



Knowing how and knowing when: unpacking public understanding of atmospheric CO₂ accumulation

Erik O. Sterner¹  · Tom Adawi² · U. Martin Persson¹ · Ulrika Lundqvist¹

Received: 6 July 2018 / Accepted: 19 March 2019 / Published online: 16 April 2019
© The Author(s) 2019

Abstract

It has been demonstrated that most people have a limited understanding of atmospheric CO₂ accumulation. Labeled stock-flow (SF) failure, this phenomenon has even been suggested as an explanation for weak climate policy support. Drawing on a typology of knowledge, we set out to nuance previous research by distinguishing between different types of knowledge of CO₂ accumulation among the public and by exploring ways of reasoning underlying SF failure. A mixed methods approach was used and participants ($N = 214$) were enrolled in an open online course. We find that ostensibly similar SF tasks show seemingly contradictory results in terms of people's understanding of CO₂ accumulation. Participants performed significantly better on stock stabilization tasks that explicitly ask about the relationship between stocks and flows, compared with a typical SF task that does not direct the participants' attention to what knowledge they should use. This suggests that people possess declarative and procedural knowledge of accumulation (knowing about the principles of mass balance, i.e., *what* and *how* to use them) but lack conditional knowledge of accumulation (knowing *when* to use these principles). Additionally, through a thematic analysis of answers to an open-ended question, we identified three overarching ways of reasoning when dealing with SF tasks: system, pattern, and phenomenological reasoning, providing additional theoretical insights to explain the large difference in performance between the different SF tasks. These more nuanced perspectives on SF failure can help inform interventions aimed at increasing climate science literacy and point to the need for more detailed explorations of public knowledge needed to leverage climate policy support.

Keywords Stock-flow failure · Climate science literacy · Mixed methods · Mental models · Knowledge · CO₂ accumulation

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10584-019-02423-8>) contains supplementary material, which is available to authorized users.

✉ Erik O. Sterner
erik.sterner@chalmers.se

Extended author information available on the last page of the article

1 Introduction

“Carbon in the atmosphere is rising, even as emissions stabilize” was the heading of a recent article in the *New York Times* (Gillis 2017). The author was puzzled by this: “If the amount of the gas that people are putting out has stopped rising, how can the amount that stays in the air be going up faster than ever?” In fact, each ton of carbon dioxide (CO₂) emitted from fossil fuel combustion increases CO₂ concentration in the atmosphere for at least thousands of years (Archer and Brovkin 2008), meaning that emissions yesterday, today, and tomorrow produce warming that lasts. Hence, the total amount of CO₂ emissions needs to be limited to avoid dangerous interference with the climate system, with net CO₂ emissions eventually coming down to zero for atmospheric concentrations to stabilize. We are rapidly approaching the amount of carbon we can emit while staying below 2 °C warming and with current levels of emissions that carbon budget would be emptied within a few decades (Goodwin et al. 2018; Peters et al. 2012).

Despite this enormous challenge, the basic relationship between CO₂ emissions and atmospheric CO₂ concentrations is poorly understood by the public. The first study demonstrating the widespread failure to grasp the fundamental relationship between stocks and flows of CO₂ in the carbon cycle—known as *stock-flow (SF) failure*—was that by Sterman and Booth Sweeney (2007). In their sample of 212 graduate students at Massachusetts Institute of Technology (MIT) within science, technology, engineering, mathematics, or economics, 84% gave answers to an SF task that violated basic mass-balance principles, assuming atmospheric carbon stocks would stabilize even if emissions exceeded removals. This is “analogous to arguing a bathtub filled faster than it drains will never overflow” (ibid. p. 216). The authors hypothesized that SF failure is due to the use of a *pattern matching* heuristic, where respondents match trends in flows and stocks, rather than accounting for the stock-flow dynamics of the system.

Since the seminal paper by Sterman and Booth Sweeney (2007), several studies have focused on SF failure, and these can be divided into three main strands of research. First, there are studies that aim to confirm the findings by Sterman and Booth Sweeney (Cronin et al. 2009; Dutt and Gonzalez 2009). Second, there are studies that alter the tasks or the setting in an attempt to establish if the poor performance depends on external factors such as task design and context and background of participants (Cronin and Gonzalez 2007; Sterman and Booth Sweeney 2002, 2007; Guy et al. 2013; Fischer et al. 2015; Newell et al. 2016). Third, there are intervention studies that aim to improve understanding among the participants, mainly through knowledge transfer from other contexts or by active learning methods (Dutt and Gonzalez 2009, 2012a, b; Moxnes and Saisel 2009). A different approach was taken by Dryden et al. (2018), who simply asked for an estimation of the atmospheric residence time for CO₂. Their results show that people estimate CO₂ to be gone from the atmosphere within decades of being emitted, which further highlights misunderstandings around CO₂ accumulation.

In this paper, we report on findings from a mixed methods study of public understanding of atmospheric CO₂ accumulation. First and foremost, we wanted to take a closer look at the common yet intriguing finding in the literature on SF failure that most people “have difficulty relating the flows into and out of a stock to the level of the stock, even in simple, familiar contexts such as bank accounts and bathtubs” (Sterman 2011, p. 817). We surmised that most people have an intuitive understanding of the concept of accumulation, but this type of understanding is *not revealed* in the kind of CO₂ stabilization task used by Sterman and Booth Sweeney (2007). We test this hypothesis by drawing on a *typology of knowledge* (Biggs 2003) that distinguishes

between three different types of knowledge that SF tasks can assess: *declarative* (knowing *what*), *procedural* (knowing *how*), and *conditional* (knowing *when*) knowledge.

