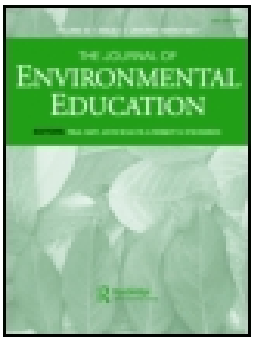

The Journal of Environmental Education

Routledge
Taylor & Francis Group

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/vjee20

Understanding and responding to challenges students face when engaging in carbon cycle pool-and-flux reasoning

Beth A. Covitt, Joyce M. Parker, Craig Kohn, May Lee, Qinyun Lin & Charles W. Anderson

To cite this article: Beth A. Covitt, Joyce M. Parker, Craig Kohn, May Lee, Qinyun Lin & Charles W. Anderson (2021): Understanding and responding to challenges students face when engaging in carbon cycle pool-and-flux reasoning, The Journal of Environmental Education, DOI: 10.1080/00958964.2020.1847882

To link to this article: https://doi.org/10.1080/00958964.2020.1847882

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

View supplementary material

Published online: 02 Mar 2021.

Submit your article to this journal

Article views: 175

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=vjee20

THE JOURNAL OF ENVIRONMENTAL EDUCATION
https://doi.org/10.1080/00958964.2020.1847882

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Understanding and responding to challenges students face when engaging in carbon cycle pool-and-flux reasoning

Beth A. Covitt[a], Joyce M. Parker[b], Craig Kohn[b], May Lee[b†], Qinyun Lin[b] and Charles W. Anderson[b]

[a]University of Montana, Missoula, Montana, USA; [b]Michigan State University, East Lansing, Michigan, USA

ABSTRACT

Carbon cycle pool-and-flux reasoning is a critical facet of climate literacy. This article begins with discussion of why this type of reasoning is both challenging and important. Results from two studies are reported. The first describes students' approaches to carbon cycle pool-and-flux reasoning. The second describes and reports results from an instructional intervention designed to scaffold secondary students' model-based pool-and-flux reasoning. Before instruction, most secondary students employed informal reasoning approaches including good versus bad and correlation heuristics to carbon cycle pool-and-flux problems. After instruction, the portion of students employing goal model-based pool-and-flux reasoning increased from 27 to 52 percent. This study builds on previous and current research to offer a promising instructional approach to scaffolding improvements in students' model-based pool-and-flux reasoning.

KEYWORDS

carbon cycle curriculum; carbon cycle pool-and-flux reasoning; climate literacy

Introduction

As environmental educators well know, climate change is one of the most urgent socioenvironmental problems facing society today. Environmental education's focus on this problem in North America is evident in the North American Association for Environmental Education's *Guidelines for Excellence: K-12 Environmental Education* (North American Association for Environmental Education (NAAEE), 2019). For example, middle school guideline 2.1.A indicates learners should be able to, "…provide an evidence-based explanation of how humans have changed Earth's atmospheric gases during the last two centuries and the consequences of those changes" (NAAEE, 2019, p. 47). The *Guidelines* also identify "systems and systems thinking" as the first essential underpinning of environmental education, noting that, "[s]ystems thinking helps make sense of a large and complex world" (NAAEE, 2019, p. 12).

Earth's complex and changing climate is a prime context in which systems thinking can help people make sense of and respond to a socioenvironmental issue. The affordances of systems thinking are also evident in other expectations in the *Guidelines*. For example, high school guideline 3.1.C suggests individuals should be able to, "[c]ritique proposed solutions using gauges such as likely impacts on society or the environment, and likely effectiveness of solving the issue" (NAAEE, 2019, p. 81). In order to achieve this goal, students need to access and use climate system thinking (e.g., by evaluating explanations, predictions, and arguments that draw on understanding of invisible dynamic processes in the system that unfold across different spatial and temporal scales) (Hmelo-Silver et al., 2007; Hogan & Weathers, 2003).

CONTACT Beth A. Covitt beth.covitt@umontana.edu spectrUM Discovery Area, University of Montana, Missoula, MT, USA.
†Lee was affiliated with Michigan State University at the time this paper was written and is currently affiliated with Rijksuniversiteit Groningen, Groningen, Netherlands.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

Growing acknowledgment of the importance of climate and climate change is also evident in recent shifts in science education standards in the United States. While climate change and global carbon cycling were largely absent from the *National Science Education Standards* released in 1996 (National Research Council), they are prominent in the more recent *Next Generation Science Standards (NGSS)* (NGSS Lead States, 2013). Relevant performance expectations, for example, address: clarifying evidence of factors that have led to a rise in global temperatures (MS-ESS3-5), developing a quantitative model to describe global carbon cycling (HS-ESS2-6), and using evidence from climate models to forecast climate change (HS-ESS3-5).

This paper focuses on one key element in students' understanding of global climate change, namely carbon cycle pool-and-flux reasoning. The *NGSS* call on students in the United States to "[d]evelop a quantitative model to describe the cycling of carbon among the hydrosphere, atmosphere, geosphere, and biosphere" (NGSS Lead States, 2013, HS-ESS2-6). This is a particularly challenging standard to achieve because, as we will explain, most people make sense of carbon cycle data in problematic ways that lead to the erroneous conclusion that addressing anthropogenic climate change will be much easier to accomplish than it actually will be.

We do not suggest that teaching students pool-and-flux reasoning addresses all elements of environmental education associated with goals such as environmental literacy, environmentally responsible behavior, or action competence (e.g., Bishop & Scott, 1998; Coyle, 2005; Hsu, 2004; Kollmuss & Agyeman, 2002; McBeth & Volk, 2009; Mogensen & Schnack, 2010). Understanding pool-and-flux reasoning may not directly impact peoples' climate-relevant personal and societal decisions or behaviors. However, we argue that carbon cycle pool-and-flux reasoning represents a "necessary but not sufficient" accomplishment for informed participation in societal decision-making related to climate change.

This argument is consistent with perspectives in environmental education. For example, in various models and approaches to environmental education, knowledge is a consistently included construct (e.g., Bishop & Scott, 1998; Coyle, 2005; Hsu, 2004; Kollmuss & Agyeman, 2002; McBeth & Volk, 2009; Mogensen & Schnack, 2010). And, while carbon cycle pool-and-flux reasoning represents an aspect of knowledge necessary for environmental literacy, it is much more than just a fact to be learned. Rather, this type of reasoning involves employing sophisticated sense making to coordinate longitudinal, global scale data with a model-based, mechanistic understanding of Earth's complex carbon cycling system. Because the potential effectiveness of different climate actions is commonly evaluated based on predicted impacts on atmospheric carbon levels, carbon cycle pool-and-flux reasoning is essential for informed engagement with responses to climate change.

In the context of climate change, NAAEE's guideline 3.1.C involves critiquing the likely impacts of different goals for emissions reductions on atmospheric carbon dioxide concentrations, which subsequently impact global temperatures and other climate indicators (IPCC, 2018). People who are able to use carbon cycle pool-and-flux reasoning may not make decisions consistent with supporting effective means of addressing global climate change, but those without access to carbon cycle pool-and-flux reasoning cannot; they lack the capacity to understand the likely effects of different choices, and thus to make evidence-informed decisions about personal and policy-related climate issues. This is particularly concerning in today's society in which people have reason to be skeptical about arguments from various sources concerning socioenvironmental issues and solutions (Barzilai & Chinn, 2020; Feinstein & Waddington, 2020; Iyengar & Massey, 2019; Stubenvoll & Marquart, 2019).

Further, while human understanding of climate science and arguments concerning responses to climate change continue to change over time, some basic ideas and models (including but not limited to carbon cycle pool-and-flux reasoning) represent fundamental aspects of preparation for future learning in this domain that will remain useful over time. After they complete their schooling, individuals need to be able to continue to learn about socioenvironmental issues (e.g., through reading news articles in the media) as both the circumstances of and our understanding of those issues change (Bransford & Schwartz, 1999; Zeidler et al., 2009; Zeidler & Kahn, 2014).

Preparation for future learning does not mean knowing everything - it means being able to judge and make sense of arguments about changing and emerging issues as need arises. Pool-and-flux reasoning positions people to critique alternative goals and strategies for emissions reductions now and in the future as aspects of our global socioenvironmental system such as levels of atmospheric carbon dioxide (CO_2); rates of emissions; available technologies and understandings of how they work; and circumstances of social, political, economic, and justice contexts change over time. Thus pool-and-flux reasoning is one essential, flexible facet of systems thinking that individuals need in order to be prepared for current and future participation as informed environmental decision-makers.

In this paper, we draw on research in the literature and our own design-based research to discuss (1) why carbon cycle pool-and-flux reasoning is crucial to addressing climate change, (2) why this type of reasoning is so challenging, (3) the more and less sophisticated ways middle and high school students reason about global carbon pools and fluxes, and (4) a promising instructional approach to improving secondary students' carbon cycle pool-and-flux reasoning. The evidence we present concerning students' ways of thinking and the beneficial effects of an instructional experience both draw from a large-scale design-based research project aimed at teaching students to trace matter through carbon transforming processes at multiple scales from atomic-molecular to global (Anderson et al., 2018; Cobb et al., 2003).

Why quantitative carbon cycle pool-and-flux reasoning is critical for addressing climate change

Figure 1 presents an iconic image that is frequently used as evidence that CO_2 concentrations in the atmosphere are increasing. Known as the Keeling Curve, it documents the increasing concentration of CO_2 in the atmosphere at Mauna Loa, Hawaii between 1958 and the present. Most students we have interviewed or who have completed written assessments for our project believe that increasing CO_2 concentrations are bad and that we should do something about them. The questions of what to do and how much difference it will make, however, are more complicated.

