@CONCORD

vol.26 • no.1 • Spring 2022

The Concord Consortium
concord.org

Perspective:
Science Starts with Wonder

By Chad Dorsey

Science is among the most fundamental of all pursuits, engaging us in elemental questions that capture and motivate us. Wonder at the world's unexpected surprises and curiosity about the unexplained galvanizes our irrepressible drive for understanding. This natural inclination holds significant potential for science education.

While many groundbreaking discoveries can be expected—from the ultimate revelation of an elusive particle to the confirmation of an essential genetic or molecular structure—the moments that truly matter can be tantalizingly elusive. A fleeting glimpse opens the door to a new paradigm. A tiny deviation causes one to question all that has come before. As Isaac Asimov famously noted, "The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka,' but 'That's funny…'"

There are countless historical precedents for such discoveries. In the late 17th century, Antonie van Leeuwenhoek discovered a world of wonder in a simple drop of water—a teeming universe of life previously unknown to all, both within the realms of science and beyond. Spurred by this unexpected discovery, he cataloged this previously invisible world for the illustrious ranks of the academy. Yet his bold claims—of countless omnipresent, minuscule living creatures—fell on skeptical ears. Unfortunately, the uneven development of microscopy tools precluded the ability of most others to see these "animalcules" for themselves. His discoveries were banished from the science mainstream for over a century and a half.

Just after the turn of the 20th century, another pioneer stumbled upon an unexpected moment of wonder. Charles Henry Turner, long fascinated by the actions of insects and other animals, moved a bottlecap next to a small hole he had dug. As an ant proceeded into this new, false burrow, Turner recognized the significance—the ant was navigating based on landmarks in the world around it, incontrovertible evidence that the insect was learning. Although Turner was the first Black man to publish a paper in *Science*, racial discrimination stood as a constant barrier to wide airing of his discoveries. For many decades, his trailblazing ideas proved too forward thinking for the narrow-minded ranks of the more traditional scientists of his time.

Common threads bind these stories together—the pure surprise of the unexpected, the beckoning pull of the unknown, and the deep inspiration of wonder. Just as these experiences have pushed scientists forward for generations, they serve as equally powerful forces in our own personal experiences. Their universal attraction makes them valuable guideposts as we consider what, and how, to encourage future generations to learn about science.

Leaning into the unexpected

Encountering the unexpected is a clear and compelling inspiration for scientific investigation. When a beaker suddenly foams or changes color, when a ring appears around the moon on a frosty evening, or when an animal moves in a seemingly unnatural way, the jolt of the unexpected provides a revelation that one's concept of the world is incomplete.

Pedagogically, this is highly useful. Science teachers routinely leverage surprising demonstrations to hook students' interest, often to great effect. Our data science education research shows that students examining data attend more carefully to unexpected features, and that they will spend considerable time working to understand the nature and origins of such anomalies.

Oddities present lessons and opportunities far beyond the classroom and into the scientific enterprise as a whole. Stanford professor Garry Nolan, a renowned cancer researcher, inventor, and co-founder of seven companies, states the importance of heeding this lesson. "It's not the data that falls in line that's so interesting," he explains. "It's the spot off the graph that you want to understand. When something is way off the graph, that's the interesting thing, because that's usually where discovery is."

While examining anomalies, learners exercise important habits of mind for scientific critical thinking. And when supported in maintaining and honing these habits over time, learners can gain the inclination, and develop the fortitude, to stop and notice the unexpected, pose the right questions, and follow where the data leads. To get there, however, we must explicitly encourage this impulse.

Chad Dorsey
(cdorsey@concord.org)
is President and CEO of the Concord Consortium.

Wonder at the world's unexpected surprises and curiosity about the unexplained galvanizes our irrepressible drive for understanding.

Following the unknown

Sadly, exploring the unknown is miles removed from most students' classroom experience. In far too many cases, we reduce education to a game in which teachers and students learn early on that success comes from closely following well-established rules and behaviors. In this traditional game of science learning, the sage teacher holds the answers and diligent students obediently follow along, remaining in their proverbial seats and prescribed roles until they are granted the appropriate knowledge and learn to parrot it back.

Just as frequently, this pattern pervades laboratories as well. Investigations abound whose goal is for students to confirm what is already known and to generate lab reports that conform to the composition rules of a carefully defined genre. These commonplace practices rob students of the most important aspect of science—the experience of the truly unknown.

Fortunately, this cycle can be broken. Rather than follow a pre-ordained process for laboratory investigations, teachers can provide students with the goal and let them reason out their own paths, just as scientists do. Contrary to the popular notion of an archetypal "scientific method," scientists tinker, explore, hit dead ends, and retrace their steps. They work out problems through noisy debates with colleagues and on long, solitary walks. Discovery of the true unknown comes through persistence, patience, and puzzling. To provide students a taste of true science, we need to make room for true discovery.

When confronted with the unexplained, an honest scientist digs in, chooses a path, and persists, following the data to uncover the wonder it reveals. Yet as history has shown, doing so can prove daunting, especially when pursuit of the unknown collides with the stigma of narrowly defined societal expectations. Whether in a professional laboratory or a high school classroom, fostering true scientific discovery requires us to remain open. Teachers must expand norms, provide students with open-ended technology tools that enable broad investigation, and encourage the early "messing around" stages that are central to the scientific process. Teaching today's students to follow their own scientific north stars may be one of the most crucial learnings we can provide. One such innovative teacher tells the story of her journey in this issue.*

Welcoming wonder

Across time immemorial, humans have looked up at the night sky and mused about the deepest questions of existence. Such innate human wondering binds us together and inspires new exploration and scientific discovery.

This is the overarching lesson for teachers of the next generation. We cannot realize the wonders of the world for our learners. However, we can guide them thoughtfully, giving them space to follow the unexpected glimmers that speak to them individually. We can provide learning environments that value what they know and encourage them in chasing that wonder wherever it leads. We can and must work to ignite the spark for *all* students.

Doing this begins by acknowledging and welcoming wonder, in its many places and forms. The Concord Consortium believes that its mission—to innovate and inspire equitable, large-scale improvements in STEM teaching and learning through technology—holds the key to discoveries that could change the course of history. In doing so, we hope learners experience the inspirational power of wonder. Only in this way can we uncover answers to the unexplained.

* Learn more about Julia Wilson's inspirational teaching in our *Teacher Innovator Interview* on page 15.

Exploring Artificial Intelligence with StoryQ

By Jie Chao

Jie Chao
(jchao@concord.org)
is a learning scientist.

Mia submitted her college applications—six schools carefully chosen with the help of her guidance counselor. Now the anxious waiting begins. But when an article on Artificial Intelligence (AI) catches her attention, her anxiety increases. She learns that some colleges are considering the benefits and risks of using AI systems to improve the admission process. Mia's college essay might be "read" by AI instead of humans. The thought of a heartless machine scanning her work gives her chills. Will her chance to get into her dream college be influenced by an AI model that predicts her likelihood to be a successful student?

Our Narrative Modeling with StoryQ project is creating tools, materials, and opportunities for young people like Mia to gain a fundamental understanding of AI as well as become a powerful voice in a society being rapidly reshaped by AI technologies. In this article, we walk through how a simple AI system works and how it is built.

Clickbait filter

Nowadays, we read most news on the Internet because it's convenient and has a wide coverage. But some headlines—called clickbait—are designed to entice us to click for content that is useless, deceptive, or misleading. Imagine if we could design an AI filter to detect and potentially remove clickbait from view. The key ingredient is a training dataset that includes a large number of examples of headlines labeled by humans as clickbait or non-clickbait (Figure 1).