We note that previous research on SF failure seems to have overlooked this aspect of task design (there is, at least, no explicit discussion of different types of knowledge). Consequently, we developed two alternative SF tasks (using the carbon cycle and a bathtub, respectively, as contexts) with lower knowledge demands,¹ so to speak, they explicitly ask about the relationship between the flows into and out of a stock for the stock to stabilize. Performance on these two alternative SF tasks was compared with performance on a task with higher knowledge demands, similar to the one used by Sterman and Booth Sweeney (2007). To further test the surmised disconnection between these types of knowledge, we used a pre- and post-test design, to investigate whether an explanation of the knowledge required to solve the tasks would have any effect on the performance on the kind of task used by Sterman and Booth Sweeney (2007).

In addition, through qualitative data, we sought to gain insight into different *ways of reasoning* when solving the SF tasks to better understand what could explain SF failure and why people seem unable to apply intuitive knowledge about accumulation in certain tasks. It is widely acknowledged that an understanding of how people make sense of concepts and principles in science is essential for effective science teaching and communication (Ambrose et al. 2010; Morgan et al. 2002). Yet, most previous research on SF failure has focused on task performance without probing how people actually reason when solving various SF tasks (Korzilius et al. 2014). One notable exception is the study by Korzilius et al. (2014), which used the think-aloud method to explore “reasoning patterns” used by people when solving SF tasks. The SF tasks in their study, however, were more generic, while our study focuses on ways of reasoning about atmospheric CO₂ accumulation and how this relates to task performance. There are several reasons for the necessity of studying ways of reasoning in the CO₂ context, ranging from the carbon cycle dynamics (which posit that the capacity for uptake of CO₂ is determined by the historical emissions) to the amount of public debate on the topic. As an example, the New York Times article mentioned earlier received more than 600 comments online.

Finally, we investigated whether there is a connection between performance on our SF tasks and stated climate policy support, as suggested by some (Sterman 2008; Chen 2011; Dutt and Gonzalez 2012a). While there is some support for the notion that climate science literacy enhances concern for climate change (Hornsey et al. 2016; Guy et al. 2014; Ranney and Clark 2016), the previous literature on SF failure in the climate context has not explicitly tested for a relationship between SF task performance and stated climate policy support.

2 Method

2.1 Study context and participants

The context of the study was a massive open online course (MOOC) entitled “Sustainability in Everyday Life,”² offered by Chalmers University of Technology between Aug 29, 2016, and Oct 16, 2016, using the EdX platform. The course was not part of

¹ The knowledge demands of a task refer to the cognitive resources needed to solve the task.

² ChalmersX, ChM002x is the edx.org course code

any university program, required no particular prior knowledge, was open to take, and free of charge for everyone with internet access. It only generated a diploma if completed. This MOOC was chosen for this study due to the relevant course content and the possibility to get a large number of respondents.

The sustainability MOOC consisted of five modules or themes: globalization, climate, food, energy, and chemicals. The performance on different kinds of SF tasks was assessed during the climate module, directly after a general introductory video on climate change, which did not address the knowledge tested by the SF tasks, and a question assessing climate policy support was included in the pre-course survey (i.e., before the students were introduced to any course contents). To motivate task completion, the SF tasks gave points that contributed to the total examination of the course regardless of performance.

Of 3540 participants enrolled in the course, 300 started the climate change module where the SF tasks were placed. Of these, 214 participated in the study by completing all of the SF tasks. A total of 49 countries were represented in the sample, with most participants from the EU/EEA (58), the USA (25), India (11), and Mexico (9). See the [supplementary material](#) for the full list. The sample included 119 females and 77 males (18 participants had not disclosed their gender). The participants' average age was 38 years. Of the 92% who stated their highest attained educational level, 81% had a bachelor's degree or higher. Admittedly, the high average education level, together with the fact that the participants have opted to take a course in sustainability, implies that our participants do not constitute a representative sample of the general public (see the [supplementary material](#) for more information on the course context and participants).

2.2 Study design

In this section, the overall design of the study is described along with the design of the tasks; in the next section we explain—by drawing on a typology of knowledge—how tasks were designed to assess different types of knowledge. Table 1 depicts the overall design of the study, summarizing the different tasks (all tasks were completed online) and the order in which they were completed—the five steps of the study design.

Prior to the SF tasks, the participants were given a question aiming to measure stated preferences with respect to climate policy (T0). Here, the participants were asked which one of the following statements came closest to their personal view:

1. Society should not take any steps to reduce emissions of greenhouse gases (such as CO₂).
2. Society should reduce emissions of greenhouse gases in the future, in response to climate impacts as they actually occur.
3. Society should take moderate actions to reduce emissions of greenhouse gases today, to reduce future climate impacts.
4. Society should take strong action to reduce emissions of greenhouse gases today, to reduce future climate impacts.
5. I do not know/I have not formed an opinion.

The alternatives were formulated to reflect attitudes of “wait and see” (2) or “go slow” (3), as discussed by Sterman (2008).

Table 1 An overview of the study design, describing the tasks' order and format, the types of knowledge assessed, and the number of participants that completed each task

Task	N	Task format	Knowledge assessed
T0: climate policy task	167	Multiple choice (text)	n.a.
T1: main task	214	Multiple choice (graph)	Declarative, procedural, and conditional (higher knowledge demand)
T2: alternative task (intervention)	A: 74 B: 77 C: 63	A and B: multiple choice (text) C: reading explanatory text	A and B: declarative and procedural (lower knowledge demand) C: n.a.
T3: main task repeated (post-test)	214	Multiple choice (graph)	Declarative, procedural, and conditional (higher knowledge demand)
T4: reasoning about T3	204	Open ended	n.a.