Most students correctly attribute the upward trend in the Keeling Curve to human activities that use fossil fuels. On the surface, this connection seems straightforward. For example, one can compare time series graphs showing the Keeling Curve and the fossil fuel flux of carbon into the atmosphere (Figure 2). Eyeballing the trends in these graphs, they look similar. If we look at the period from 1958 through 2010, we see that in both cases, trends are going up steeply over time. This leads to a seemingly logical conclusion: If we can reduce CO_2 emissions (i.e., get the lines in Figure 2 to start going down), then CO_2 concentration (the line in Figure 1) will start going down too. Unfortunately, the relationship between CO_2 emissions and CO_2 concentration is not that simple. Figure 3 shows why.

Global carbon cycling involves the multiple processes (photosynthesis, cellular respiration, combustion, etc.) that move carbon among connected pools in the geosphere, hydrosphere, atmosphere, and biosphere. When these systems have balanced carbon fluxes, the sizes of carbon pools remain the same over time. When fluxes are imbalanced, pool sizes change over time. What's more, it only takes a small imbalance in fluxes to make a large change in a pool's size over time. Figure 3 shows that the flux from burning fossil fuels (10 GtC/year) is far smaller than most other fluxes into and out of the atmosphere, but it is unbalanced. We can calculate the overall carbon flux using the Figure 3 model by summing the annual fluxes into the atmosphere (208 GtC/year), summing the fluxes out of the atmosphere (200 GtC/year), and comparing the two; this yields a net flux of 8 GtC/year into the atmosphere.

Pool-and-flux reasoning shows us that simply reducing emissions will not reduce or even stabilize the atmospheric carbon pool. With reduced emissions, the atmospheric CO_2 concentration will continue to grow at a slower rate. This is the crux of why pool-and-flux reasoning is so important. Stabilizing the concentration of CO_2 in the atmosphere will require not just reducing emissions, but reducing them to an extent that will sustain the global carbon cycling system at or near a balanced-flux state indefinitely (i.e., fossil fuel emissions will need to be close to zero or else other actions will need to be taken to move

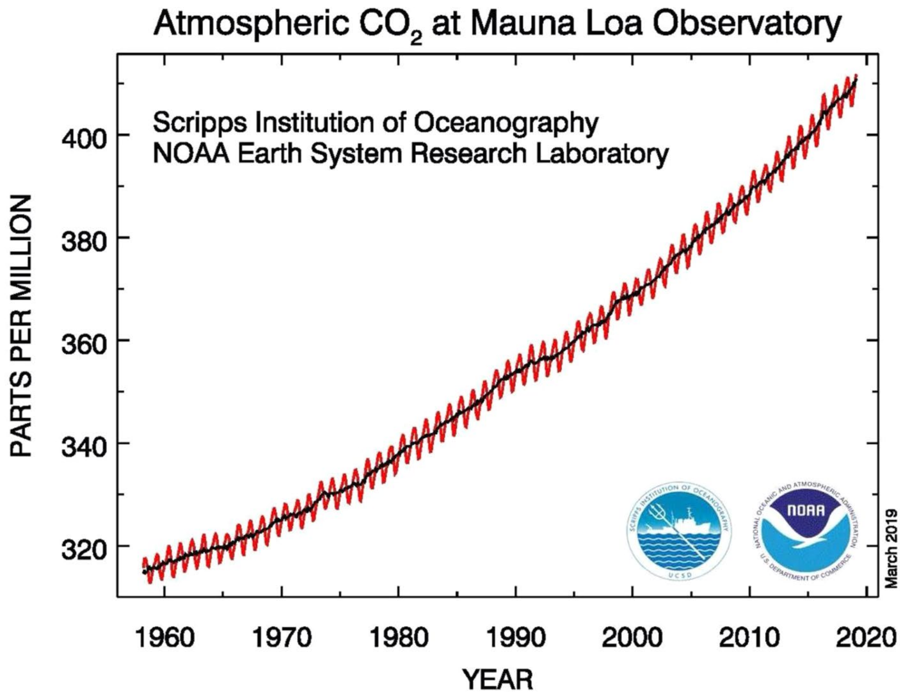

Figure 1. Keeling Curve (record of atmospheric CO_2 concentration at Mauna Loa) (National Oceanic & Atmospheric Administration, 2019).

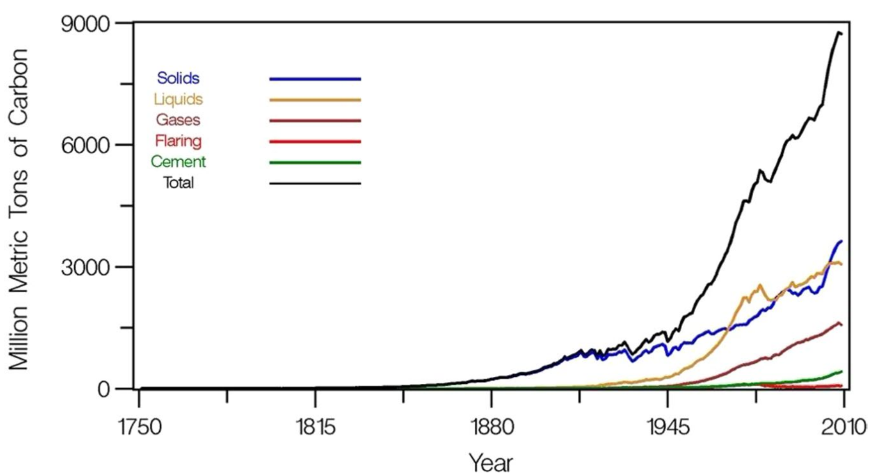

Figure 2. Global fossil fuel carbon emissions (Boden et al., 2015).

more CO_2 from the atmosphere back to terrestrial and ocean systems to balance the fossil fuel flux into the atmosphere).

This is why the *NGSS* emphasize quantitative modeling of global carbon cycling as a key goal. Students (and people in general) need to recognize the actual problem we are facing with respect to addressing climate change in order to make informed decisions concerning the changes that are required to avert the most catastrophic projections for climate change. In the next section, we discuss research from related fields that explains why people, spanning from middle school students through science experts, have so much trouble with pool-and-flux reasoning.

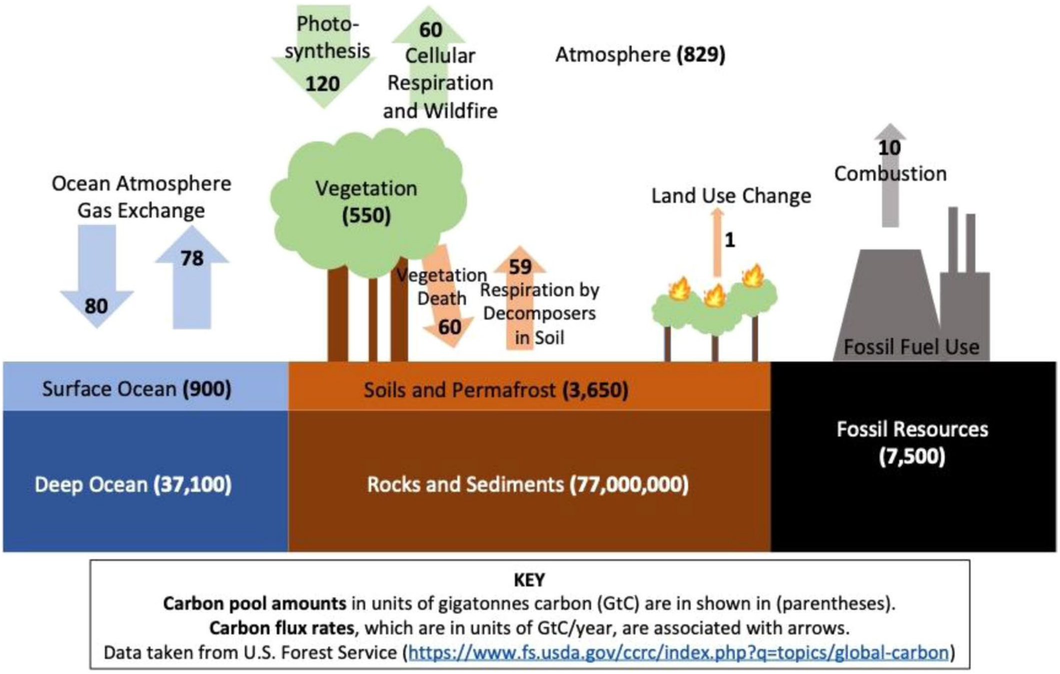

Figure 3. Global carbon cycle model.

The challenge of carbon cycle pool-and-flux reasoning

Studies of pool-and-flux reasoning in different contexts

Studies conducted over the past several decades provide an illustration of the kinds of trouble people encounter when they reason about pool-and-flux problems. This research has been conducted with a variety of participants, though often with university undergraduate and graduate students. The work has been conducted using a range of pool-and-flux problems including water in a bathtub, oil in a tank, people in a building, air in a balloon, dollars of national debt, distance between cars, and CO_2 in the atmosphere (Cronin et al., 2009; Dutt & Gonzalez, 2012; Guy et al., 2013; Moxnes & Saysel, 2009; Reichert et al., 2014, 2015; Sterman & Sweeney, 2007).

