.F. Shipley, 2024. Wrapping your head around environmental problems by leveraging feedback loop thinking. Workshop website for the Earth Educators' Rendezvous: https://serc.carleton.edu/earth_rendezvous/2024/program/morning_workshops/w6/index.html

Kastens, K. A., and L. S. Liben (2007). Eliciting self-explanations improves children's performance on a field-based map skills task. *Cognition and Instruction*. v. 25, pp. 45-74.

Kastens, K. A., & Shipley, T. F. (June 2021). Linguistic clues for spotting feedback loops in the wild. *Creative Learning Exchange Newsletter, 30*(2). Retrieved from http://static.clexchange.org/ftp/documents/x-curricular/CC2021_LinguisticClues.pdf

Kastens, K.A., & Shipley, T.F. (2023), How the human mind thinks and learns about feedback loops. *Proceedings of the International Systems Dynamics Conference,* https://proceedings.systemdynamics.org/2023/papers/P1208.pdf

Kastens, K. A., Shipley, T. F., Boone, A., & Straccia, F. (2016). What geoscience experts and novices look at, and what they see, when viewing data visualizations. *Journal of Astronomy & Earth Science Education, 3*(1), 27-58. Retrieved from http://www.cluteinstitute.com/ojs/index.php/JAESE/article/view/9689

Kenzie, E.S., Wakeland, W., Jetter, A., Lich, K. H., Seater, M., Gunn, R., & Davis, M. M. (2024). Protocol for an interview-based method for mapping mental models using causal-loop diagramming and realist interviewing. *Evaluation and Program Planning 103*, 102412. doi: https://doi.org/10.1016/j.evalprogplan.2024.102412

Lyneis, D. (2001 revision). Rubrics for Understanding: Using Systems Dynamics Tools. *Creative Learning Exchange*. http://static.clexchange.org/ftp/documents/implementation/IM2001-03RubricsForSDTools.pdf

Mason, J. (2022). Sowing seeds of hope in food deserts. *SustainWeb.* https://www.sustainweb.org/realbread/articles/mar22-microbakery-role-in-food-deserts/

Meadows, D. H. (2015). *Thinking in Systems.* Chelsea Green Publishing.

Reddy, Y. M., & Andrade, H. G. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education, 35*(4), 435-448.

Resnick, I., Kastens, K., & Shipley, T. F. (2018). How students' reason about visualizations from large professionally collected data sets: An eye-tracking study of students approaching the threshold of data proficiency. *Journal of Geoscience Education, 66*, 55-76.

Richardson, G. P. (2011). Reflections on the foundations of system dynamics. *System Dynamics Review, 27*(3), 219-243.

Richardson, G. P. (2020). Core of Systems Dynamics. In B. Dangerfield (Ed.), *Systems Dynamics: Theory and Applications* (pp. 11-20): Springer.

Schirber, M. (2015). "Snowball Earth" might have been slushy. NASA Goddard Institute for Space Studies, online at https://www.giss.nasa.gov/research/features/201508_slushball/, reprinted from Astrobiology Magazine.

Senge, P. M. (2006). *The fifth discipline: the art and practice of the learning organization.* Rev. and updated. New York, Doubleday/Currency.

stardiverdwr (n.d). Personal System Causal Loop Diagram. https://www.rcampus.com/rubricshowc.cfm?code=HX9W2W8&sp=yes&.

Sterman, J.D. (n.d). Fine-tuning your Causal Loop Diagrams--Part I, *The Systems Thinker*, https://thesystemsthinker.com/fine-tuning-your-causal-loop-diagrams-part-i/

Sterman, J.D. (n.d). Fine-tuning your Causal Loop Diagrams--Part II, *The Systems Thinker*, https://thesystemsthinker.com/fine-tuning-your-causal-loop-diagrams-part-ii/

Sterman, J.D. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin McGraw-Hill, Boston.

Stevens, Dannelle D., and Antonia J. Levi. Introduction to Rubrics : An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning, Taylor & Francis Group, 2012. ProQuest Ebook Central, https://ebookcentral-proquest-com.tc.idm.oclc.org/lib/teacherscollege-ebooks/detail.action?docID=4438553.

Stuntz, L., Costello, W., & Lyneis, D. (2000). Rubrics for Understanding. *Proceedings of the 18th International Conference of the Systems Dynamics Society*, https://proceedings.systemdynamics.org/2000/PDFs/stuntz43.pdf

Tabuchi, H. (2024). Air-conditioning use will surge in a warming world, U.N. warns. *New York Times*, online at: https://www.nytimes.com/2023/12/05/climate/air-conditioning-electricity.html#:~:text=As%20the%20planet%20warms%2C%20installed,grids%2C%20particularly%20in%20developing%20economies.