#	Headlines	Labels
1	Can you match the actor to their animated film roles?	Clickbait
2	Covid in the Northeast	Non-clickbait
3	32 gift ideas for the Canadian in your life	Clickbait
4	Space Shuttle Discovery back in Florida	Non-clickbait
(996 more rows)		

Figure 1. A small subset of a training dataset for an AI filter on clickbait.

#	"you"	"the"	"in"	*(394 more columns)*	Labels
1	✓	✓		-----	Clickbait
2		✓	✓	-----	Non-clickbait
3		✓	✓	-----	Clickbait
4			✓	-----	Non-clickbait
(996 more rows)					

Figure 2. A feature table consisting of unique words frequently used.

In our StoryQ app, developed as a plugin for our Common Online Data Analysis Platform (CODAP), we can extract many unique words from the headlines and transform the headlines into lists of words organized in a *feature table* (Figure 2).

When we run this feature table through a machine learning program, the program discovers the rules of how these words are related to the labels and encapsulates these rules in a model. With this very basic approach, the model correctly predicts the labels for almost 90% of headlines. The approach is by no means state of the art, but it helps us see the inner workings of AI systems. But before we look at how the model works, it's useful to reflect on how our brains work.

Consider this headline: *The Struggles of Being a Sleepwalker.* Curious? Want to know more? Tempted to click on it?

You are not alone! Most people are intrigued by the chance to peek into this unusual syndrome. However, the headline's hyperlink leads to a distasteful short video surrounded by a blanket of advertisements and more clickbait—exactly what the headline was designed to accomplish.

What makes us take the bait? The words "struggles" and "sleepwalker" bring to mind a topic that we know a little bit about and are interested in learning more. The content words do the "baiting" to our human brain.

Does the model work the same way? Yes and no. While the model correctly labels it as clickbait, the words "struggles" and "sleepwalker" are completely irrelevant. In fact, they are not even in the model. What matters are the words "the," "of," and "being," which are function words absent of any content.

The following annotated version is similar to what the model "sees." The grayed-out words are ignored because they are not part of the model at all. The model only knows the underlined words, the so-called features, and uses them to make the prediction.

The Struggles of Being a Sleepwalker

For each feature, the model has learned from numerous examples how strongly it is associated with clickbait while considering the presence of other features. This relationship is captured as a numeric value called weight. The word "the" has a weight of 0.505, which is a fairly large positive weight, given that the weights of all features range from -0.636 to 0.952. The word "of" has a small negative weight of -0.133. "Being" is almost neutral, with a weight of 0.003. The three features and their weights, summarized in Table 1, are all the model knows about the headline.

Features	Weights
the	0.505
of	-0.133
being	0.003

Table 1. The features and weights of an AI model.

With such seemingly limited information, how does the model use these features to make a prediction? Consider a bar graph where bar lengths represent the feature weights (Figure 3). Blue bars are the features pulling to the clickbait side and the orange bar pulls to the non-clickbait side. On average, there is a strong pull to the clickbait side.

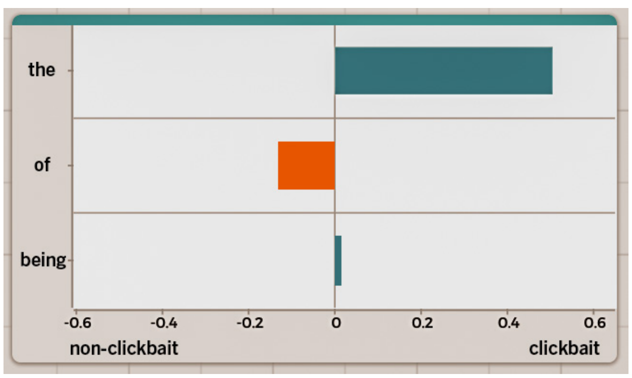


Figure 3. Bar graph of the feature weights.

The model goes through similar reasoning by plugging the numbers into a formula, computing the probability of the headline being clickbait, and labeling it as clickbait based on a predefined probability threshold.* Counterintuitive as it can be, this is how the model predicts clickbait, sometimes using contentless words.

How did the machine learn?

The intriguing thing is how the machine learned this trick. Why do contentless words like "the" help predict clickbait? Recall that the machine learning program ran through the feature table and computed a weight for each feature. On the surface, the model just looks like a list of features, each with a weight (Table 1).

How does this simple list capture the rich meaning in the data? To get some insight into this question, we need to look at the features in context. For example, the word "the" seems to be a neutral word—we use it to make grammatically correct sentences. Surprisingly, in the clickbait model, the word "the" carries a large positive weight, making a headline much more likely to be clickbait if it contains this word. Why does a contentless word indicate the likelihood of clickbait? Let's look at a few examples:

How well do you actually know the Addams family?

How well do you remember Season 5 of the Walking Dead?

Can you identify the Janet Jackson music video from a single screengrab?

Notice that what follows "the" are popular culture topics, such as TV shows and celebrities, which are very common in clickbait. Let's look at a few more:

18 times Squidward perfectly captured the dating struggle

40 country songs that defined your life in the early 2000's

The recession may be a boon to book sales

Look around the word "the." Most people are familiar with "dating struggle," "early 2000's," and "recession" as common social phenomena. One final set of headlines:

The hardest Thanksgiving poll you will ever take

We know which celebrity you dislike the most based on your zodiac sign

The toughest Dragon Ball A quiz you'll ever take

It's not difficult to notice the exaggeration of "hardest," "most," and "toughest," typical in clickbait. According to English grammar, such adjectives must be preceded with "the."

So, does "the" capture any meaning? The answer is no if we consider the word by itself. But if we look at why it is used in certain contexts and what meanings it is associated with, the answer is yes. In these cases, the ties between "the" and popular culture topics, social phenomena, or comparative words come from grammatical rules in English.

Many features, especially those carrying large positive or negative features, capture meanings beyond their own definition and have more to do with clickbait as a special genre. For instance, "you" is a strong clickbait feature, frequently used to convey an invitation or imperative to the reader. The word "are" is also a strong indicator frequently used before adjectives or to create questions, which are common styles for clickbait. Numbers are also common clickbait indicators used in listicle headlines. While a computer model does not feel enticed or curious as we humans would, it can "know" what each feature signals about the mechanism underlying our perceptions.

Exploring AI with StoryQ

In real practice, AI systems are far more sophisticated, though they are developed in much the same way. Mia may not feel better about the college admission process knowing how AI systems are created, but we hope such background emboldens her to question the appropriateness of the AI applications she encounters in her life. StoryQ is designed to help students learn about AI's power and its fallibility.

* The StoryQ app allows learners to train text classification models using the logistic regression model. When in use, the model computes the probability of a new headline being clickbait and labels it as clickbait if the probability exceeds 50%.

LINKS

StoryQ – concord.org/storyq

StoryQ app – concord.org/storyq/app

StoryQ curriculum – learn.concord.org/storyq

Clickbait filter – http://short.concord.org/lqi

(continued on p.6)

(continued from p.5)

Q&A with Carolyn Rosé, Ph.D.

Carolyn Rosé is a Professor of Language Technologies and Human-Computer Interaction at Carnegie Mellon University, Interim Director of the Language Technologies Institute, and Co-Principal Investigator of the StoryQ project

What is the goal of the StoryQ project? And what is your role?

My Ph.D. students and I have taught the StoryQ team about text mining and we serve as curriculum advisors on the project. In addition to developing and researching curriculum for high school students, the goal of the StoryQ project is to help people think more critically about what they read regarding AI in the media. There's a lot of unproductive hype and fear out there and we want students to know enough about AI that they neither believe the hype nor succumb to fear. Another goal is to help students envision future careers in AI. In our professional development trainings with teachers, we describe the programs at Carnegie Mellon and the different career paths possible for students.