Findings have been consistent. People, including those with technological expertise and training, are generally poor pool-and-flux reasoners. Instead of recognizing fluxes as rates of change and pools as amounts of materials, people often oversimplify these problems and view fluxes and pools as having a simple linear relationship. This tendency has been labeled both "correlation heuristic" and "pattern matching" (Cronin et al., 2009; Dutt & Gonzalez, 2012; Moxnes & Saysel, 2009; Sterman & Sweeney, 2007). Basically, when dealing with pool-and-flux problems, individuals will often assume that if a flux has a positive trend then a pool will have a positive trend, and vice versa. As noted by systems scientists, this simplifying heuristic can lead individuals to grossly underestimate how much we will have to reduce CO_2 emissions to stabilize or reduce the atmospheric carbon pool (Sterman & Sweeney, 2007).

Other studies provide evidence of additional informal reasoning approaches, aside from the correlation heuristic. For example, Sweeney and Sterman (2007) found that middle school students sometimes consider inflow but not outflow in pool-and-flux problems. Niebert and Gropengiesser (2013) analyzed metaphors that scientists and high school students use to understand climate change; they found that students viewed anthropogenic CO_2 as "bad" because it is made by people rather than being natural. Similarly, in our research (Covitt & Anderson, 2018), we have found that high school students often use informal approaches to making judgments and predictions about phenomena related to climate and climate change. These include, for example, covering law approaches (Braaten & Windschitl, 2011), which describe things such as pollution and climate change just going together without explaining underlying mechanisms and qualifications (e.g., how does pollution impact climate, which types of pollutants, from which sources, and to what extent). We have also observed fast thinking heuristics (Kahneman, 2011) such as eyeballing graphs and simply extending patterns and trends in graphs to make predictions for future CO_2 levels.

Informal approaches to judgments served our prehistoric ancestors well and have become prevalent among the human population (Gigerenzer & Todd, 1999; Payne et al., 1993). In many quotidian contexts, quick and decisive approaches to making judgments are desirable (Kahneman, 2011). Without quick thinking, people would get bogged down in every little decision (e.g., what should I have for breakfast today?) and find it difficult to complete larger and more significant tasks. In most of our everyday experiences and contexts, the correlation heuristic is an effective approach. Sterman and Sweeney (2007) provide a few examples such as kettle whistling correlates with water boiling, and eating certain mushrooms correlates with becoming ill. Unfortunately, quick thinking approaches like the correlation heuristic are insufficient for the task of making informed critiques of proposed solutions to climate change.

Studies of instructional interventions focusing on pool-and-flux reasoning

Van Dooren et al. (2007) found that oversimplified correlational or linear reasoning was highly prevalent among sixth graders, and that it was reinforced by the common use of word problems in school that prompt students to identify linear relationships. Van Dooren and colleagues also found that interventions that required students to go beyond verbal and text writing performances (e.g., to undertake drawing or manipulating objects) helped students to avoid misapplied linear reasoning. However, on a subsequent posttest, students in all conditions returned to linear reasoning strategies. The interventions helped disrupt linear reasoning about a particular example but did not change students' overall tendency to apply linear reasoning.

As with Van Dooren and colleagues' (2007) interventions seeking to disrupt linear reasoning, attempts to help individuals achieve more sophisticated pool-and-flux reasoning have shown that some approaches can have significant impacts, but also that students often revert to applying the correlation heuristic. Some approaches that have been shown to at least modestly improve pool-and-flux reasoning include providing feedback (Cronin et al., 2009), interacting with pool-and-flux simulations (Dutt & Gonzalez, 2012), employing analogies (Guy et al., 2013; Moxnes & Saysel, 2009; Reichert et al., 2015), introducing a cognitive conflict (Moxnes & Saysel, 2009; Reichert et al., 2015), and employing cognitive flexibility principles (Reichert et al., 2015). Other approaches have demonstrated mixed results. For example, Guy et al. (2013) found that employing graphs in problems can lead to relatively worse reasoning outcomes. Cronin and colleagues (2009), however, found that employing graphs did not negatively influence reasoning. Approaches including simplifying problems and providing motivational incentives have also been shown to be ineffective in some experiments (Cronin et al., 2009).

Studies of instruction about climate change

To date, few studies have examined or documented changes in pool-and-flux reasoning among secondary students as a result of learning experiences. Thus, little evidence has been presented to suggest that secondary students can learn to successfully use pool-and-flux reasoning, especially in the context of the carbon cycle. A search of both research and practice literature suggested that much of the work at the secondary level has focused on either describing students' understanding of climate change without examining learning (e.g., Chang & Pascua, 2016; Düsing et al., 2019; Özdem et al., 2014; Shepardson et al., 2009, 2011, 2014; You et al., 2018) or describing climate change and/or carbon cycle instruction without addressing or examining learning related to pool-and-flux reasoning (e.g., Bofferding & Kloser, 2015; Pruneau et al., 2003).

Some curricular materials we found in the literature focused on the pathways carbon moves through without requiring students to engage in quantitative pool-and-flux reasoning, which is required for making sense of changes in pool sizes over time (e.g., Hoover, 2019; Peel et al., 2017). One study examined secondary students' reasoning relevant to pools and fluxes using a qualitative approach that provided useful insights but did not provide a more generalizable examination of whether and how educational experiences might support significant learning in this domain among secondary students (Niebert & Gropengiesser, 2013). Another study found only 20% of students achieved qualitative model-based carbon cycle reasoning as a result of instruction (pool-and-flux reasoning was not explicitly addressed in the study) (Zangori et al., 2017). In summary, research to date has not produced evidence of or from effective approaches for scaffolding secondary students' learning of carbon cycle pool-and-flux reasoning.

One other issue to note with regard to educational implications is that carbon cycling is a particularly complex pool-and-flux reasoning problem when compared with many other examples (e.g., pools of national debt and fluxes of revenue and spending, pools of money in a savings account and fluxes of deposits and withdrawals, pools of water in a bathtub and fluxes of water entering and exiting). While the carbon cycle comprises multiple pools and fluxes moving carbon through a complex system, in all the examples above, there is only one pool and two fluxes (one in and one out).

Summary

Past research on pool-and-flux reasoning surfaces several key points. First, pool-and-flux reasoning has been recognized as an important learning target in several different fields. Second, difficulty with this type of reasoning tends to arise when people rely on simplified heuristics that produce quick but sometimes inaccurate conclusions. Third, teaching students when and how to use pool-and-flux reasoning is hard. And finally, research on teaching climate change has generally not recognized the important role of pool-and-flux reasoning or documented successful strategies for teaching it.

Background and research questions

Learning progressions and design research

The two studies reported in this article represent work situated in learning progressions theory (Duncan & Rivet, 2013) and the methodological approach of design-based research (Cobb et al., 2003; Collins et al., 2004). These theoretical and methodological lenses are leveraged to examine and respond to the educational challenge of teaching pool-and-flux reasoning with secondary students.

"Learning progressions are descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time" (National Research Council, 2007, p. 214). Development of empirically grounded learning progressions has been shown to hold promise for advancing and informing multiple aspects of research-based education efforts including in areas of formative assessment, measurement of student learning, creation of responsive curriculum materials, and design of effective teacher professional development (Gotwals, 2012).

Our learning progressions research uses grounded evidence from students' own performances to characterize students' ways of talking, thinking, and writing as they make sense of the world as they experience it (Gee, 1991). Knowing how students make sense of the world provides a critical lens for designing learning experiences that are responsive to students' ways of reasoning and that can support students in developing more sophisticated knowledge and practice over time.

Because we focus on just a few assessment items in this article, the research and evidence presented here does not represent a complete learning progression on its own. However, this study does build on and fit within the body of our previous learning progressions work that describes less and more formal ways that students make sense of environmental phenomena and systems (Covitt & Anderson, 2018; Gunckel et al., 2012; Mohan et al., 2009). While the results of this study are consistent with the methods and findings of our previous work, they are also unique; we have never published data or results specifically addressing students' pool-and-flux reasoning before.

The methodological approach of design-based research aims to "blend empirical educational research with theory-driven design of learning environments … [to understand] … how, when and why educational innovations work in practice" (Design-Based Research Collective, 2003, p. 5). In collaboration with schools and teachers that has extended for over a decade, we have used a design-based research approach to develop, test, and refine learning progressions and learning progression-informed instructional approaches addressing environmental science literacy (Anderson et al., 2018).

Research context

The carbon TIME project

For over a decade, the *Carbon TIME* project has enacted a design-based research partnership aimed at studying, testing, and refining a learning progression-based approach to teaching carbon cycling in the United States at the middle and high school levels (Anderson et al., 2018). The *Carbon TIME* curriculum comprises six instructional units: *Systems & Scale*, *Animals*, *Plants*, *Decomposers*, *Ecosystems*, and *Human Energy Systems* (all *Carbon TIME* materials are freely available at carbontime.bscs.org).

In the sequence of *Carbon TIME* units, students learn to trace matter and energy through processes such as photosynthesis, biosynthesis, cellular respiration, and combustion at multiple scales—from atomic molecular through global. In the curriculum, carbon cycle pool-and-flux reasoning comes at the end—in the *Ecosystems* and *Human Energy Systems* units. Thus, before *Carbon TIME* students encounter the challenge of global carbon cycle pool-and-flux reasoning, they have had experience with tracing carbon through smaller systems including animals, plants, engines, and ecosystems.

Carbon TIME teachers participated in a professional development (PD) course of study that was embedded in a local professional network (i.e., professional learning community). The course of study, which involved 75 hours of participation over two years, included both face-to-face and online PD experiences with activities including but not limited to experiencing, analyzing, and critiquing units; enacting units and reflecting on instruction; analyzing and responding to student performances; and collaboratively working on problems of environmental science literacy instruction. While tracing matter and energy through systems was emphasized throughout the PD course of study, very little PD time focused specifically on global pool-and-flux reasoning.