In the first SF task (T1), participants completed a task, which we will refer to as the main SF task that was designed to be similar to the task used by Sterman and Booth Sweeney (2007).³ The main SF task consists of a short introductory text, graphs of the annual historic emissions and uptake of CO₂, a graph of a scenario with a stabilized amount of CO₂ in the atmosphere, and a multiple choice question (see Fig. 1). Participants were asked to choose, among four alternative graphs, the graph depicting emissions and uptake trajectories that is consistent with the scenario for CO₂ stabilization. The correct answer is alternative 3 (marked with a green symbol).

Although the main SF task (see Fig. 1) was designed to be similar to the task used by Sterman and Booth Sweeney (2007), our version of the task contained less superfluous information, both in text and graphs, to avoid cognitive overload. However, we added more elaborate information about the CO₂ uptake, which was given the same attention as the emissions. For the first period of the graphs (i.e., 1900–2015), the CO₂ emissions and uptake values were produced using a simple climate model (Sterner and Johansson 2017), which simulates the carbon cycle response. For this, widely used “historic emissions” that give a realistic impression were used (Meinshausen et al. 2011).

No feedback on task performance is provided to the participants throughout the full set of tasks. In the second SF task (T2), participants were randomly assigned to complete one of three alternative tasks, T2A–C (see Table 1). In contrast to the main SF task, these tasks were designed to direct the participants' attention towards the principles of accumulation. This was done by explicitly asking questions about (T2A–B) or describing (T2C) the relationship between the flows into and out of a stock in order for the stock to stabilize at a certain level. As a consequence, and as we argue in the next section, these tasks differ from the main SF task in terms of their *knowledge demands*—that is, in terms of the type of knowledge they assess. The first task (T2A) uses the carbon cycle as context (see Fig. 2), while the second (T2B) uses a bathtub as context (see Fig. 3). These two tasks are central to our hypothesis (stated in the introduction) as they allow us to investigate whether participants perform better on stock stabilization tasks that explicitly ask about the relationship between the flows into and out of a stock (T2A–B), compared with the kind of task used in previous studies (Dutt and Gonzalez 2012a; Guy et al. 2013; Newell et al. 2016; Sterman and Booth Sweeney 2007) (T1). The third task (T2C), not involving a question, uses a bathtub analogy to explain atmospheric CO₂ accumulation in a simple way (see figure in the [supplementary material](#)); in T2C, the respondents were only asked to confirm that they had studied the analogy. This task, in

³ Sterman and Booth Sweeney (2007) used a text from the summary for policymakers in the IPCC's Third Assessment Report (Houghton et al. 2001).

Task 1

The amount of CO₂ in the atmosphere is affected by two flows of CO₂, one into the atmosphere (emissions) and one out of the atmosphere (uptake). CO₂ emissions are mainly caused by the burning of fossil fuels and lead to an increase in the amounts of CO₂ in the atmosphere. CO₂ is taken up by forests and oceans, causing a decrease in the amounts of CO₂ in the atmosphere. In the last century, emissions of CO₂ have exceeded uptake and the amount of CO₂ in the atmosphere have increased. Figure I shows historical levels of CO₂ emissions and estimated levels of CO₂ uptake until year 2015. Current levels of emissions are about 36 billion tons of CO₂ per year and uptake is about 18 billion tons of CO₂ per year.

Figure II shows historical amounts of CO₂ in the atmosphere until 2015 followed by a scenario of future amounts of CO₂ in the atmosphere. In this CO₂ scenario the amount of CO₂ in the atmosphere gradually rises and stabilizes at a level about 10 % higher than today in the year 2050, as shown in Figure II.

What would the levels of emissions and uptake look like for the rest of this century in order for the amount of CO₂ in the atmosphere to follow the scenario in Figure II?

Provide your answer by selecting one of the alternatives A-E with curves representing the emissions and the uptake, respectively, for the period 2015-2100.

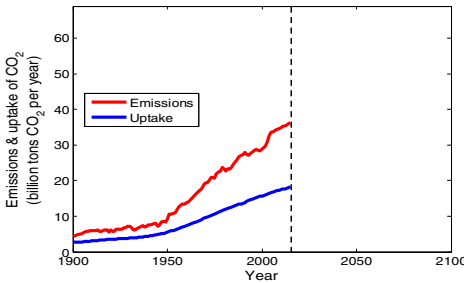


Figure I. Historical CO₂ emissions (red) and uptake (blue) in billion tons of CO₂ per year.

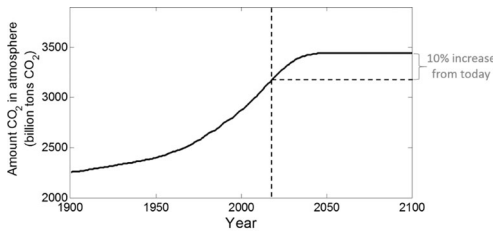
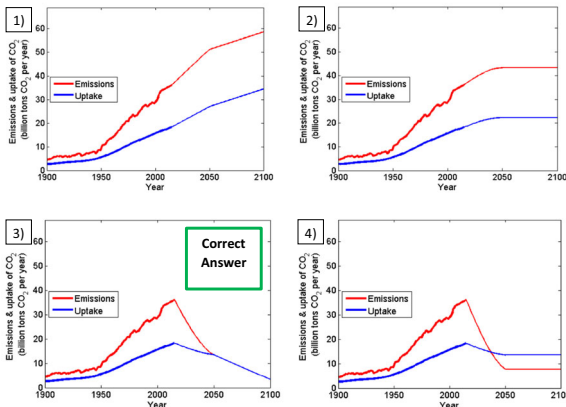


Figure II. Amount of CO₂ in the atmosphere, historically up to today followed by a scenario of the future amounts of CO₂ in the atmosphere.



◀ **Fig. 1** The main SF task (T1/T3), which also included an answer alternative 5: “I don’t know.” The correct answer is alternative 3 in which emissions and uptake meet—which causes the atmospheric CO₂ amount to stabilize—after which they jointly diminish over time (since lower emissions causes uptake to fall)

contrast to T2A–B, presented the participants with the knowledge that is needed to solve the main SF task.