How would you describe AI to a novice? What's the difference between AI and machine learning?

The field of AI started around 1950 with an effort to get computers to behave in intelligent ways through reasoning that was automated. Machine learning is about pattern recognition and it's based in the field of statistics. Machine learning is just one aspect of AI.

"Exploring Artificial Intelligence with StoryQ" posits the idea of a college admissions office utilizing AI. What would you say to Mia to make her less anxious about their use of AI?

I would tell her emphatically not to worry! It's important to realize that admissions essays are a very different kind of writing. The College Board has used automated assessment on writing samples that are very topically focused and very narrow in terms of content, things you could build a model to make a prediction on. College essays are very broad and personal, and it's not entirely clear what a college admissions committee wants to learn from them. If they were simply looking for "good writing," I would trust a machine, which can check for grammar and well-formed sentences. This highlights an issue of AI in the media in general—people don't see these distinctions. They don't realize how the data and the nature of the judgment are different. Just because a feature got a high weight, that doesn't mean it's finding out what makes people click on clickbait.

So, are words like "the" and "of" actually indicators of clickbait?

I don't think you can conclude that just because you know which things people clicked on that you are identifying what made people click. You're probably identifying something that correlates with what made them click, and that might not even be in the writing at all. It might be an abstraction over the writing, but not in any individual word. So, if you represent the data in terms of features that are words, then you don't have features to put weight on that are the real reason why somebody clicked on it. The word "the" is thus probably what we would call "misplaced weight." It's one existing feature that correlates with the real thing, which might not be anywhere in the feature space. The key point we want students to learn is that behavior that looks intelligent is not necessarily driven by a human kind of intelligence.

The article concludes that AI is both powerful and fallible. Where do you see its power and its failure?

Computer programs have the ability to process a lot of information very quickly. That seems very powerful. For instance, when you start typing in Gmail, it shows you what you might want to say next. That comes from a giant language model, but because people see that as a prediction about what they were thinking, they feel like it's reading their mind. Humans are good at finding regularities, but not as good as computers. But humans can do creative problem solving way better.

How can a teacher use StoryQ?

We believe that feature engineering is a great fit for English Language Arts classes at the high school level. Students can learn about language constructs and what it is about language that makes it seductive or interesting.

You can try the curriculum at **learn.concord.org/storyq**

StoryQ

Monday's Lesson:

Tephra in the Wind

By Chris Lore

Chris Lore
(clore@concord.org)
is a research assistant.

The world was rocked by several large volcanic eruptions in the past year, from the explosive activity of Mount Sinabung in Indonesia to the incredible lava flows of La Palma in the Canary Islands and Fagradalfjall in Iceland. Eruptions of all types display the incredible power of volcanoes and the unstoppable hazards they create. Volcanic ash particles called tephra ranging in size from tiny glass fragments to large boulders explode from these eruptions, harming people and destroying property. In this Monday's Lesson, students explore environmental factors that influence where tephra falls.

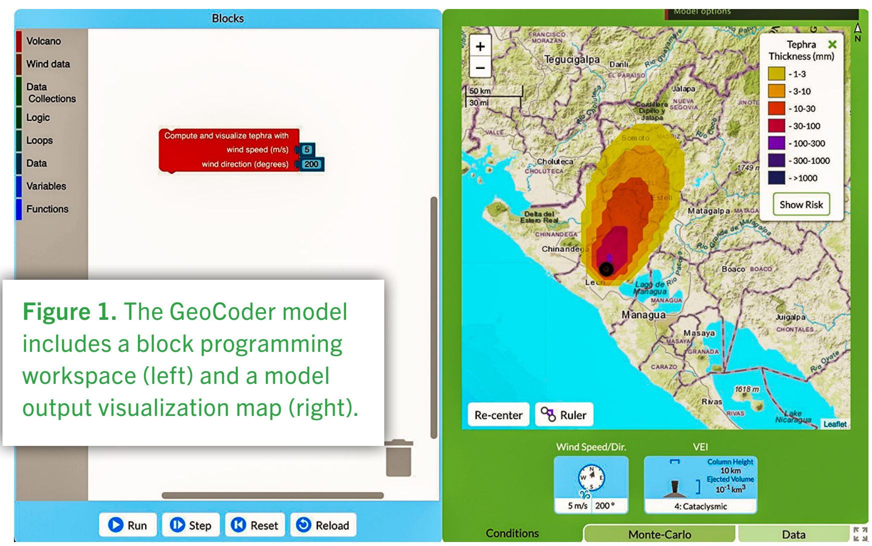


Figure 1. The GeoCoder model includes a block programming workspace (left) and a model output visualization map (right).

Our GeoCoder model allows students to experiment with both wind speed and direction to see how these factors affect the tephra distribution around a volcano. The model focuses on the Cerro Negro volcano in Nicaragua, an active cinder cone volcano that has erupted 10 times in the last 60 years. Using block code, students set values for wind speed and direction, which the simulation uses to produce a tephra distribution. The model output visualizes the area covered in tephra as well as the thickness of tephra on a map of the region around Cerro Negro.

1 Open the GeoCoder

Go to **geocode-app.concord.org**

In the Blocks menu on the left, go to the "Volcano" tab. Drag the block called "Compute and visualize tephra with wind speed (m/s) and wind direction (degrees)" into the program workspace. Click the "Data" tab and pull a number block into the slot next to wind speed. (They fit like puzzle pieces and lock into place.) Set the wind speed by clicking the number to change it to any value between 0 and 30 meters per second (or up to about 70 mph). Next, drag another number block and put it in the wind direction slot. Set the value between 0 and 360, which defines the direction the wind is blowing from. Finally, run the model.

2 Analyze the tephra distribution

After the code runs, a multicolored tephra distribution appears on the map as colors representing the thickness of tephra (Figure 1). (See the key in the upper right of the map.) Zoom in to see the towns and cities covered by tephra and use the ruler to measure how far the wind blew the tephra from the volcano.

3 Discover patterns

Reset the model, then change wind speed and/or wind direction to explore their effects on tephra distribution. How does the shape of the distribution change as wind speed increases? Which wind direction causes the most towns to be covered in tephra? To simulate multiple eruptions in a single run, stack two "Compute and visualize" blocks on top of each other with different inputs.

4 Dig deeper

To experiment with the size of the eruption, change the Volcanic Explosivity Index (VEI) value of eruptions (0–8) following the steps above to add a volcano block with VEI and number blocks. VEI accounts for column height and ejected volume, which affect the distribution of tephra.

5 Discuss

What are the impacts of tephra falling on people's homes, cars, and their agricultural land? Only a few millimeters of tephra ash can devastate vegetation by covering leaves and preventing photosynthesis. Thicker tephra deposits (100-300 mm) can collapse roofs and limit ground and air transportation.

Looking for more?

Created by our GeoCode project, the GeoCoder has been embedded in a weeklong curriculum unit for middle and high school classes. Students explore tephra distributions and wind data to make risk assessments for the towns surrounding Cerro Negro.

LINKS

Tephra module – **learn.concord.org/geocode-tephra**

Addressing Socioscientific Issues

While Studying Natural Hazards

By Amy Pallant and Trudi Lord

Amy Pallant
(apallant@concord.org)
is a senior research scientist.

Trudi Lord
(tlord@concord.org)
is a senior project manager.

"The specter of climate change threatens worsening natural disasters, rapid urbanization, forced migration, and economic hardship for the most vulnerable."