Carbon TIME has addressed the *NGSS* performance expectation for carbon cycle pool-and-flux reasoning (HS-ESS2-6) through conducting research on students' carbon cycle reasoning and through instructional design and implementation based on our own and others' research. In this paper, we present results from two studies. The first was a pilot study that analyzed patterns in students' responses to a pool-and-flux problem. The results of the first study contributed to revisions of the *Human Energy Systems* unit. The second study examined the impact of that unit on students' carbon cycle pool-and-flux reasoning.

Research questions

Study One (Pilot) Research Question

What are different (and more and less sophisticated) ways students reason about how a 50% reduction in combustion of fossil fuels would affect future atmospheric CO_2 concentrations?

Study Two Research Questions

1. How does viewing a diagrammatic carbon cycle model influence students' pool-and-flux explanation and prediction performances?
2. How does engaging in an instructional unit that scaffolds carbon cycle pool-and-flux reasoning affect students' explanation and prediction performances?

Study One: Learning progression research on students' predictions and explanations

Methods

Context and data sources

In Study One we drew on a convenience sample of students of different ages and levels of experience with the purpose of eliciting and describing a spectrum of approaches to carbon cycle pool-and-flux reasoning. Interviews were conducted with 25 undergraduate students (mostly non-science majors) and 5 graduate medical students. Written responses were collected from 93 high school students including 42 ninth grade students and 51 twelfth grade students. Some, but not all of the high school students had previously completed *Carbon TIME* units. All data were collected in a Midwest state.

In both interviews and written responses, we asked students to evaluate different predictions for how a 50% reduction in combustion of fossil fuels would affect atmospheric CO_2 concentrations over time. The question (Figure 4) depicts part of the Keeling curve with dashed lines showing five predictions for atmospheric CO_2 concentration from 2016 to 2065. The students were asked to agree with one of five predictions for future CO_2 levels and to explain their choices.

We asked students to choose a prediction and explain their choice both before and after they saw Figure 5, which is a quantitative carbon cycling model representation from the Intergovernmental Panel on Climate Change (IPCC) (2001). This model is similar to Figure 3 but uses older data. The carbon cycle pool-and-flux reasoning required in both models is the same. The rationale for asking students to respond both before and after viewing the carbon cycling model representation stems from the use of this type of representation in climate change education and media sources aimed at student audiences, for example, in educational materials presented by The Globe Program (retrieved May 9, 2020) and Project Learning Tree (retrieved May 9, 2020). These programs present the diagrams with minimal consideration of challenges associated with pool-and-flux reasoning, suggesting that the authors expect students to be able to interpret and use the diagrams without much additional support.

If students have difficulty using carbon cycle pool-and-flux diagrams as reasoning tools, as we suspected they likely would, this would suggest that educators who make use of these diagrams in lessons and other materials will need to be aware that in many cases, students may not take away from such lessons the learning outcomes (i.e., understanding how the carbon cycle imbalance affects CO_2 concentrations in the atmosphere over time) that educators hope students will achieve. More directed and intensive learning experiences that go beyond just showing and/or explaining the models to students would be needed.

Analysis

We analyzed students' prediction selections and explanations from interview and written responses using established learning progression research methods (Black et al., 2011; National Research Council, 2006). These methods involve iterative cycles of assessment development, implementation, and analysis with combinations of deductive and inductive coding aimed at articulating empirically grounded levels or categories of ways of reasoning about a topic. Consistent with learning progression research approaches, the reasoning categories presented in this study were developed with reference to both emergent themes arising from this study's data and past research including both our own (e.g., Covitt & Anderson, 2018; Mohan et al., 2009; Parker et al., 2015) and others' (e.g., Cronin et al., 2009).

Our analyses were conducted in several cycles beginning with implementation and analysis of interviews in 2015 and 2016 followed by implementation and analysis of written responses in 2017. Across these assessment implementations, we found that students generally responded to pool-and-flux reasoning questions in one of three ways: *pool-and-flux model-based reasoning*, *correlation heuristic reasoning*, and *good versus bad heuristic reasoning*. These categories are described in the Results section.

After categories were developed using first the interview data and then samples from the written response data, two authors separately coded 80 of the remaining written responses (including responses from both before and after viewing the IPCC model) to establish interrater reliability. Weighted Cohen's Kappa for interrater reliability was 0.65, which is considered substantial (Landis & Koch, 1977). The authors compared and discussed codes, came to consensus for disparate codes, refined the coding exemplar, and one author coded the remaining written responses.

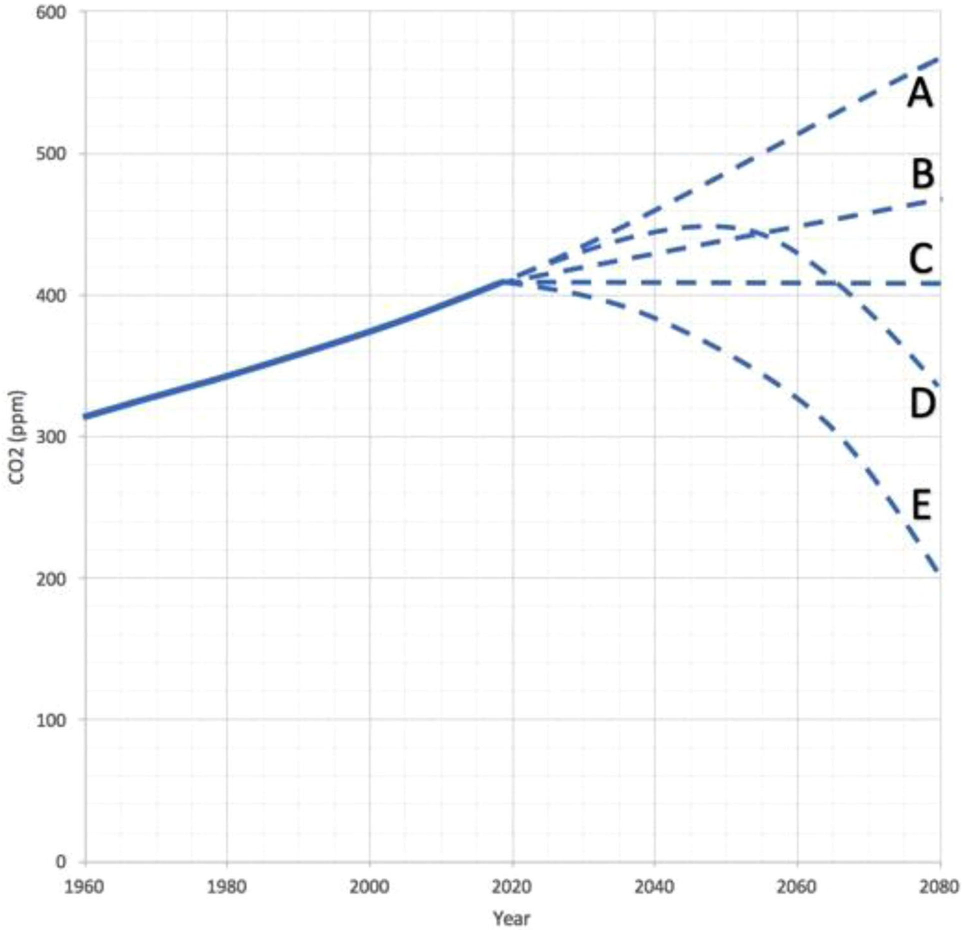

Figure 4. Atmospheric CO_2 prediction question, part 1.
The solid line in the graph shows how carbon dioxide (CO_2) concentrations in the atmosphere changed between 1960 and 2016. If the world were suddenly able to cut its use of fossil fuels in half tomorrow and maintain that low level of use, what would be the effect on the concentration of atmospheric CO_2? Which line best describes what you think would happen to CO_2 levels: A B C D E
Explain your answer.

Results

The students' predictions and explanations fit into three general patterns: pool-and-flux model-based reasoning, correlation heuristics, and good versus bad heuristics. We describe each pattern below, then conclude this section with a discussion of implications for instruction.