Thereafter, the participants were asked to complete the main SF task again (T3) (see Table 1 and Fig. 1). The logic behind this was that the alternative tasks, T2A–C, would help participants by pointing to the knowledge needed for solving the main SF task, thus allowing us to investigate whether these three tasks could serve as educational interventions that improve performance on the main SF task.

In addition to testing people’s performance on SF tasks with different knowledge demands, we aim to unpack public understanding of CO₂ accumulation by exploring people’s ways of reasoning when solving SF tasks. We did this by, in task T4, asking participants to provide a short, written explanation of how they reasoned when choosing to keep or change their answer when completing the main SF task again (T3). Collecting the combined data of how people answer on SF tasks and how they reason while doing so, we aim to study the mental representations used by the participants when answering the main SF task. Mental representations are similar to mental models (which are “personal, internal representations of external reality that people use to interact with the world around them”) (Jones et al. 2011) but are here used instead of mental models to emphasize that their nature is not seen to be stable or static to the same extent that mental models are sometimes viewed.

2.3 Task design and knowledge demands

As noted above, the tasks—the main SF task (T1/T3) and the alternative tasks (T2A–B)—were designed to assess different types of knowledge. While knowledge can be classified in many ways (Alexander et al. 1991), we draw on a typology described by (among others) Biggs (2003), comprising three types of knowledge:

1. *Declarative* knowledge, which refers to “knowing *about* things [such as facts, concepts, and principles], or knowing *what*” (p. 41)
2. *Procedural* knowledge, which refers to “knowing *how* to do things, such as carrying out procedures or enacting skills” (p. 42)⁴
3. *Conditional* knowledge, which refers to “knowing *when* to do these things [...] under what conditions one should do this as opposed to that” (p. 42)

These types of knowledge are “characterized by the *function* they fulfil in the performance of a target task” (de Jong and Ferguson-Hessler 1996, p. 106). To put it differently, we are interested in *knowledge-in-use* (ibid. p. 110).⁵ Moreover, while “it is certainly possible to

⁴ It is worth noting that this definition of procedural knowledge, also endorsed by others (e.g., Alexander et al. 1991; de Jong and Ferguson-Hessler 1996), does not restrict procedural knowledge to being *tacit*—in contrast to some definitions, primarily in other scientific domains include both tacit and explicit knowledge.

⁵ This theoretical stance—and the three-part typology of knowledge described here—is perhaps best understood in light of the long-standing criticism leveled at university teaching for placing too much emphasis on declarative knowledge (Biggs 2003), or even procedural knowledge (Turns and Van Meter 2011). We join Turns and Van Meter (2011) in arguing that one way of mitigating this problem is by anchoring teaching (and educational research) in the typology of knowledge described here.

know the what of a thing without knowing the how or when of it” (Alexander et al. 1991, p. 323), successful problem solving requires the use of all three of these types of knowledge (Turns and Van Meter 2011). With these theoretical deliberations in mind, we now turn to an *epistemological demand analysis* (de Jong and Ferguson-Hessler 1996)—i.e., an analysis of the knowledge demands—of our SF tasks.

Tasks T2A (climate context) and T2B (bathtub context) were designed to assess declarative and procedural knowledge of accumulation. That is, in these tasks, participants first have to recall *what* the principles of accumulation (i.e., principles of mass balance) say—thus demonstrating *declarative* knowledge. Next, they have to figure out *how* to apply these principles to arrive at the relationship between the emissions/inflow and uptake/outflow for the amount of CO₂ or water to stabilize at a certain level—thus demonstrating *procedural* knowledge.⁶ The difference between T2A and T2B is mainly the familiarity of the context, where the more familiar context of a bathtub may make it easier to draw on knowledge that is relevant for solving the problem.

In the main SF task (T1/T3), on the other hand, participants not only have to *apply* the principles of accumulation—thus demonstrating declarative and procedural knowledge (as in T2A–B)—but also have to realize that this is what the task *requires them to do*—thus demonstrating *conditional* knowledge. Note that the main SF task does not direct the participants’ attention towards the principles of accumulation; that is, it does not explicitly ask about the relationship between the emissions and uptake for the amount of CO₂ to stabilize. As such, one can argue that the main SF task (T1/T3) poses *higher demands on knowledge*, compared with tasks T2A–B.

2.4 Data analysis

In addition to descriptive statistics, a chi-square test of homogeneity was used to determine if the rate of success was significantly different between any pair of groups on the same task or any pair of tasks for the same group.

An *inductive thematic analysis* (Braun and Clarke 2006) was used to analyze the participants’ written answers to the open-ended question, “Briefly explain how you reasoned when choosing to keep or change your answer.” In line with this kind of qualitative analysis, a set of themes was identified after coding the data and sorting and sifting the codes in an iterative way. (For a more detailed account of the analysis, see the [supplementary material](#).) These themes provided a deeper understanding of the ways of reasoning being used when answering the main SF task and made it possible to relate the performance on the different SF tasks to different ways of reasoning.

3 Results

3.1 Performance on SF tasks with different knowledge demands

Table 2 shows that there was a large difference between participants’ performance on the SF tasks that assessed different types of knowledge and SF tasks with different knowledge

⁶ That is, they have to carry out the following calculation (procedure): $\dot{A} = 0 \rightarrow E = U$, where A stands for the amount of CO₂ or water, E for emissions/inflow, and U for uptake/outflow.

Task 2A (T2A)

Consider the emission and uptake setting in Figure 1 [see Figure 1].

What is required of the relationship between the emissions and uptake of CO₂, in order for the amount of CO₂ in the atmosphere to stop increasing and stabilize at a certain level in the future?
Provide your answer by selecting one of the alternatives A-E.