– Tedros Adhanom Ghebreyesus, Director-General of the World Health Organization

As global temperatures continue to rise, so do the number, size, and impact of climate-fueled natural hazards, transforming lives across the planet. *The Washington Post* reported that in 2021, one in three Americans had experienced a natural hazard. Climate scientists warn that there will be more frequent and more extreme events. For today's students, the idea of climate change is no longer an abstract concept.

Teaching about natural hazards such as wildfires, floods, and hurricanes and their relationship to climate change poses an opportunity and a challenge. Students are now likely to have some direct experience with living through or knowing someone who has lived through a natural hazard. This makes the study of natural hazards personally relevant and meaningful. Along with this real-world experience, however, many young people are also anxious and angry about the insufficient ways in which climate change is being addressed. The GeoHazard: Modeling Natural Hazards and Assessing Risks project has identified opportunities for students to critically engage in discussions about natural hazards and considerations related to how these hazards are affecting people.

Three of our online curriculum modules about hurricanes, wildfires, and inland flooding are now publicly available. Each GeoHazard module supports student inquiry with embedded Earth systems models that allow students to explore the factors that influence the formation, progression, and severity of each hazard, as well as the factors that contribute to potential risks to people and their communities. During three years of pilot testing the modules, we learned that it is important to help teachers engage students in these topics despite the difficult concepts and emotions that often accompany these discussions.

Situating student learning in the context of socioscientific issues

A Framework for K-12 Science Education identifies the importance of connecting natural hazards to human activities. ESS3B: Natural Hazards and ESS3:D Global Climate Change explore the human activities that have significantly altered the environment and the major factors responsible for these environmental changes. They charge students with investigating human vulnerabilities and understanding human behavior.

Our hurricane, wildfire, and flood modules all address the following questions:

1. What environmental factors play major roles in the development of this natural hazard?
2. How does this natural hazard impact people and their communities?
3. What do rising global temperatures mean for the risks and impacts humans face from this hazard?

Natural hazards are both awe-inspiring and terrifying, and they can spark students' natural curiosity about the power of Earth's systems. Because they are personally relevant to so many students, they also provide a rich context to explore socioscientific issues. This means that in each module, students explore both the scientific phenomena of a natural hazard and the impacts on people and their communities, including the disproportionate consequences for vulnerable members of the population. In the hurricane module (Figure 1), for example, students investigate the reasons why some members of a population cannot evacuate prior to a hurricane, considering how illness, poverty, and occupation (for example, first responders and firefighters) might prevent evacuation. These topics can generate deep discussions and allow students with personal experience to share their stories.

While many teachers are interested in teaching about natural hazards, some find it difficult to discuss the socioscientific issues in their classrooms. Through teacher tips embedded in the modules, we support teachers in handling socioscientific issues, focusing on places where they can facilitate discussions of students' experiences. A separate guide provides additional information on establishing norms for discussions around personal experiences and supporting students to construct explanations and ideas that are based on evidence as much as possible.

Similarly, we prompt teachers to discuss issues around climate change and what people might do to reduce the impacts of future hazards. Helping students learn about the part people play in causing climate change, as well as their role in addressing it, is an important part of the Next Generation Science Standards.

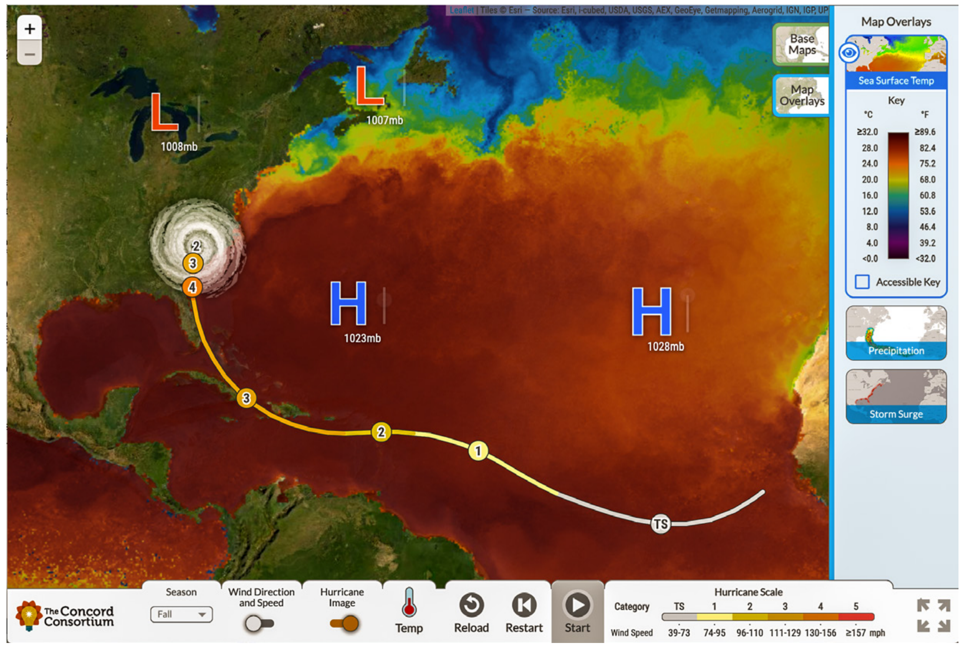


Figure 1. In the hurricane module, students use the Hurricane Explorer to investigate factors that affect hurricane strength.

Inviting students to share experiences

Each module begins with a prompt inviting students to describe their lived experiences with the natural hazard and to share their questions. The goal is to provide an opportunity for students to express their full range of experiences and to support teachers in holding a conversation that could, for example, include small, local floods to catastrophic flood events. Student responses to this question in the flood module (Figure 2) ranged from those who had never witnessed a flood to some who had firsthand experience.

"I never experienced flooding but I could imagine how people were impacted by this."

"I have never been seriously impacted by flooding but I have been stuck in my house and out of my house for multiple days because the water was over the road."

"My grandma's house got flooded and she had to live with us for four months till she would get her house fixed."

"Someone I know has been impacted by a flood. The flood of 2021, one of my sister's friend's house was destroyed by the flood waters and they had lost everything they have owned. They had to quickly evacuate their home with nowhere to stay."

"Yes the last time it flooded our whole yard was covered from our Creek being so high. It was coming in through the side of our house and we had to move everything that was in the den."

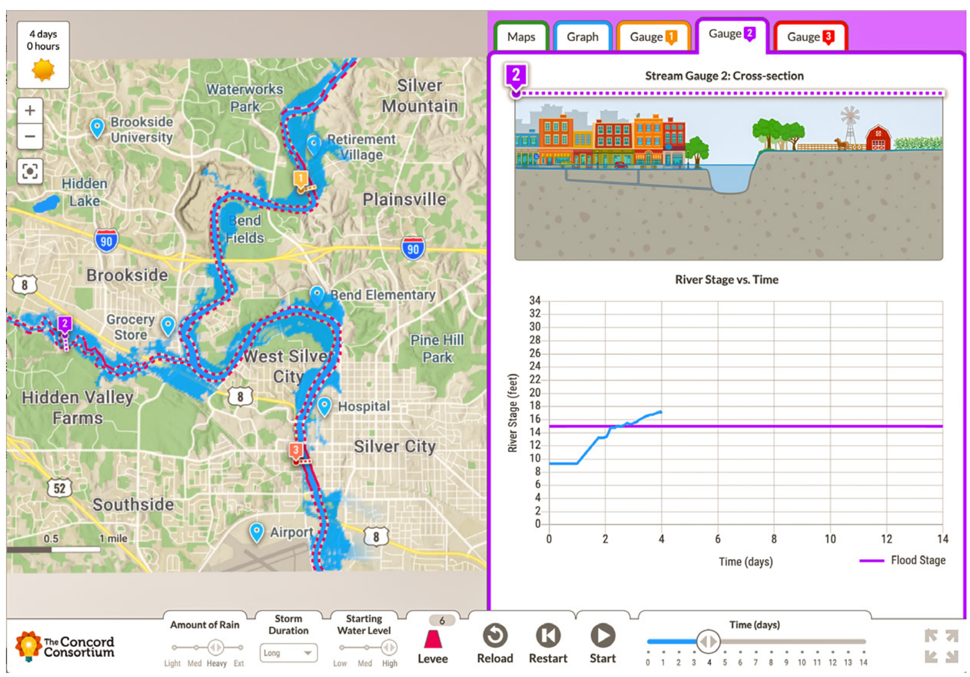


Figure 2. In the flood module, students use the Flood Explorer to consider the effects of flooding on different communities.