Pool-and-flux model-based reasoning

The most sophisticated student responses used the arrows in the IPCC model to calculate a net flux if fossil fuel use were cut in half while the other fluxes were unchanged. These students chose C or D and used the numbers in the model to calculate the net flux of CO_2 into the atmosphere given a starting level of emissions of 3.15 GtC per year. Calculating in conjunction with the other fluxes shown in the model, if emissions were to be cut in half, the net flux would be about 0 (or 0.05 GtC per year out of the atmosphere if students included multiple digits in their calculations[1]). In an example response representing this type of reasoning, the student wrote, "*Cutting CO_2 from fossil fuels in half would mean 3.15 from processes in the atmosphere. The ocean takes up −2 Gt (88 – 90), land use takes up −0.2 (0.7 – 1.9), and −1 Gt from GGP (119-120). This shows that the atmosphere carbon levels SHOULD go down 0.05 Gt a year.*"

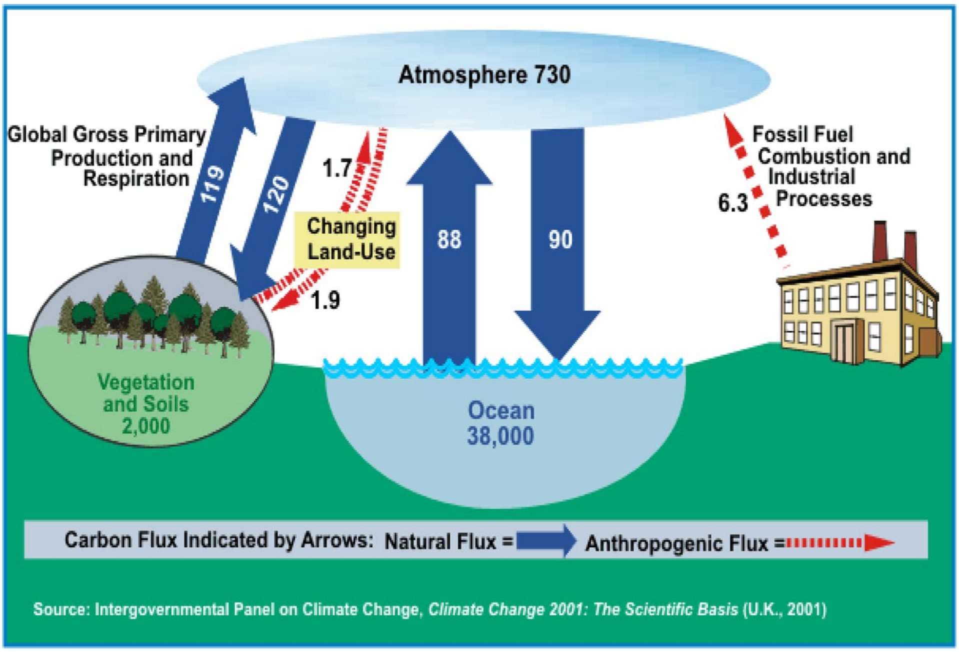

Figure 5. Global carbon cycle model.
The figure to the right shows part of the global carbon cycle. It shows some of the different places or reservoirs where carbon is found on the planet and the amount of carbon in gigatonnes (Gt) in each of those places. The arrows show the number of gigatonnes of carbon that move in and out of the atmosphere every year.

An even more sophisticated level of understanding (which we did not observe in responses from students) would involve choosing response B and explaining that with a reduction in emissions, other fluxes in the model would change as well. For example, the flux arrow from the atmosphere to the ocean would likely decrease due to a negative feedback loop, resulting in the CO_2 concentration in the atmosphere continuing to rise at a less rapid rate over time.

Some students agreed with B and used pool-and-flux reasoning to make a reasonable prediction without doing a calculation. This type of response was evident both before and after the students viewed the IPCC model. These students recognized that changing a flux changes the *slope* of the line on the graph rather than the value on the Y axis, which represents CO_2 concentration. These students explained that if we cut fossil fuel use in half, we would still be using fossil fuels—just not as much. Therefore, atmospheric CO_2 concentrations would continue to rise, but at a slower rate. This student's written response is representative of this type of model-based reasoning, "*We'd still be producing more CO_2 than what gets taken out. So only the rate would slow*." While it does not include a calculation, this students' response still represents model-based pool-and-flux reasoning that recognizes the distinction between amount of atmospheric CO_2 and rate of CO_2 flux into the atmosphere.

Correlation heuristic reasoning

Other students chose D or E and reasoned about pools and fluxes in quantitative but inaccurate or incomplete ways. These students often applied the correlation heuristic, conflating changes in flux (slope of the graphed line) with changes in pool size (value on the Y-axis). The following written response reflects this type of thinking, "*D because fossil fuels help to produce CO_2 so if we cut it in half it would decrease*." Note how this student used "it" twice in the same sentence, perhaps without recognizing that each "it" had a different meaning:

…*if we cut it* (CO_2 emissions—the flux arrow) *in half,*

> …*it* (CO_2 concentration—a measure of the size of the atmospheric CO_2 pool) *would decrease.*

This approach often led to spurious quantitative reasoning, such as when another student conflated a change in flux with a change in pool size, saying, "*I guess it would definitely be down here, like 200. … Because we're at 400 right now, so in half.*"

Good versus bad heuristic reasoning

Other students reasoned in ways that ignored the numbers from the graph and the model. They used an informal frame to explain their ideas about what would happen. These students did not attend to quantitative pools, fluxes, or concentrations at all. Instead, they described things that happen to the environment as good (e.g., less pollution) or bad (e.g., using fossil fuels). For instance, some students chose D or E, connecting good actions (e.g., cutting fossil fuel use) with good outcomes without referencing carbon cycle mechanisms: "*If it's cuts down and maintain a low level use, the air will clear up and it will be good for animals and humans to breath clean air.*" Some students chose A based on connecting bad actions to bad outcomes. For example, one student wrote, "*[b]ecause I think we've reached a point where we've done too much damage to earth, personally. And I don't think we can come back from that.*"

Implications for instruction

Over two-thirds of high school students provided responses consistent with the good versus bad heuristic or the correlation heuristic both before and after they saw the IPCC model. Generally speaking, students who provided good versus bad and correlation heuristic type responses before seeing the model did not subsequently use the IPCC model to make pool-and-flux model-based predictions. This suggests, as suspected, that seeing a quantitative pool-and-flux model is not particularly helpful for most students who rely on good versus bad or correlation heuristics.

In general, students who demonstrated capacity to engage in model-based pool-and-flux reasoning were successful with the following three practices:

1. Reasoning using mechanisms (i.e., fluxes between pools) rather than good or bad factors (e.g., pollution) that influence CO_2 concentrations.
2. Recognizing and distinguishing between carbon pools and carbon fluxes.
3. Reasoning quantitatively (which does not necessarily require calculations) about multiple fluxes.

With respect to preparation for making informed critiques of solutions as advocated by high school guideline 3.1.C (NAAEE, 2019), we observe a large and meaningful difference in the preparedness of students who engage in "good versus bad" and/or "correlation heuristic" types of reasoning compared with students who engage in "pool-and-flux" reasoning. We concluded that scaffolding pool-and-flux reasoning about global carbon cycling should be a high priority for the *Human Energy Systems* unit.

Study Two: Design-based research on teaching pool-and-flux reasoning

Methods

Context

Study Two examined students' performances before and after they studied the *Human Energy Systems* unit (carbontime.bscs.org/human-energy-systems). This unit builds on findings from Study One as well as knowledge and practices that students develop in the previous *Carbon TIME* units. The first five units support students in developing a repertoire of explanations and evidence-based arguments for tracing matter and energy in combustion and life science contexts at the atomic-molecular, macroscopic, and ecosystem scales. While this repertoire provides a critical precursor, it is not sufficient for employing model-based, global pool-and-flux reasoning. Therefore, the *Human Energy Systems* unit was designed to scaffold the important practices needed for pool-and-flux reasoning identified in Study One.

The *Human Energy Systems* unit is divided into two phases. The first phase comprises three lessons in which students look at related time series patterns in data about Earth systems: global temperatures, changes in sea level, Arctic sea ice, and atmospheric CO_2 concentrations. Students study the relationships among these patterns, eventually concluding that through the greenhouse effect, *CO_2 is the driver*; changes in CO_2 concentrations are driving the changes in the other variables.

This leads to a key question that students subsequently answer in Phase 2 (Lesson 4)—W*hat drives the driver (i.e., what causes CO_2 concentrations to go up every year)*? Students begin by sharing their own ideas and questions about what is happening. The *Human Energy Systems* unit is designed to respond to those ideas and questions through engaging students in multiple experiences in which they enact the practices needed for pool-and-flux reasoning while modeling carbon cycling.

Consistent with our iterative, design-based research approach, one significant change made to the unit as a result of Study One was the development of two Global Carbon Cycling models described below. The first model, which students manipulate on their desks, provides a less quantitatively complex introduction. The second, online model, is designed to support students in modeling and observing the effects of changes in fluxes on the size of global carbon pools over time. These models scaffold students in all three important practices described above: (1) observing and reasoning with mechanisms (i.e., photosynthesis, cellular respiration, and combustion), (2) distinguishing between carbon pools and carbon fluxes, and (3) observing and reasoning about quantitative changes in pools and fluxes over time.

Phase 2 begins with students offering and discussing their own initial explanations and questions concerning the cause of increasing CO_2 concentrations. Next, students play a Tiny World Modeling Game (Figure 6), in which they move markers representing carbon atoms among three carbon pools. The carbon fluxes are carbon transforming processes that they have studied in previous units: photosynthesis, cellular respiration, and combustion. In the Tiny World Game, students model (1) a steady state, in which the fluxes are balanced; (2) an annual cycle, in which the photosynthesis flux changes with the seasons; and (3) scenarios that include an unbalanced flux from combustion of fossil fuels.

In a subsequent activity, students use the online Global Carbon Cycling Model (Figure 7) to make global scale, quantitative predictions about effects of changes in fluxes on pool sizes. The computer model has the same pools and fluxes as the Tiny World Model, but pool and flux sizes are based on current global-scale data (Figure 3). Students can control the size and timing of changes in fluxes and see projections of the long-term effects across 50 years. In combination, these activities are designed to scaffold students in developing model-based explanations and predictions for pool-and-flux carbon cycling at the global scale. Students can employ their explanations and predictions to answer the question of what causes atmospheric CO_2 concentrations to increase each year.

Data sources

Data for the second study come from matched pre and post unit assessments for 415 students who completed the *Human Energy Systems* unit in 2019. The sample included 77 middle school students and 338 high school students. Students were from schools in three states in the Midwest, Mountain West, and Northwest. Students completed the *Human Energy Systems* unit in Biology and Environmental Science courses.