- A – Emissions and uptake should continue growing but keep their current difference
- B - Emissions and uptake should stop growing and keep their current difference
- C - Emissions should reduce to and stay equal to the uptake
- D - Emissions should reduce to and stay at a level below the uptake
- E - I don't know

Fig. 2 A description of task T2A, directing participants’ attention towards the principles of accumulation in the original carbon cycle context. T2A was designed to have a lower knowledge demand compared with the main SF task: it (only) assesses declarative and procedural knowledge of accumulation

demands. The main SF task—both as a pre-test and post-test—had a significantly lower success rate than the two alternative tasks, T2A (carbon cycle context) and T2B (bathtub context), that directed the participants’ attention towards the principles of accumulation and hence did not assess conditional knowledge. The success rate for the participants who were assigned T2A went from 26 on the main SF task to 54% on the alternative task. For the T2B group, the success rate increased from 17 to 70%. These differences are statistically significant ($p < 0.001$) and indicate a high level of intuitive understanding—declarative and procedural knowledge—of the principles of accumulation. The level of education also seems to be

Task 2B (T2B)

Consider the inflow and outflow setting in Figure Tub.

What is required of the relationship between the inflow and outflow of water, in order for the amount of water in the bathtub to stop increasing and stabilize at a certain level?
Provide your answer by selecting one of the alternatives A-E.

- A – Inflow and outflow should continue growing but keep their current difference
- B - Inflow and outflow should stop growing and keep their current difference
- C - Inflow should reduce to and stay equal to the outflow
- D - Inflow should reduce to and stay at a level below the outflow
- E - I don't know

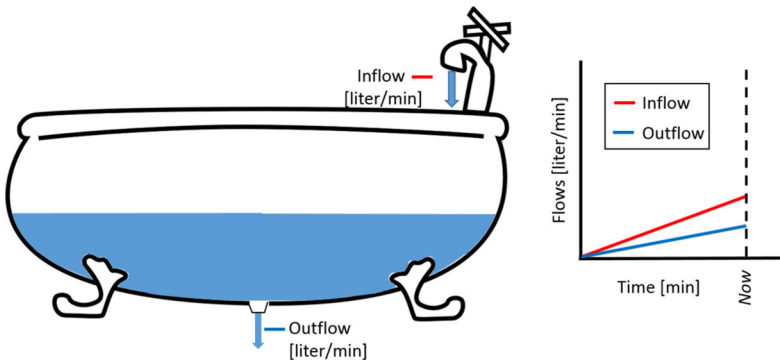


Figure Tub. Illustrative figure and graph of a bathtub with inflow (solid red line) and outflow (solid blue line) in liter/min.

Fig. 3 A description of task T2B, directing participants’ attention towards the principles of accumulation in a bathtub context. T2B was (like T2A) designed to have a lower knowledge demand compared with the main SF task: it (only) assesses declarative and procedural knowledge of accumulation

positively correlated with performance (see the [supplementary material](#)) but was not analyzed further because it is outside the scope of this study.

3.2 Efficacy of the interventions

For the full sample, the success rate on the main SF task was 21% in T1 and 28% in T3, after the alternative tasks, serving as interventions (see Table 2). This difference is not statistically significant ($p = 0.14$). Only one of the three interventions had a weakly statistically significant ($p = 0.08$) impact on the participants' performance on the main SF task: the alternative task that directed the participants' attention towards the principles of accumulation in the bathtub context (T2B). The task (T2C) that involved reading about the bathtub as an analogy for atmospheric CO₂ accumulation (see the [supplementary material](#)) did not improve the participants' success rate on the main SF task, even though it presented them with the knowledge needed to answer the task, using both text and visuals.

3.3 Ways of reasoning

Five different ways of reasoning when answering the SF tasks (from answers on task T4) were identified, and these could be grouped into three main categories: *system* reasoning (with three subcategories), *pattern* reasoning, and *phenomenological* reasoning. These reflect different mental representations of the tasks (and possibly different levels of ambition in dealing with the tasks). Below, we describe what the participants focused on when using a certain way of reasoning, with Table 3 showing the frequency of responses that were classified to belong to the different categories of reasoning and some illustrative quotes for the different ways of reasoning.

Participants who used *system* reasoning focused on the system in terms of a relationship between emissions and uptake. We identified three different ways of conceptualizing this relationship:

1. *Conservation of mass*, which correctly posits that emissions must equal uptake for CO₂ stabilization
2. *No accumulation*, which incorrectly posits that the difference between emissions and uptake must be constant for CO₂ stabilization. Some participants claimed that the amount of CO₂ in the atmosphere is equal to the annual difference between emissions and uptake (i.e., $A = E - U$). Consequently, this way of reasoning does not take into account the amount of CO₂ in the atmosphere at the start of each year that remains from past years

Table 2 Share of correct answers for SF tasks and a chi-square test of homogeneity, in which statistically significant ($p < 0.1$) differences are marked in italics

Task group (N)	Share of correct answers			Chi-square homogeneity test (p values)		
	Main task, pre-test (T1)	Alternative task (T2)	Main task, post-test (T3)	T1–T2	T2–T3	T1–T3
Full sample (214)	21%	62%*	28%	<i>3E–15</i>	<i>4E–11</i>	0.14
A: CO ₂ question (74)	26%	54%	24%	<i>4E–04</i>	<i>0.0002</i>	0.85
B: bathtub question (77)	17%	70%	29%	<i>3E–11</i>	<i>2E–07</i>	<i>0.08</i>
C: bathtub description (63)	22%	–	30%	–	–	0.31

*This is the average success rate for the alternative tasks T2A and T2B together

3. *Historic debt balancing*, which incorrectly posits that emissions must go below uptake for CO₂ stabilization. According to this way of reasoning, emissions have historically been above uptake and all emitted CO₂ needs to be taken up for CO₂ stabilization (i.e., $\dot{A} = 0$ only if $\int(E) = \int(U)$)

Participants who used *pattern* reasoning inappropriately focused on matching graphical patterns between the amount of CO₂ in the atmosphere and the annual emissions or uptake. Alternatively, they focused on the notion of “stabilization,” without being explicit about in what sense.