As these responses make clear, students are already grappling with natural hazards and the complex ways in which these hazards intersect with their daily lives. Rather than leaving these social aspects out of the science classroom, the GeoHazard modules explicitly integrate instructional strategies for addressing them. By exploring natural hazards in the context of personal and community vulnerabilities, students can expand their notion of why studying these real-world topics is important for tackling the consequences of global warming and adapting to future natural hazard emergencies.

LINKS

GeoHazard
concord.org/geohazard

Data Story Bytes:
Examining Healthy Food Through Data

Hollylynne S. Lee, Michelle H. Wilkerson, David Stokes, and Bill Finzer

Hollylynne S. Lee (hollylynne@ncsu.edu) is Distinguished Professor of Mathematics and Statistics Education at NC State University.

Michelle H. Wilkerson (mwilkers@berkeley.edu) is Associate Professor of Education at the University of California, Berkeley.

David Stokes (djstokes@ncsu.edu) is a doctoral student in Mathematics and Statistics Education at NC State University.

Bill Finzer (wfinzer@concord.org) is a senior scientist.

Every day, many students and their families make choices about what foods to eat. Families tell stories around the table, including stories about the food they're preparing and consuming. But what about the food's nutritional value or its popularity—stories informed by data? The Writing Data Stories project developed curriculum to help students unpack stories about everyday issues like food told through data.

Our goal is to investigate important socioscientific issues such as food, the environment, and climate by analyzing scientific datasets using "data storytelling." We created curricular activities called Data Story Bytes or simply Data Bytes, intended to be "bite-sized" data activities completed in 30 minutes or less to help students interpret data graphs and visualizations related to various STEM concepts. They are similar to the popular *New York Times* feature "What's Going On in This Graph?" and "Data Talks" developed by Youcubed, but are uniquely structured to help frame students' interrogation of data. Using a four-part questioning framework, students consider where the data came from, who is represented or omitted from the data and other equity issues, and the implications of the dataset for their personal lives (Figure 1).

Data Bytes are available as Google Slides for students in grades 6–12 in different subject areas (e.g., science, social studies, health, math, English, business), and they include directions in both English and Spanish to support multilingual learners. The Teacher's Guide includes resources for using and creating your own Data Bytes. In this article, we explore one Data Story Byte in detail, taking a bite out of food data in "What's healthy?"

Making sense of trends and relationships in the data or visualization, what these patterns mean, and how they connect to key science concepts.

Building personal connections by considering how students' own lives and communities may be impacted by or reflected by the patterns found in data.

Reflecting on the context and history of the data, how it was collected, by whom (including what gets "counted" and why), how it is visualized, what might be missing/hidden, and what questions the data can and cannot answer.

Envisioning future uses of data and visualization to expand the investigation, include and explore different perspectives, and highlight the importance of understanding what's happening in the world around us in multiple ways.

Figure 1. The four-part questioning framework for each Data Story Byte.

What's healthy?

Set the stage by asking students to name foods they eat that they consider healthy or unhealthy, creating two lists on a white board or shared Google Document. Do any students disagree with the foods listed as healthy or unhealthy? Why might people disagree?

Explore the graph

Next, share the graph published in the *New York Times* article "Is Sushi 'Healthy'? What About Granola? Where Americans and Nutritionists Disagree" (Figure 2). The graph plots 52 foods by the percent of "everyday Americans" who say the food is healthy vs. the percent of nutritionists who say the food is healthy.

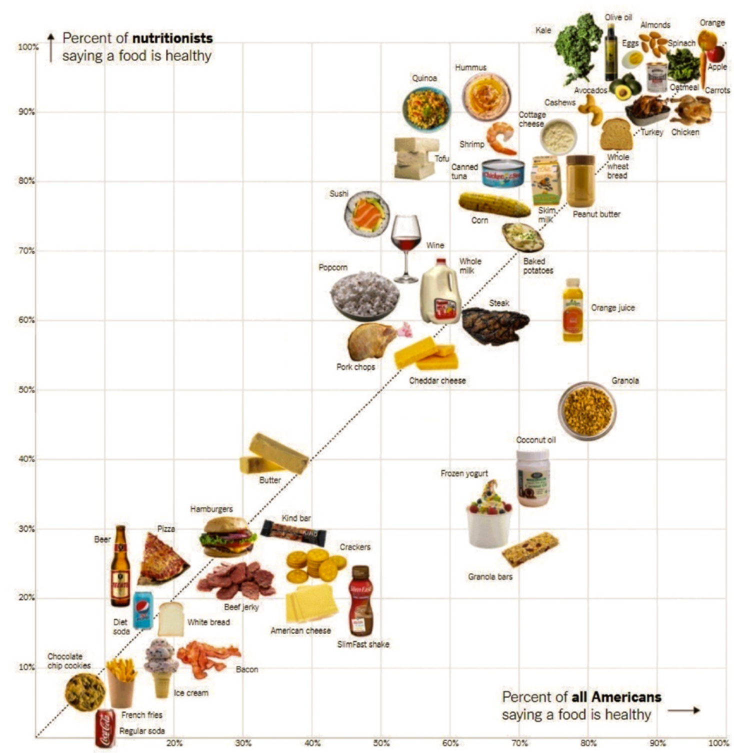


Figure 2. Different foods plotted by the percent of Americans who say the food is healthy vs. the percent of nutritionists who say the food is healthy. Source: The *New York Times*.

Using a copy of the Google Slide Deck for this Data Bytes activity, ask students to describe what they notice in the graph. Encourage students to look for information in the news article that can help contextualize how and why the data was collected, and from whom.

Use the following questions or develop your own to prompt students to think more deeply about the graph. Encourage students to move beyond the data to connect the patterns they see to their personal lives and communities, to consider what perspectives are included or left out, and to develop new questions. Students can work individually, in small groups, or in a whole class discussion.

Making sense of trends and relationships

What is the first thing you notice or wonder when you see the graph?

Were any foods on our healthy or unhealthy list represented in this graph?

What relationships or trends do you notice?

What does the line represent?
Why are some foods closer to the line than others?
What can you say about foods located above the line versus ones below the line?

Personal connections

What personal connection do you have with the data or patterns in this graph?

Can you locate any foods that you eat? If so, explain where they are located in the graph and the approximate percentage of each group considering it a healthy food.
Where do you think your favorite food might be located in this graph?

Which people or groups do you think would feel more or less of a connection with the data or patterns here?

How might this graph change for people with dietary restrictions?
Which foods are missing that are important to you?
Why might they be missing?

Context and history

Who do you think made this graph? Why did they collect this information and create this graph? What did they want to know?