Appendix 1: Rubric for Evaluating Student Causal Loop Diagrams (CLDs)

		Exceeds Expectations (7)	Meets Expectations (5)	Room for improvement (3)	Unacceptable (1)
Nodes	(1a) Nodes on the CLD		• All nodes that are part of a loop depict something that can potentially increase or decrease, such as a quantity (e.g. # of people), state (e.g. anger, temperature), or attribute (e.g. strength), *and* • Nodes do not state or imply the direction of change.	• Nodes are a mixture of unacceptable and meets expectations forms.	Most nodes in loop: • depict something that cannot increase or decrease *or* • indicate the direction of change, usually as a subject-verb pair (e,g. hostility rises)
Causal Relationships / Connections	(2a) Links/arrows on the CLD		• All arrows have either a +/- or an S/O *and* • The polarity of every link matches the narrative *and* • The number of links and nodes is sufficient to convey all of the relationships that are important for the intended use of the CLD.	• Some arrows lack +/- or S/O. • The polarity of some links does not match the narrative. • There are not enough links and nodes to convey all of the relationships that are important for the intended use of the CLD.	• CLD has lines instead of arrows. • Arrows have heads on both ends. • Arrows bifurcate or merge. • All arrows lack +/- or S/O.
	(2b) Narrative description of causal links	(for some loops only) The narrative indicates that the loop can be driven in either direction, depending on the initial nudge; i.e. + loop can go down-down or up-up; - loop can go down-then-up or up-then-down.	The narrative progresses methodically around the loop, describing how a change in each node results in a change in the next node downstream. The narrative may start at any point in the loop, but must unambiguously return to the starting node. Explicitly causal language is used for at least some links.	The narrative mentions most of the relationships in the diagram, but skips around and the logic can be hard to follow.	• The narrative does not discernably connect to the diagram, e.g. omits most of the links that are in the diagram. • The narrative does not "close the loop" (return to the starting point).
	(2c) Connection of depicted links to real-world system	For crucial link(s) in the CLD, the narrative states a plausible mechanism and/or empirical evidence that the depicted relationship exists in the real-world system.	For every link in the CLD, the narrative expresses a causal relationship that does exist in the real-world system or that could plausibly exist in a hypothesized system, as judged by a coder/ instructor with knowledge of the system.	For some links in the CLD, the narrative expresses a causal relationship that meets expectations. For other links, the connection to the real world is garbled, ambiguous, or wrong.	The narrative does not demonstrate that the writer understands why or how entities depicted in some nodes influence entities depicted in other nodes in the real-world system.

Net Effect of Loop as a Whole	(3a) Diagram depiction of feedback loop		• Diagram includes text nodes and arrows which form a closed loop; arrows all point in the same direction (CW or CCW around the loop) *and* • There is a R/B or +/- label for the net outcome of the loop in the center of the diagram *and* • The type of loop indicated in the center of the diagram matches the cumulative effect of stepping around the loop.	Diagram includes nodes and links which form a closed loop, *but*: • there is no indicator of the existence or type of loop (R/B or +/- or caption), *or* • the type of loop indicated in the center of the diagram disagrees with the cumulative effect of stepping around the loop.	• Nodes and arrows do not form a closed loop.
	(3b) Understanding of impact of this loop on a broader system within which it is embedded.		• The narrative clearly describes the net effect or impact of the loop taken in its entirety. • The impact of the loop on a broader system within which the loop is embedded is explicitly discussed, and is an accurate representation as judged by a coder/instructor with knowledge of the real-world system.	• The narrative says something about the net effect of the loop, but it is garbled or incomplete, or it is unclear why the sequence of nodes and links in the CLD would lead to that outcome. • The narrative has a generalized statement of net effect that could apply to any + or - loop, but does not address the impact of this specific loop in the context of a broader system.	• The narrative does not address the impact or net effect of the loop at all, *or* • The net effect or outcome attributed to the loop is not a plausible outcome of the sequence of nodes and links shown in the CLD, as judged by a coder/instructor with knowledge of the real-world system.

Notes: CLD = Causal loop diagram. "R" stands for "reinforcing" and is synonymous with "+". "B" stands for "balancing" and is synonymous with "-". "Real-world system" is used to refer to the system of which the CLD and the narrative are representations or models; for some assignments, the system being represented could be hypothetical or imaginary rather than literally true to life at the current time.