In this study, we focus on two items from the *Human Energy Systems* unit pre- and post-assessments (the full unit assessments include six items). The two items we report on in Study Two are similar to those used in Study One in that they ask students about atmospheric CO_2 concentration given a 50% reduction in fossil fuel emissions and in that students respond to the first item before viewing the IPCC model and the second item after viewing the model (Figures 4 and 5). In the first item, students were asked to choose one of five predictions for future atmospheric CO_2 concentrations and explain their choice. In the second item, they were again asked to choose a prediction and then they were asked to explain why they did or did not change their previous prediction after seeing the model.

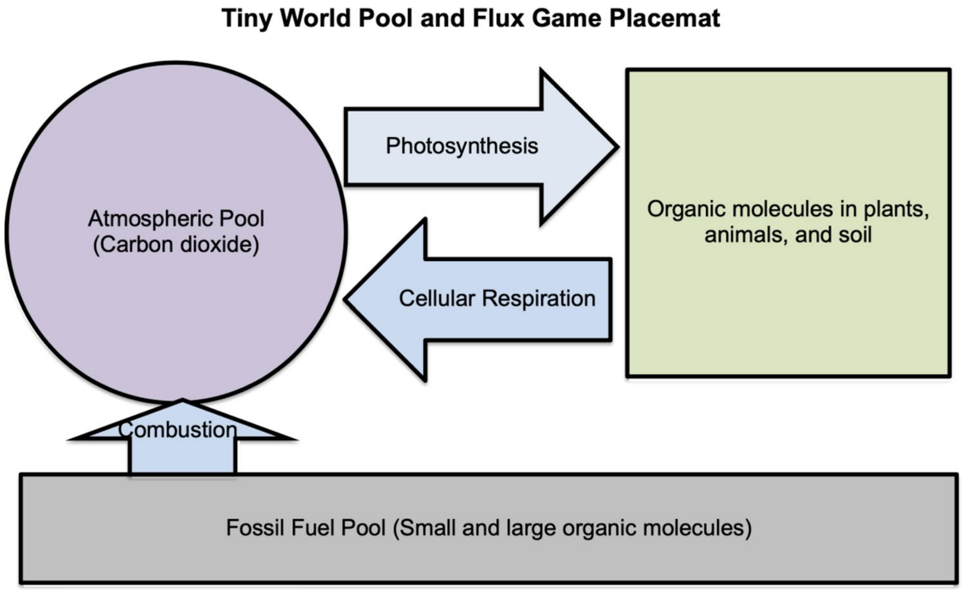

Figure 6. In the Tiny World Modeling Game, students move markers representing carbon atoms through carbon cycle pools and fluxes.

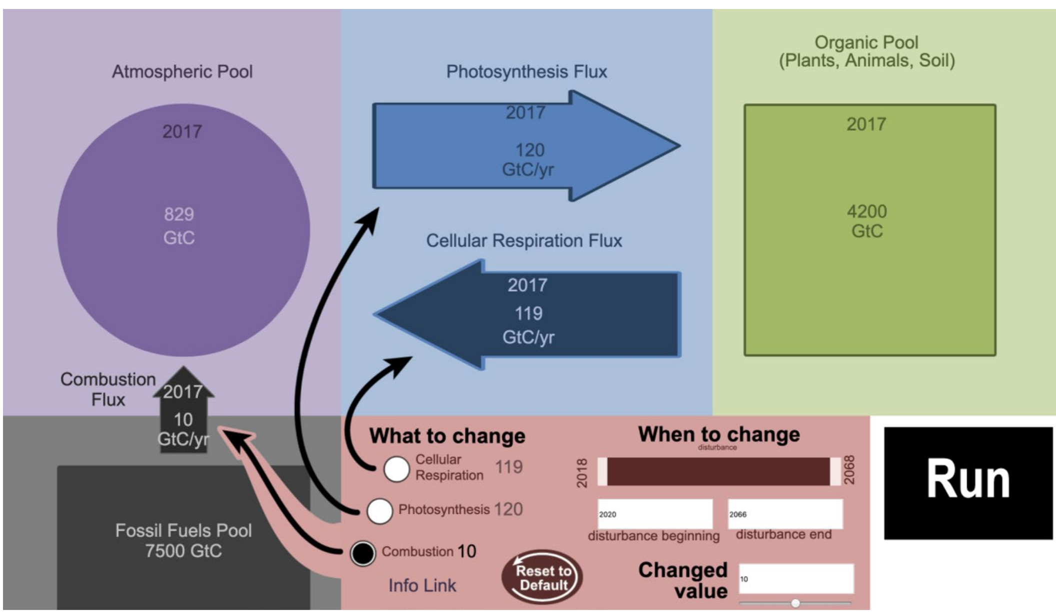

Figure 7. In the Global Carbon Cycling Model, students make predictions for future sizes of carbon pools based on the current size of pools and experimental manipulations of fluxes.

Analysis

Students' item responses were coded into three reasoning levels corresponding to the reasoning levels from Study One. Coding was completed using the *Carbon TIME* machine scoring system developed in collaboration with ACT, an education research and assessment organization. Development, implementation, and validation of the machine scoring system are described in Thomas (2020) and Thomas et al. (2020).

The machine scoring system is based on iterative development and refinement of item rubrics with indicators of each type of reasoning. Rubrics were initially developed and refined with human coders and then machine learning was used to train the Open Source machine-learning engine, Lightside Researcher's Workbench (Mayfield & Rosé, 2013), to code student responses. Machine coding was refined and checked against human coding until a standard of a quadratic weighted kappa (QWK) of at least 0.7 was achieved (Landis & Koch, 1977). Item coding rubrics are available in supplementary materials. Briefly, the descriptors of the coding levels are as follows:

- *Level Three (Pool-and-flux model-based prediction and explanation):* Responses explain that reducing emissions reduces the rate of increase in CO_2 concentrations (the slope of the line).
- *Level Two (Correlation heuristic):* Responses describe incomplete or inaccurate quantitative relationships between CO_2 emissions and CO_2 concentration.
- *Level One (Good versus bad heuristic):* Responses focus on normative and immaterial ideas about consequences of changing fossil fuel use while disregarding numbers and information about carbon pools and fluxes.

Once data were coded, we applied linear probability models to test several effects described in the Results section below. We examined within-student variation so that students' characteristics are not confounded with the results. A statistical comparison between middle and high school students could not be made because the sample only included middle school students from one teacher. However, it is worth noting that running the probability models with and without the middle school students in the sample did not lead to different results.

Results

Research question 1: Effect of seeing the model

We examined the effect of viewing the diagrammatic model (Figure 5) on students' likelihood of moving to a higher reasoning level by comparing students' performance before and after viewing the model, within the same test. On both the pretest and the posttest, we see only a small change in the percentage of students responding at any given level before viewing the IPCC model (black bars in Figures 8 and 9) compared with after viewing the model (grey bars). The probability of a student responding with good versus bad reasoning decreased by 0.0361 ($p < 0.01$) after viewing the model in a test, regardless of whether it was a pretest or posttest—a difference that we judge to be statistically but not educationally significant. The probability of a student responding with pool-and-flux reasoning did not change after viewing the model, again, regardless of whether it was a pretest or posttest.

These findings provide further evidence that offering the diagrammatic carbon cycling model is not very helpful to students who are using the good versus bad heuristic or the correlation heuristic. Instead, it seems that model-based carbon cycle pool-and-flux reasoning is a prerequisite for being able to use the IPCC model in a productive way (i.e., by calculating a net flux and using the net flux to make a prediction for future atmospheric CO_2 concentration).

Research question 2: Effect of completing the human energy systems unit

The second prominent finding is that completing the *Human Energy Systems Unit* did have a significant impact. Analysis of the students' performances on the full unit pre and post assessments showed substantial learning gains associated with completing the *Human Energy Systems* unit (average pre to post increase of 0.779 logits representing a paired t value of 16.398, SE = 0.047, $p < 0.001$, effect size = 0.799).

With regard to the focal assessment items, while only 27% of students provided responses consistent with model-based pool-and-flux reasoning on the pretest (Figure 8), about 52% of students did so on the posttest (Figure 9). This change reflected an increase in the probability of a student using

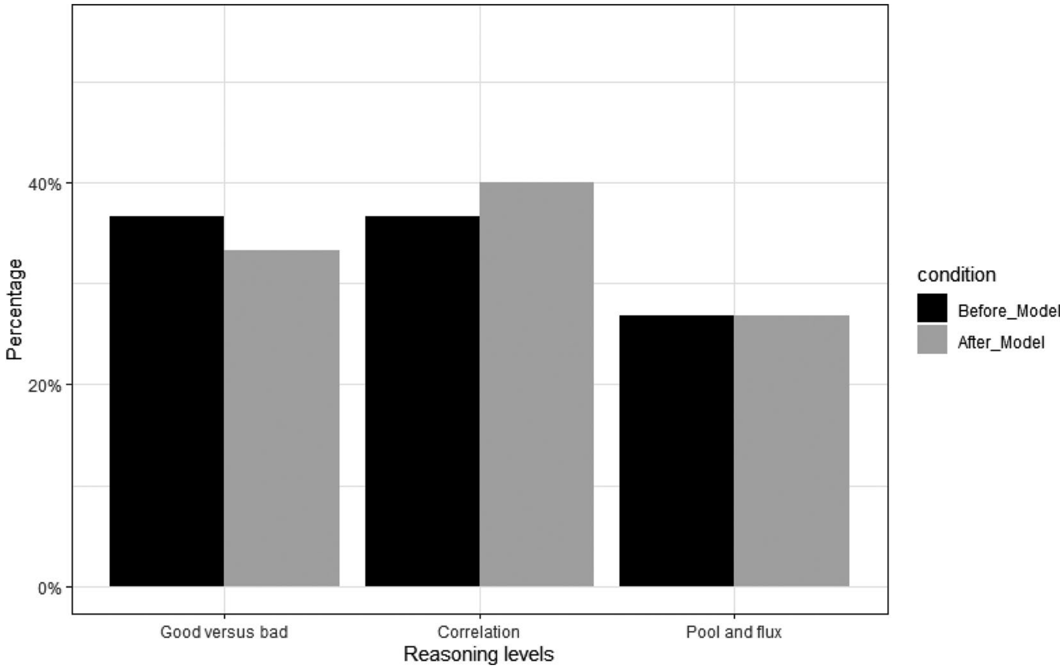

Figure 8. Students responding at each reasoning level on the unit pretest.