How were the 52 foods chosen for the survey?
What are the strengths and weaknesses of using that method to decide which foods to include?
Who may have been surveyed as part of the "all Americans" group?
Do you think this graph represents foods from a particular culture or a particular part of the world?

How might these patterns be different if the data had been collected about different people or in a different time period?

What might you say about the group of foods towards the lower left of the graph?
How do you think the graph would change if the "all Americans" survey results were replaced with survey results from our class?

Future uses of data

What are some questions you can (or cannot) use this data or graph to answer?

Can you say whether or not your diet is considered healthy based on this graph?
Why might nutritionists disagree about which foods are healthy or not?

What could you do to make this more useful for yourself or others who might not be included?

Would you collect more data, or group or graph the data differently?

Do you think there is a relationship between food cost and its location on the graph?
What data would you need to answer that question?
How would you collect it?

Extension activities

With answers to the above questions as well as new questions inspired by the Data Bytes activity, students make deeper connections between data and their lives. Students can use the Common Online Data Analysis Platform (CODAP) to extend their learning.

- The graph focuses on whether nutritionists and "everyday Americans" agreed that a given food is healthy. But not everyone surveyed entered a response for every food, and some people responded "No" or "Don't Know/No Opinion" for a given food. Explore a CODAP document created from the survey data files linked in the article footnote (Figure 3).
- Although 71% of the public rated granola bars as "healthy," only 28% of nutritionists agreed. However, all granola bars are not the same. Explore a CODAP document with a dataset of nutritional values for 33 different granola bars.

We hope that Data Story Bytes help students critically analyze and interpret data visualizations in ways that connect to their lives and to important issues in society. Perhaps the data will even make for dinnertime conversation about the foods students are eating.

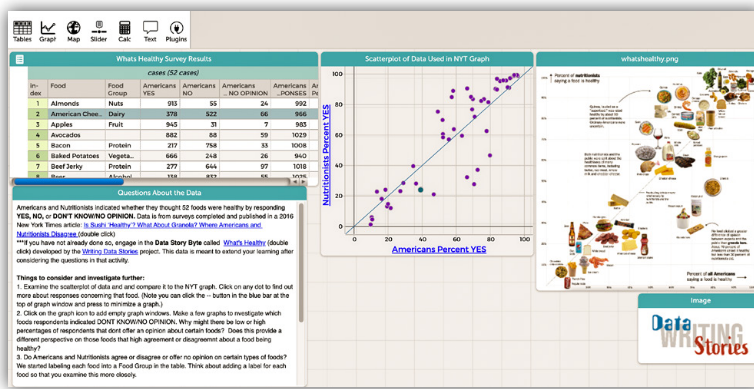


Figure 3. CODAP document created from the survey data files linked from the *New York Times* article "Is Sushi 'Healthy'? What About Granola? Where Americans and Nutritionists Disagree."

LINKS

Data Bytes Teacher's Guide
https://bit.ly/WDSdatabytes

What's Healthy Student Slide Deck
https://bit.ly/DataByteHealthy

CODAP Document of "What's Healthy?" data
https://bit.ly/CODAPHealthy

CODAP Document of granola bars data
https://bit.ly/CODAPGranola

Writing Data Stories project
https://www.fi.ncsu.edu/projects/data-stories

Can Elementary Students Reason About the Invisible?

By Carolyn Staudt, Jamie Broadhead, Ala Samarapungavan, and Lynn A. Bryan

An ice cube melts when warmed, then refreezes when cooled. This simple phenomenon offers kindergarten students the opportunity to learn about solids and liquids, and the change between states. But what if young students could do more than observe macroscopic events? What if they could also develop and use models to make sense of the invisible? To find out, the Sensing Science Through Modeling Matter project developed four apps and a curriculum for kindergarten students to explain states of matter and phase change from a particulate view of matter.

Carolyn Staudt
(cstaudt@concord.org)
is a senior scientist.

Jamie Broadhead
(jameliss.jb@gmail.com)
is Head of Development and Support at Videatives, Inc.

Ala Samarapungavan
(alasam@purdue.edu)
is Professor of Educational Psychology at Purdue University.

Lynn A. Bryan
(labryan@purdue.edu)
is Professor of Science Education at Purdue University.

Our project research focused on kindergarten students' learning about matter. Do their concepts of matter change as they interact with the Sensing Science apps and curriculum? Are different apps associated with differences in students' learning? Are students' conception of particles consistent as they explain varied macroscopic phenomena?

The curriculum

The Sensing Science Through Modeling Matter curriculum includes three multiday inquiry-based lessons around modeling, states of matter, and phase change. In the first lesson, students learn about the use of models and modeling across scales. For instance, globes and maps both model the Earth. Models can also represent invisible things, such as the particles that make up matter.

Students use an app called the Thermoscope to see inside matter; two on-screen circles act like "magnifying glasses." When fast-acting temperature probes are placed into hot and cold water, students use the Thermoscope to observe the relationship between temperature and speed of particle movement (Figure 1). Next, students read the animated Land of Bump story, which illustrates what happens when hot and cold dancers mix together on a dance floor (Figure 2). The animated characters demonstrate a scientifically accurate computational model of water at different temperatures, and act as a metaphor for the motion of particles.

In the second lesson, students learn about states of matter and the relationship between the macroscopic properties of each state of matter and the behavior and arrangement of the microscopic particles that make up matter using one of two apps. The Particle Modeler is designed for students to discover the patterns of particles that adhere to known physical laws (Figure 3). The Thermonator allows students to set rules for the way particles move and interact, including examples that are counter to physical laws (Figure 4). Then students use their bodies to model the arrangement and motion of particles in a particular state of matter, for example, by wiggling in place or moving more freely to represent solids and liquids, respectively.

In the third lesson, students investigate melting, freezing, evaporation, and condensation. First, they observe macroscopic examples and draw their predictions of the particle model for each phase change. Next, they use the same apps to observe the microscopic particle model and revise their predicted models. Finally, they construct human models of each phase change as in the previous lesson.

Kindergarten students from seven classrooms (n = 139) in three U.S. public schools participated in the study in the spring of 2019. In Technology Group 1 (n = 95), students used the Thermoscope, the Land of Bump, and the Particle Modeler. In Technology Group 2 (n = 44), they used the Thermoscope, the Land of Bump, and the Thermonator.

Figure 1. The Thermoscope, which works with or without temperature probes, is a simplified visualization of particle movement that makes the temperature differences between two materials visible.

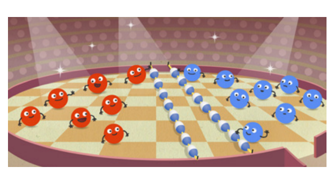


Figure 2. The Land of Bump is an animated story that introduces the motion of particles related to temperature and energy transfer in an accessible, interactive way.

Research with kindergarteners

Students worked in pairs on iPads, which allowed us to record their interactions with the apps as well as their conversations with one another through screencast recordings. Students were asked to draw and describe their understanding of the states of matter and phase change in individual project notebooks. We also interviewed students before and after they used the curriculum. Using this data, we compared changes in students' models of matter before and after completing the curriculum.

To analyze students' responses to interview questions we used a coding scheme based on both "top down" theoretically derived coding and "bottom up" or inductive coding based on an analysis of the data. Our coding scheme included component coding and coherence coding. The component-level coding focused on students' responses to individual question sequences. For example, states of matter questions were coded for students' descriptions of the a) composition of the material, b) arrangement of the material's smallest component pieces, and c) motion of the material's smallest component pieces.

After the component-level coding was completed, we conducted a coherence analysis to examine the degree of consistency and accuracy in students' use of particle models across states of matter and phase change question sequences.