Participants who used *phenomenological* reasoning focused on a variety of aspects of phenomena related to climate change that are not needed for solving the SF tasks.⁷ Examples of such phenomena can be found in the illustrative quotes for this way of reasoning in Table 3 but include population growth and sources of emissions and uptake. Based on these phenomena related to climate change, participants seemingly or explicitly inferred what will or should happen to emissions and uptake in the future, rather than dealing with the task as it is formulated.

3.4 Relation between ways of reasoning and answers on the main SF task

Figure 4 shows how the five ways of reasoning, identified from the answers on task T4, are related to answers on the main SF task in the post-test (T3). While some of the participants who chose the first or second (incorrect) alternatives of increasing or stable emission scenarios reasoned in terms of *no accumulation*, the majority of those who chose the second alternative used *pattern* reasoning. The vast majority of those who chose the third (correct) alternative used *conservation of mass*. The majority of those who chose the fourth (incorrect) alternative, where emissions plummet below uptake, reasoned in terms of *historic debt balancing*. In summary, Fig. 4 shows that apart from *phenomenological* reasoning—which appears in all four alternative answers—there is a dominant way of reasoning behind each alternative. The occurrence of *phenomenological* reasoning in all alternative answers in the post-test suggests that the participants struggled to create a correct mental representation of the main SF task; that is, they struggled to judge what prior knowledge is relevant for the SF task at hand.

We note that among those who managed to create or utilize a mental representation that guided them to the correct answer, only a couple used *phenomenological* reasoning. The largest shares of unclassified explanations fell into the first two answer alternatives which also had the largest shares of pattern matchers. This may indicate that an unproportionally large share of the answers for alternatives 1 and 2 is less thought through than the average answer, since the main reason for not being classified was that explanations given were too brief to be classifiable (which we reason is a sign of the tasks being given little thought) and since pattern matching is considered to be a general solution heuristic (Gilovich and Savitsky 2002) requiring little cognitive effort.

Lastly, we note that among those answering alternative 4 (in which emissions go below uptake), a higher than average number of participants were categorized into more than one way of reasoning. Most often they reasoned both about what they want to happen or what needs to

⁷ In keeping with a common use of the term (e.g., Redish 2003), we use the term *phenomenological* here to reflect the numerous references to the “real world” in this category, as opposed to the more abstract and mathematically oriented discourse in the other categories.

Table 3 The participants' answers to the open-ended task (T4) were classified into five ways of reasoning, which are summarized into three overarching categories. The frequencies reported are the fraction of the 214 answers that were classified to belong to a given category or way of reasoning. These do not sum up to 100% since some answers were classified as belonging to several ways of reasoning. The ways of reasoning are exemplified using illustrative quotes

Category/subcategory	Frequency	Illustrative quotes
System reasoning	44%*	
Conservation of mass	23%	<p>“In order to get a concentration of CO₂ stable, we want a net flow = 0, thus we want uptake = emission.”</p> <p>“For the amount to stabilise, input and outflow have to have the same value. The only graph showing this is the third one. The absolute values are irrelevant. The trend could as well be positive, providing the lines for input and outflow are coincident.”</p>
No accumulation	7.5%	<p>“The amount CO₂ in the atmosphere is dependent on inflow minus outflow. In order to stabilize the total, you need to stabilize this difference, as seen in [alternatives] 1 and 2.”</p>
Historic debt balancing	7.5%	<p>“The historical CO₂ emission shows that the difference between intake and uptake has been increasing and is getting bigger over the years. This means that in order for the level to stabilize, the intake needs to make up for all of these past bigger increases and that can only happen if over the coming years intake is inferior to uptake.”</p>
Pattern reasoning	26%	<p>“The leveling off in [alternative] 2 seems to match the graph in my answer.”</p> <p>“If CO₂ stabilizes then everything stabilizes.”</p>
Phenomenological reasoning	17%	<p>“The emissions levels will keep rising on our current course and uptake will stay the same because of deforestation and population growth.”</p> <p>“My reasoning is based on the premise that at the early stages of human existence, there was less population and less pressure on the environment because early humans were basically hunter gatherers who moved from one place to another and depended less on the environment. As the population increases there became an immediate need to sustain the growing population, accompanied by industrial revolution with increasing technology. All these resulted to a systematic increase in emission of Carbon dioxide into the atmosphere because the forest is systematically exploited, creating a scenario where the emission of carbon dioxide far exceed the absorptive capacity. Maintaining the emission capacity from now until the end of the century means that exploitation of natural resources that emits carbon dioxide will systematically be reduced, and at the same time maintain the absorptive capacity of carbon dioxide.”</p>
Not categorized	24%	

*Includes a 6% that cannot be placed into either of the three subcategories

happen in terms of human development (as opposed to in terms of emissions, uptake, and amount)—i.e., phenomenological reasoning—and about the need for emissions to go below uptake for the amount to stabilize—historic debt balancing.