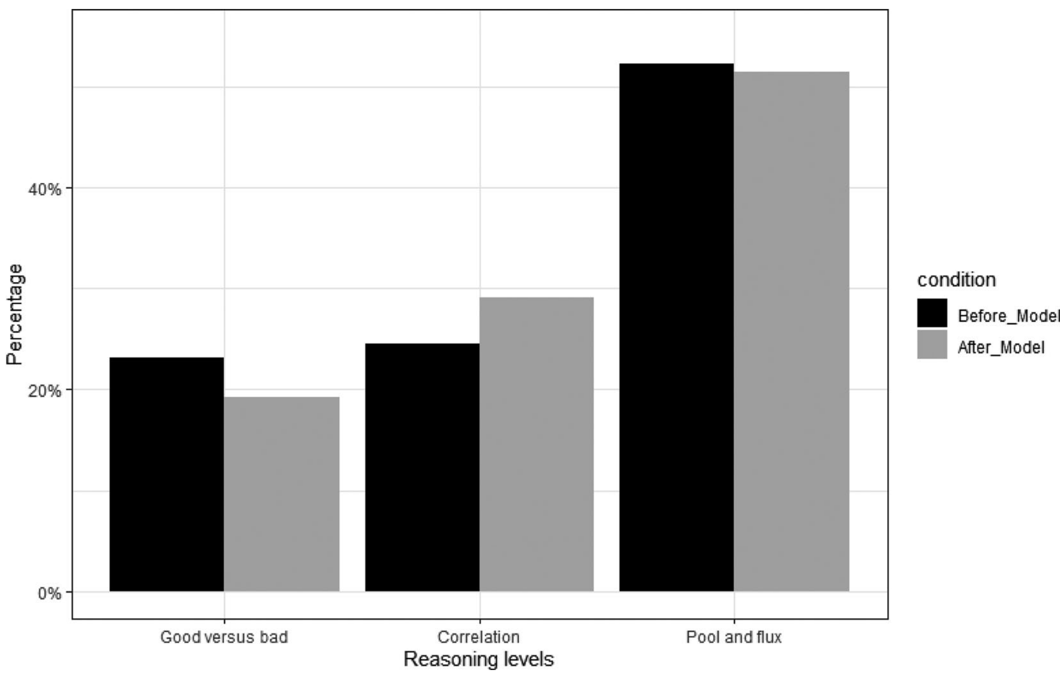

Figure 9. Students responding at each reasoning level on the unit posttest.

pool-and-flux reasoning of 0.252 ($p < 0.001$). The probability of a student relying on the good versus bad heuristic showed a decrease of 0.137, $p < 0.001$).

While it is encouraging to see an increase from 27% to 52% of students who provided responses consistent with model-based pool-and-flux reasoning, it is important to acknowledge that this result also shows that after completing the unit, about 20% of students still provided responses at the good versus bad heuristic level and about one quarter still provided responses at the correlation heuristic level. These results, while promising, are consistent with previous research studies that have shown the entrenched nature of informal approaches to pool-and-flux reasoning.

Discussion

It is tempting, but problematic, to assume that the meanings of representations like the Keeling Curve (Figure 1) or carbon cycling models (Figures 2 and 5) are transparent to students. The results of our studies are consistent with past research and further elucidate the challenges students face in interpreting and using these representations. The correlation and good versus bad heuristics that we describe above are sometimes useful to all of us; these heuristics help us understand that combustion of fossil fuels is problematic. However, neither of these reasoning approaches helps people understand how multiple fluxes affect CO_2 concentration in the atmosphere. In order to evaluate the costs and benefits of different decisions or actions, people need to be able to predict the quantitative impact of changing carbon fluxes on atmospheric CO_2 concentration.

We found that without instruction, almost three-fourths of high school students relied on good versus bad or correlation heuristics, even in a situation where the heuristics were inappropriate. This was true even when they were provided with a diagrammatic pool-and-flux model. After completing an instructional unit—*Human Energy Systems*—in which students used both physically manipulated (Tiny World) and computer-based pool-and-flux models, approximately double the percentage (over half) of students could successfully use a pool-and flux model on the posttest. The percentage of students relying on the least sophisticated good versus bad heuristic decreased significantly as well; only about one fifth of students relied on this type of reasoning on the post assessment. Thus, we found that with strategic instructional approaches aimed at scaffolding important practices, most secondary students could apply model-based pool-and-flux reasoning to make sense of and predict changes occurring within Earth's carbon cycle.

Carbon TIME aims to help students recognize problems that require more than heuristic reasoning, and to be able to use model-based pool-and-flux reasoning when they need to. While it is encouraging that *Carbon TIME* learning experiences helped many students develop capacity for model-based pool-and-flux reasoning, we are interested in exploring how educational experiences can be more successful in this respect. To that end, we will continue our efforts to examine how students make sense of carbon cycle pool-and-flux reasoning in the context of interactions with multiple types of models. We hope to find ways to further refine the unit to support greater facility with important pool-and-flux reasoning practices. Ultimately, we would like to see all participating students benefit from these activities by developing model-based pool-and-flux reasoning that they can use in problem solving throughout their lives.

Note

1. While calculating with this precision is problematic because the model is inconsistent with respect to precision of fluxes, we focused our analysis on the conceptual use of the IPCC model as a reasoning tool rather than on the issue of significant figure standards.

Acknowledgment

We would like to express our appreciation and thanks to Jay Thomas for his work in developing the machine scoring system used in Study Two and to Jennifer Doherty for pointing us to some key pool-and-flux literature.

Funding

This study was supported by National Science Foundation, Grant/Award Number: NSF 1440988. Any opinions, findings, and or conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

ORCID

Beth A. Covitt http://orcid.org/0000-0003-0490-8342

References

Anderson, C. W., de los Santos, E. X., Bodbyl, S., Covitt, B. A., Edwards, K. D., Hancock, J. B., Lin, Q., Morrison Thomas, C., Penuel, W. R., & Welch, M. M. (2018). Designing educational systems to support enactment of Next Generation Science Standards. *Journal of Research in Science Teaching*, *55*(7), 1026–1052. https://doi.org/10.1002/tea.21484

Barzilai, S., & Chinn, C. A. (2020). A review of educational responses to the "post-truth" condition: Four lenses on "post-truth" problems. *Educational Psychologist*, *55*(3), 107–119. https://doi.org/10.1080/00461520.2020.1786388

Bishop, K., & Scott, W. (1998). Deconstructing action competence: Developing a case for a more scientifically-attentive environmental education. *Public Understanding of Science*, *7*(3), 225–236. https://doi.org/10.1088/0963-6625/7/3/002

Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, *9*(2-3), 71–123.

Boden, T. A., Marland, G., & Andres, R. J. (2015). *Global, regional, and national fossil-fuel CO2 emissions*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U. S. Department of Energy.

Bofferding, L., & Kloser, M. (2015). Middle and high school students' conceptions of climate change mitigation and adaptation strategies. *Environmental Education Research*, *21*(2), 275–294. https://doi.org/10.1080/13504622.2014.888401

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, *95*(4), 639–669. https://doi.org/10.1002/sce.20449

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*(1), 61–100.

Chang, C. H., & Pascua, L. (2016). Singapore students' misconceptions of climate change. *International Research in Geographical and Environmental Education*, *25*(1), 84–96. https://doi.org/10.1080/10382046.2015.1106206

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. https://doi.org/10.3102/0013189X032001009

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, *13*(1), 15–42. https://doi.org/10.1207/s15327809jls1301_2

Coyle, K. (2005). *Environmental literacy in America: What ten years of NEETF/Roper research and related studies say about environmental literacy in the US*. National Environmental Education & Training Foundation.

Covitt, B., & Anderson, C. (2018). Assessing scientific genres of argument, explanation and prediction. In A. Bailey, C. Maher, & L. Wilkinson (Eds.), *Language, literacy and learning in the STEM disciplines: How language counts for English learners* (pp. 206–230). Routledge.

Cronin, M. A., Gonzalez, C., & Sterman, J. D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes*, *108*(1), 116–130. https://doi.org/10.1016/j.obhdp.2008.03.003

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8.

Duncan, R. G., & Rivet, A. E. (2013). Science education. Science learning progressions. *Science*, *339*(6118), 396–397. https://doi.org/10.1126/science.1228692

Düsing, K., Asshoff, R., & Hammann, M. (2019). Tracing matter in the carbon cycle: Zooming in on high school students' understanding of carbon compounds and their transformations. *International Journal of Science Education*, *41*(17), 2484–2507. https://doi.org/10.1080/09500693.2019.1686665

Dutt, V., & Gonzalez, C. (2012). Decisions from experience reduce misconceptions about climate change. *Journal of Environmental Psychology*, *32*(1), 19–29. https://doi.org/10.1016/j.jenvp.2011.10.003

Feinstein, N. W., & Waddington, D. I. (2020). Individual truth judgments or purposeful, collective sensemaking? Rethinking science education's response to the post-truth era. *Educational Psychologist*, *55*(3), 155–166.