Results

When we analyzed pre- and post-interview data, we found significant gains in the kindergarteners' ability to understand and use simple particle models to explain material phenomena. There were gains in component score for materiality (i.e., the ability of students to describe matter), states of matter, and phase change, as well as total score (Table 1).

TABLE 1	Minimum	Maximum	Mean	Standard Deviation
Pre Total	12	150	43.47	12.93
Post Total	12	150	79.05	24.53
Pre Materiality	2	15	5.59	2.16
Post Materiality	2	15	8.43	2.86
Pre States of Matter	3	96	20.97	9.91
Post States of Matter	3	96	46.79	19.40
Pre Phase Change	7	39	16.92	4.84
Post Phase Change	7	39	23.83	7.44

Table 1. Descriptive statistics of pre- and post-interview total and component scores. N=139.

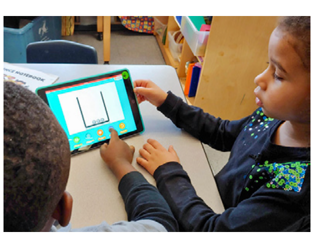


Figure 3. The Particle Modeler was designed for open-ended discovery. Students can drag particles and increase or decrease the temperature to observe their behavior.

Because we had designed the Particle Modeler and the Thermonator with different approaches based on different learning theories, we were also interested in the different technologies. While students in the two technology groups were equivalent in their initial knowledge of states of matter and phase changes, we found a very small but statistically significant effect on gains from pre- to post-interview total scores for technology. Students who used the Thermonator as part of Technology Group 2 had slightly higher total scores and component scores than those who used the Particle Modeler in Technology Group 1 (Table 2).

TABLE 2	Sample size	Mean	Standard Deviation
Pre Total Technology Group 1	95	42.03	12.89
Pre Total Technology Group 2	44	46.59	12.60
Post Total Technology Group 1	95	75.92	25.81
Post Total Technology Group 2	44	85.79	20.16

Table 2. Descriptive statistics of pre- and post-interview total and component scores by technology.

Students from all seven classrooms showed a shift towards more coherent particle model use for both states of matter and phase change phenomena. The frequency of macroscopic states of matter models decreased while the frequency of microscopic particle models increased. A similar pattern appears to hold for phase change models, although relative to the states of matter models, a greater proportion of the phase change models remained macroscopic. In particular, there appears to be a greater frequency of macroscopic or unclear models on the phase change prediction activities, suggesting that children start with less understanding of phase change phenomena than their understanding of states of matter.

Conclusion

Our study found that kindergarteners can learn to use simple particle models to explain the microscopic features and behaviors that characterize matter in solid, liquid, and gas states, and during phase changes. While this study is small and further research is needed to examine the role of different modeling practices, we believe it sheds important light on the value of introducing simple particle models to early elementary students.

Figure 4. The Thermonator allows students to add and arrange particles inside a virtual container and test normative and non-normative ideas about particle behavior.

LINKS

Sensing Science Through Modeling Matter
concord.org/sensing-science

Under the Hood:

Characterizing Systems Model Structure

By Dan Damelin

Dan Damelin
(ddamelin@concord.org)
is a senior scientist.

SageModeler, our free, web-based systems modeling tool developed with our partners at the CREATE for STEM Institute at Michigan State University, supports middle and high school students and teachers in understanding *systems and system modeling,* one of the seven crosscutting concepts in the Next Generation Science Standards. With new automated model tagging in SageModeler, we can now provide teachers more information about their students' understanding of systems.

To represent key components of a system, students first place nodes on the SageModeler canvas, then link the nodes together, specifying the semi-quantitative effect of one node on another to describe the web of relationships. They are thus able to build a computational model of their own conceptual model, then simulate the model and generate output to compare to the real-world behavior of that system.

When simulated, the model's structure and the set of relationships between model components dictate model behavior. To assist teachers and researchers in understanding certain characteristics of model structures, we developed algorithms that can automatically generate tags about a model's structure, including branching chains, linear chains, and feedback loops.

We started with an open-source graph introspection library called graphlib, which assumes graphs with structures consisting only of nodes and directed links (Figure 1). This graph is analogous to structures students make with a static equilibrium model in SageModeler (Figure 2).

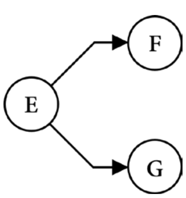


Figure 1. Example generic graph consisting of nodes and directed links. Source: https://github.com/dagrejs/graphlib/wiki/API-Reference#alg-components

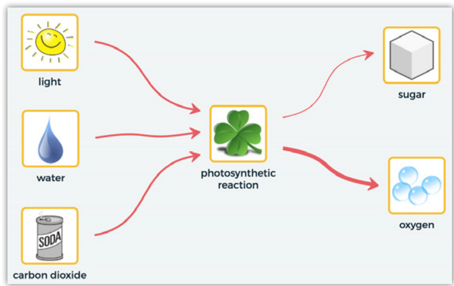


Figure 2. Example static equilibrium model of photosynthesis reaction.

We extended the graphlib library and created our own open-source npm package to analyze dynamic time-based models in SageModeler, which include special variables called "collectors" (traditionally called stocks) and flows. Figure 3a shows a flow between two collectors. Notice the single arrow from one collector *(carbon in fossil fuels)* through the valve (which controls the rate of flow) to the second collector *(CO2 in atmosphere).* The valve node describes how much to subtract from the source collector, which is then added to the destination collector.

In order to utilize features of standard graph analysis packages like graphlib, we first convert such structures behind the scenes to something that is semantically identical but represented using just nodes and directed arrows (Figure 3b). This allows us to automate the process of tagging particular characteristics in the models.

Students build and revise their models throughout our SageModeler curricular units. It would be daunting to analyze the thousands of student models created across our research cohort of teachers. Using the new model tagging algorithms, we can process large quantities of models and learn how model structures tend to evolve over the course of model revisions as well as how different scientific phenomena (e.g., predator and prey or chemical reactions) elicit different kinds of model structures.

Automated model tagging can also help teachers, allowing them to scan a list of student models and know something about the structure of each model. Did students use feedback loops to show exponential or logarithmic growth? Are models branched to show two outputs? Selecting and discussing various ways students model phenomena can propel class discussion.

Our goal is to support student understanding of modeling and systems. As machine learning continues to develop, model tags may go beyond simple structures and include key ideas represented (or missing from) models.

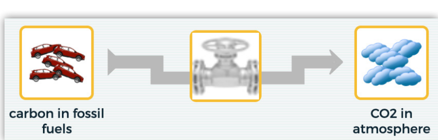


Figure 3a. Flow of carbon from fossil fuels to the atmosphere represented by using a transfer relationship.

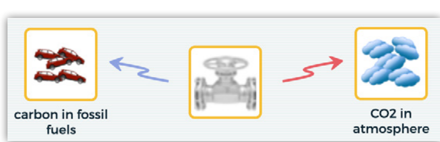


Figure 3b. Semantic representation of flow using nodes and directed links. The flow node has a negative impact on the source collector (blue arrow) and a positive impact on the destination collector (red arrow).

LINKS

SageModeler
sagemodeler.concord.org

graphlib
github.com/dagrejs/graphlib

topology tagger—github.com/concord-consortium/topology-tagger

Teacher Innovator Interview:
Julia Wilson

High school chemistry teacher
Portsmouth, Rhode Island


Julia Wilson tests how many drops of water, isopropyl alcohol, and acetone can fit on a penny during an InquirySpace professional development session.