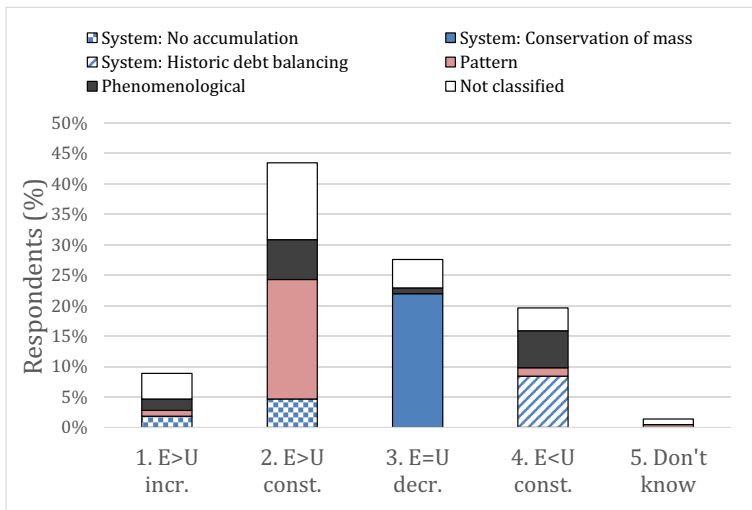


Fig. 4 Relation between ways of reasoning and answers on the post-test. Pattern reasoning is in orange and shows up in alternative 2 (which match the pattern of emissions with that of amount) as expected. Phenomenological reasoning is gray and is distributed between the different answers. The system reasoning category is marked with different patterns of blue to highlight that the answers almost exclusively belong to one of three different subcategories of system reasoning: no accumulation (small white dots), conservation of mass (chess squares), and historic debt balancing (diagonal stripes)

3.5 Relation between performance on SF tasks and stated climate policy support

The stated support for stringent climate policies was very strong in our sample (see the [supplementary material](#)), with 93% of the 167 participants that answered both the SF tasks and the climate policy question agreeing with the statement that “society should take strong action to reduce emissions of greenhouse gases today.” This clearly shows that our sample participants constitute an interested and pro-climate policy group of the general public. Given this lack of variance in stated climate policy support, we were unable to explore potential correlations between different types of knowledge (or understanding) of climate physics and stated policy support. However, these results suggest that at least the type of knowledge tested in the main SF task is not a prerequisite for stated support for stringent climate policy.

4 Discussion

4.1 Probing SF failure: knowing how and knowing when

Interestingly, but in line with our hypothesis that SF tasks with lower knowledge demands would result in higher success rates, participants performed significantly better on the SF tasks that directed their attention towards the principles of accumulation (T2A–B), compared with the main SF task (T1/T3). As Newell et al. (2013) pointed out, “Given the low-base of accurate performance in [SF tasks], any manipulation which leads to over 50% of the sample getting the answer (approximately) correct is newsworthy” (p. 3143). Our finding nuances the common finding in the literature on SF failure that most people “have difficulty relating the flows into and out of a stock to the level of the stock, even in simple,

familiar contexts such as bank accounts and bathtubs” (Sterman 2011, p. 817). Instead, we found that most participants were able to successfully solve SF tasks (T2A–B) assessing *declarative* and *procedural* knowledge of accumulation (knowing *what* and knowing *how*) but struggled with *conditional* knowledge (knowing *when*) in relation to the main SF task. To put it in simpler terms, our finding suggests that people do “understand” the principles of accumulation and how to use them but do not understand that it is this knowledge they should apply in the main SF task. This finding is in line with research on problem solving in physics, indicating that students find it difficult to create a correct mental *representation* of a new problem by combining the information provided in the problem statement with *relevant* background knowledge (Savelsbergh et al. 2002).

Yet, the idea that different kinds of SF tasks may assess different types of knowledge of accumulation seems to be largely overlooked in the literature on SF failure; there is, at least, no explicit discussion of different types of knowledge or what it means to “understand” accumulation. Indeed, we note that the high success rates on several SF tasks reported by Fischer et al. (2015) could be a result of what type of knowledge they assess, rather than the particular format (without graphs), as suggested by the authors.

4.2 Efficacy of the interventions

Only one of the three alternative tasks that directed the participants’ attention towards the principles of accumulation had a (weakly) statistically significant impact on performance on the main SF task in the post-test: the alternative task that used the bathtub analogy as context (T2B). This finding supports the notion that while analogies can be an effective teaching tool (Podolefsky and Finkelstein 2006), active learning methods, such as answering a question, are more conducive to learning compared with just reading or hearing an explanation (Freeman et al. 2014). However, the rather small improvement in the success rate for the main SF task suggests that additional scaffolding is needed to overcome the challenges inherent in the main SF task.

4.3 Ways of reasoning provide additional theoretical insights into SF failure

We identified five ways of reasoning when dealing with the main SF task, and these could be grouped into three categories: *system* reasoning, *pattern* reasoning, and *phenomenological* reasoning. These ways of reasoning provide additional theoretical insights to explain the large difference in performance between the different kinds of SF tasks. More specifically, they provide insights into what background knowledge participants drew on to create a mental representation of the main SF task. Our results therefore support the interesting hypothesis that SF failure “may be less a matter of incorrect knowledge and more a matter of incorrect problem representation” (Cronin and Gonzalez 2007, p. 15).

System reasoning consists of three subcategories, which we have termed *conservation of mass*, *no accumulation*, and *historic debt balancing*. The “no accumulation” subcategory supports the claim made by Cronin and Gonzalez (2007, p. 11) that some people “will look at the difference between the inflow and outflow when thinking about the stock [...], but they will ignore current accumulation in the stock”.

Pattern reasoning involves using the *correlation heuristic* as a problem solving strategy, “erroneously assuming that the behavior of a stock matches the pattern of its flows” (Cronin et al. 2009, p. 1). While the correlation heuristic has been forwarded as the

dominant reason for SF failure (Cronin et al. 2009), it remained an untested hypothesis until recently. As Korzilius et al. (2014) noted:

Thus far, research on stock-flow performance has focused on the outcomes of reasoning processes and inferred that individuals use correlational reasoning while estimating stock-flow behavior, assuming that the flow(s) immediately and directly affect the stock. The actual reasoning process of participants remained hidden from the researchers. [...] We may say that the correlation heuristic has the status of a hypothetical idea, a presumption that still has to be tested in research (p. 269).