Gee, J. (1991). Socio-cultural approaches to literacy (literacies). *Annual Review of Applied Linguistics*, *12*, 31–48.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Gotwals, A. W. (2012). Learning progressions for multiple purposes: Challenges in using learning progressions. In *Learning progressions in science* (pp. 461–472). Brill Sense.

Gunckel, K., Covitt, B., Salinas, I., & Anderson, C. (2012). A learning progression for water in socio-ecological systems. *Journal of Research in Science Teaching*, *49*(7), 843–868. https://doi.org/10.1002/tea.21024

Guy, S., Kashima, Y., Walker, I., & O'Neill, S. (2013). Comparing the atmosphere to a bathtub: Effectiveness of analogy for reasoning about accumulation. *Climatic Change*, *121*(4), 579–594. https://doi.org/10.1007/s10584-013-0949-3

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences*, *16*(3), 307–331. https://doi.org/10.1080/10508400701413401

Hogan, K., & Weathers, K. C. (2003). Psychological and ecological perspectives on the development of systems thinking. In A. Berkowitz, C. Nilon, & K. Hollweg (Eds.), *Understanding urban ecosystems* (pp. 233–260). Springer.

Hoover, K. S. (2019). The carbon cycle and climate change. *The Science Teacher*, *87*(1), 22–28.

Hsu, S. J. (2004). The effects of an environmental education program on responsible environmental behavior and associated environmental literacy variables in Taiwanese college students. *The Journal of Environmental Education*, *35*(2), 37–48. https://doi.org/10.3200/JOEE.35.2.37-48

Intergovernmental Panel on Climate Change (IPCC). (2001). *Climate change 2001: The scientific basis*. The Press Syndicate of the University of Cambridge.

IPCC. (2018). Summary for policymakers. In Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, & T. Waterfield (Eds.), *Global Warming of 1.5 °C. An IPCC Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (p. 32). World Meteorological Organization.

Iyengar, S., & Massey, D. S. (2019). Scientific communication in a post-truth society. *Proceedings of the National Academy of Sciences*, *116*(16), 7656–7661. https://doi.org/10.1073/pnas.1805868115

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kollmuss, A., & Agyeman, J. (2002). Mind the gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research*, *8*(3), 239–260. https://doi.org/10.1080/13504620220145401

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Mayfield, E., & Rosé, C. P. (2013). LightSIDE: Open source machine learning for text. In M. Sherrmis & J. Burstein (Eds.), *Handbook of automated essay evaluation* (pp. 146–157). Routledge.

McBeth, W., & Volk, T. L. (2009). The national environmental literacy project: A baseline study of middle grade students in the United States. *The Journal of Environmental Education*, *41*(1), 55–67. https://doi.org/10.1080/00958960903210031

Mogensen, F., & Schnack, K. (2010). The action competence approach and the 'new' discourses of education for sustainable development, competence and quality criteria. *Environmental Education Research*, *16*(1), 59–74. https://doi.org/10.1080/13504620903504032

Mohan, L., Chen, J., & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching*, *46*(6), 675–698. https://doi.org/10.1002/tea.20314

Moxnes, E., & Saysel, A. K. (2009). Misperceptions of global climate change: Information policies. *Climatic Change*, *93*(1-2), 15–37. https://doi.org/10.1007/s10584-008-9465-2

National Oceanic and Atmospheric Administration. (2019). *Global carbon dioxide growth in 2018 reached 4th highest on record [news feature]*. Retrieved November 13, 2020 from https://www.noaa.gov/news/global-carbon-dioxide-growth-in-2018-reached-4th-highest-on-record

National Research Council. (1996). *National science education standards*. National Academies Press.

National Research Council. (2006). Systems for state science assessment. Committee on Test Design for K-12 Science Achievement. In M. R. Wilson & M. W. Bertenthal (Eds.), *Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education*. The National Academies Press.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press.

Niebert, K., & Gropengiesser, H. (2013). Understanding and communicating climate change in metaphors. *Environmental Education Research*, *19*(3), 282–302. NAAEE. https://naaee.org/eepro/publication/excellence-environmental-education-guidelines-learning-k-12.

North American Association for Environmental Education (NAAEE). (2019). K-12 Environmental education: Guidelines for excellence. NAAEE. https://naaee.org/eepro/publication/excellence-environmental-education-guidelines-learning-k-12.

Özdem, Y., Dal, B., Öztürk, N., Sönmez, D., & Alper, U. (2014). What is that thing called climate change? An investigation into the understanding of climate change by seventh-grade students. *International Research in Geographical and Environmental Education*, *23*(4), 294–313. https://doi.org/10.1080/10382046.2014.946323

Parker, J. M., De Los Santos, E. X., & Anderson, C. W. (2015). Learning progressions & climate change. *The American Biology Teacher*, *77*(4), 232–238.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

Peel, A., Sadler, T., Kinslow, A., Zangori, L., & Friedrichsen, P. (2017). Climate change as an issue for socio-scientific issues teaching and learning. In *Teaching and learning about climate change* (pp. 153–165). Routledge.

Project Learning Tree. (n.d.). *12 videos to help us understand climate change*. Project Learning Tree. https://www.plt.org/educator-tips/videos-climate-change-middle-school.

Pruneau, D., Gravel, H., Bourque, W., & Langis, J. (2003). Experimentation with a socio-constructivist process for climate change education. *Environmental Education Research*, *9*(4), 429–446. https://doi.org/10.1080/1350462032000126096

Reichert, C., Cervato, C., Larsen, M., & Niederhauser, D. (2014). Conceptions of atmospheric carbon budgets: Undergraduate students' perceptions of mass balance. *Journal of Geoscience Education*, *62*(3), 460–468. https://doi.org/10.5408/13-052.1

Reichert, C., Cervato, C., Niederhauser, D., & Larsen, M. D. (2015). Understanding atmospheric carbon budgets: Teaching students conservation of mass. *Journal of Geoscience Education*, *63*(3), 222–232. https://doi.org/10.5408/14-055.1

Shepardson, D. P., Niyogi, D., Choi, S., & Charusombat, U. (2009). Seventh grade students' conceptions of global warming and climate change. *Environmental Education Research*, *15*(5), 549–570. https://doi.org/10.1080/13504620903114592

Shepardson, D. P., Niyogi, D., Choi, S., & Charusombat, U. (2011). Students' conceptions about the greenhouse effect, global warming, and climate change. *Climatic Change*, *104*(3-4), 481–507. https://doi.org/10.1007/s10584-009-9786-9

Shepardson, D. P., Roychoudhury, A., Hirsch, A., Niyogi, D., & Top, S. M. (2014). When the atmosphere warms it rains and ice melts: Seventh grade students' conceptions of a climate system. *Environmental Education Research*, *20*(3), 333–353. https://doi.org/10.1080/13504622.2013.803037

Sterman, J. D., & Sweeney, L. B. (2007). Understanding public complacency about climate change: Adults' mental models of climate change violate conservation of matter. *Climatic Change*, *80*(3-4), 213–238. https://doi.org/10.1007/s10584-006-9107-5

Stubenvoll, M., & Marquart, F. (2019). When facts lie: The impact of misleading numbers in climate change news. In W. Leal Filho, B. Lackner, & H. McGhie (Eds.), *Addressing the challenges in communicating climate change across various audiences* (pp. 31–46). Springer.

Sweeney, L. B., & Sterman, J. D. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *System Dynamics Review*, *23*(2-3), 285–311. https://doi.org/10.1002/sdr.366

The Globe Program. (n.d). *Biosphere*. The Globe Program: A worldwide science and education program. https://www.globe.gov/get-trained/protocol-etraining/etraining-modules/16867717/3099387

Thomas, J. (2020, April). *Using automated scoring to monitor and improve the assessment system*. Annual meeting of the National Council on Measurement in Education (Conference canceled).

Thomas, J., Draney, K., & Bathia, S. (2020, April). *Using machine learning to make assessment of NGSS based three-dimensional science scalable*. Annual meeting of the American Educational Research Association (Conference canceled).

Van Dooren, W., De Bock, D., Janssens, D., & Verschaffel, L. (2007). Pupils' over-reliance on linearity: A scholastic effect?*The British Journal of Educational Psychology*, *77*(Pt 2), 307–321. https://doi.org/10.1348/000709906X115967

You, H. S., Marshall, J. A., & Delgado, C. (2018). Assessing students' disciplinary and interdisciplinary understanding of global carbon cycling. *Journal of Research in Science Teaching*, *55*(3), 377–398. https://doi.org/10.1002/tea.21423

Zangori, L., Peel, A., Kinslow, A., Friedrichsen, P., & Sadler, T. D. (2017). Student development of model-based reasoning about carbon cycling and climate change in a socio-scientific issues unit. *Journal of Research in Science Teaching*, *54*(10), 1249–1273. https://doi.org/10.1002/tea.21404

Zeidler, D. L., & Kahn, S. (2014). *It's debatable. Using socio-scientific issues to develop scientific literacy*. National Science Teachers Association.

Zeidler, D. L., Sadler, T. D., Applebaum, S., & Callahan, B. E. (2009). Advancing reflective judgment through socioscientific issues. *Journal of Research in Science Teaching*, *46*(1), 74–101. https://doi.org/10.1002/tea.20281