Julia Wilson knows that her 10th grade chemistry students wonder, "Why am I sitting in this classroom counting atoms on a worksheet?" But she is optimistic that she can help them see the purpose in science. Online resources and a new approach to labs are helping expand her students' knowledge.

Julia, who teaches at Portsmouth High School, laughs, "It's very odd to be teaching where I grew up." Chemistry is the "central science" at the physics first school, tucked between physics in 9th grade and biology in 11th. According to Julia, chemistry is a gray area for students. "Kids understand physics in everyday life and they understand the biology of what they're living, breathing, experiencing." Her goal is to build bridges between those two subjects and help students see the importance of chemistry for explaining the world.

Julia is a research participant in our National Science Foundation-funded InquirySpace project, which offers tools and a curriculum sequence that guides students through conducting open-ended investigations. She admits, however, that when she tried to simply pop an InquirySpace activity into her class, it failed. Reflecting on the experience, she says, "I was trying to make inquiry look like a traditional lab." Julia now focuses on providing supports for her students while letting them figure things out on their own.

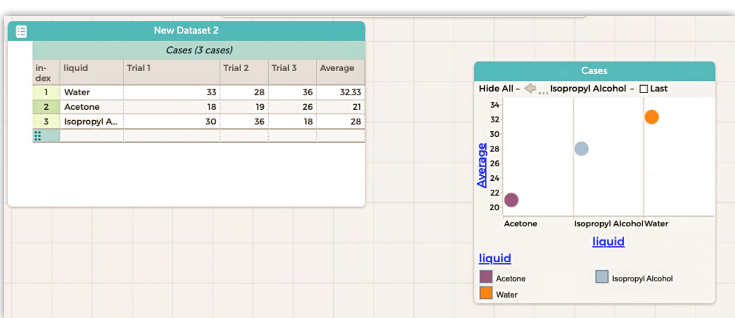
Passionate about teaching, Julia describes her path to the classroom as non-traditional. With several medical professionals in her family, she mapped a pre-med route, majoring in biology and working as an EMT at Bates College. But it was there that she realized she wanted to teach. Although she took some education classes, she did not have the traditional teaching credentials when she graduated, so she signed up with Teach for America. After a six-week summer preparation program, she was assigned to a school in St. Louis where she taught geometry for a semester then chemistry for a year and a half.

When she moved back to the East Coast to be closer to family, she was hired as the inaugural chemistry teacher at a new charter school in Lynn, Massachusetts, and had the opportunity to design the curriculum. But it was a physics class that changed Julia's classroom practices. She recalls, "There were so many opportunities to let kids play with things." For example, she gave students a ping pong ball and a meter stick and had them prove Newton's Second Law. "That shifted my thinking about what science could look like and what students are capable of doing."

In Portsmouth, Julia now strives to teach chemistry through that same lens. The InquirySpace curriculum units are aligned with this approach, each focusing on a driving question, such as "Why do I feel cold after getting soaked with water, even on a hot day?" When her students recently completed this unit, Julia was thrilled that they were able to apply the same scientific concepts to a new context when she asked them why ice keeps their soda cold.

She also credits the Next Generation Science Standards for shaping what science can look like, though she believes teachers need support in implementing inquiry while achieving content standards. In her own classroom, she aims for that "sweet spot between content and inquiry," working to transform otherwise cookie cutter labs into inquiry-based opportunities. For instance, she gives students vinegar and baking soda to investigate the conservation of mass. "Here are your materials," she tells them. "Have fun."

It's important to Julia that her students understand that they're learning not only science concepts, but a way of thinking. She says, "A big part of InquirySpace is getting students to see the why."



Sample student analysis of the InquirySpace evaporative cooling curriculum unit with experimental results displayed in CODAP.

The Concord Consortium
25 Love Lane, Concord, MA 01742

Upgrading CODAP and Strengthening Data Science Education Outreach

Data fluency—the ability to explore, interpret, visualize, and transform data into actionable next steps—is critical for both STEM learners and future STEM workers as well as for participation in modern society. Our Common Online Data Analysis Platform (CODAP) is a free, open-source, web-based learning environment aimed at bringing data fluency to everyone, especially learners in grades 5-14. Designed to support data exploration in a user-friendly, visual, and accessible manner, CODAP received a Research-Based Design Product Certification from Digital Promise for its basis in learning sciences research.

CODAP has been pursuing its mission steadily, and we're excited to note that this year marks its tenth birthday! In the decade since it was first introduced as the data analysis environment for our InquirySpace project, the CODAP codebase has grown to include a robust suite of features, scaffolds, and tools. Given CODAP's ongoing success, we are especially thrilled to announce that we are now re-engineering CODAP to modernize its underlying codebase and provide a fully open, flexible, and customizable online learning platform prepared to support data science education well into the future.

The re-architecture of CODAP will include technical upgrades to enable its long-term scalability and sustainability, and will integrate the Concord Consortium's latest work on collaborative tools, laying the groundwork for students to work together on larger and more challenging data exploration problems. As part of this new design work, we will engage K-14 audiences and youth in community-based citizen science projects, build out support resources and opportunities, and partner with dozens of organizations to ensure CODAP meets the needs of learners and programs that seek a free, open-source, and extensible data learning platform.

In addition, we are pleased to announce complementary work supporting data science education research. Thanks to new support from the Valhalla Foundation, we will work to build, support, and catalyze the data science education research community. We are excited to see these complementary initiatives yielding improved data-driven inquiry experiences, fostering and deepening research into different aspects of data education, and strengthening data fluency across a vibrant networked community of users and partners in both formal and informal science learning spaces within K-12 education.

Innovative Technology in Science Inquiry for Yup'ik Students

Two new projects—focused on grades 3-5 and 6-8, respectively—are supporting Yup'ik students in Hooper Bay, Alaska. We are engaging community partners, teachers, and students in adapting Concord Consortium STEM units by including local phenomena and Universal Design for Learning (UDL) features. The goal of both projects is to create dual-language, place-based STEM curricula based on Innovative Technology in Science Inquiry resources. The curricula will support the needs of English language learners while simultaneously providing a means for students to learn in their Yugtun Alaska Native language.

We will embed UDL and Yugtun language features such as a glossary in both English and Yugtun, videos with Elders, and speech-to-text and text-to-speech capabilities. The curricula will follow the multidisciplinary and multi-age 7E model of instruction at the core of Hooper Bay Charter School, with design features that are adaptable to other languages and cultures. Students will receive trade books and other print resources related to the STEM units for home use and sharing. Teachers will participate in remote professional development and learn to use data-based evidence about their students' learning through the Concord Consortium's reporting and dashboard systems to individualize instruction.

The projects aim to increase teachers' self-efficacy for multidisciplinary science instruction for dual language learners; provide training for critical evaluation of culturally relevant curriculum; and improve student interest and achievement in literacy, math, and science, as well as their reported sense of cultural connectedness.

Editor: **Chad Dorsey** | Managing Editor: **Cynthia McIntyre** | Design: **Jan Streitburger** | **concord.org**
The Concord Consortium | **Concord, MA** | **Emeryville, CA** | **978-405-3200** | **fax: 978-405-2076**
@Concord is published two times a year by The Concord Consortium, a nonprofit research and development organization dedicated to innovating and inspiring equitable, large-scale improvements in STEM teaching and learning through technology.
This material is based upon work supported by the Valhalla Foundation, OpenSea, the U.S. Department of Education (Grant Nos. T365C210008 and S356A210002) and the National Science Foundation under several grants (DRL-1949110, DRL-1841928, DRL-1812362, DRL-1900606, DRL-1621299, DRL-1842035, and DRL-1621301).