Our study provides empirical evidence, both quantitative and qualitative, for the claim that people use the correlation heuristic as a problem solving strategy. In the main SF task, the answer alternative that was selected by most participants (about 45%) was the pattern matching alternative, and pattern reasoning was the most frequently used explanation for choosing this alternative. This finding is in line with previous research, demonstrating a strong tendency for pattern matching (e.g., Dutt and Gonzalez 2013; Reichert et al. 2015; Cronin et al. 2009; Sterman 2008).

To our knowledge, *phenomenological reasoning* has not been documented in the literature on SF failure. What distinguishes phenomenological reasoning from the other types of reasoning is a strong focus on the context of the SF task and various phenomena related to climate change. Previous research on SF failure has viewed *contextual* knowledge as something that might be *lacking* and hence a potential explanation for the poor performance on SF tasks (Cronin et al. 2009; Newell et al. 2013). Interestingly, in our study, the problem was rather the opposite: It is not that participants knew “too little” about the context—it is rather that they knew “too much” and got “lost in the complexity of the context,” to borrow a phrase from Eggert et al. (2017). The crux of phenomenological reasoning is echoed in an observation made by the Spanish novelist Pérez-Reverte (1998):

There are no innocent readers anymore. Each overlays the text with his own perverse view. A reader is the total of all he’s read, in addition to all the films and television he’s seen. To the information supplied by the author he’ll always add his own. And that’s where the danger lies: An excess of references (p. 335).

By unearthing several such “references” and putting phenomenological reasoning next to the other ways of reasoning, we provide novel insights into climate change *domain-specific* challenges related to solving the kind of SF task used by Sterman and Booth Sweeney (2007).

Our findings have important implications for teaching and climate change communication. First of all, it is unlikely that a single learning activity or explanation will help all people—with their different ways of reasoning—to understand atmospheric CO₂ accumulation. People using no accumulation reasoning need help to realize that the CO₂ that was present last year does not magically disappear, so to speak. Those using historic debt balancing would likely benefit from being reminded that we are opting for stabilizing the CO₂ amount at a higher level (compared with pre-industrial times) and that if all CO₂ emitted by humans (since industrialization) were taken up, we would fall back to pre-industrial atmospheric CO₂ levels. People using phenomenological reasoning, and potentially also those using pattern matching, would likely benefit from having a guided step-by-step comparison of the carbon cycle with a carefully chosen analogical system. This could help them focus on the principles of accumulation. Having been told or reminded of how the principles work in a contained and familiar analogical context, the learners should have a

chance to follow an assisted transfer of knowledge back to the CO₂ context. This may help them realize how the principles are applicable in the climate context which by itself may previously have caused them to lose track of their reasoning around accumulation.

A limitation of the thematic analysis presented here was the brevity of the answers provided by most participants to the open-ended question. Thus, a next step could be to conduct semi-structured interviews with a smaller sample to explore in more detail what conceptual and mathematical difficulties people experience when dealing with SF tasks that assess different types of knowledge. Investigating deeper psychological mechanisms behind the different ways of reasoning identified in this study is also a possible next step. The substantial fraction of answers which included people's attitudes about what they want to see happen suggests that how people answer and reason is affected by more than mere task-specific cognitive reasoning. A large fraction of the participants seems to have unconsciously substituted the cognitively demanding SF task with a simpler question and answered that question instead—what Kahneman and Frederick (2002) call *attribute substitution*. We hypothesize that attribute substitution may explain why people tend to use pattern reasoning and phenomenological reasoning, and thus an inappropriate mental representation of the SF tasks.

4.4 Link between knowledge and stated policy support

Our results clearly demonstrate that performing well on the main SF task is not a necessary condition for stated support for climate policy. This should perhaps come as no surprise, given the extensive evidence that there is a host of other factors, beyond knowledge, that influence people's attitude and behavior in relation to climate change mitigation and adaptation, such as values, social norms, science skepticism and literacy, and political orientation (Hornsey et al. 2016; Hamilton et al. 2015; Gifford 2011; Wibeck 2014).

On the other hand, in no way do our results rule out that a better understanding of (some aspects of) climate science could affect support for climate policy or that understanding could be important for actual (or revealed) climate policy support. The existing evidence on the connections between climate science literacy and climate policy support does show that greater understanding of climate science correlates with greater *belief in* or *acceptance of* climate change (Hornsey et al. 2016; Guy et al. 2014; Ranney and Clark 2016) and that greater belief in turn is associated with stronger support for climate policy (Hornsey et al. 2016), though the latter effect is relatively small. Hence, we agree with Eggert et al. (2017) who argue that conceptual understanding of climate physics “is an important prerequisite to change individuals' attitudes towards climate change and thus to eventually foster climate literate citizens” (p. 137).

A key question—related to the main focus of this study—is what (type of) knowledge has the largest potential to leverage climate policy support. For instance, the Climate Literacy Framework presented by the US Global Change Research Program lists no less than 39 points that climate literate citizens should know in order to make informed decisions on climate change; a better understanding of which of these points are more important for fostering support for climate policies would help promote more effective climate change communication. The results presented by Shi et al. (2016) show that there can be differences in how knowledge in different domains of climate science—such as basic physics, causes, and impacts—can affect attitudes to climate risks. However, this and other studies on the links between climate literacy and concerns have solely focused on different facets of declarative knowledge (i.e., climate science facts). The results presented in this study suggest that it would

also be interesting to further explore the relationship between other types of knowledge (procedural and conditional) and climate policy support.

5 Conclusions

The question of whether people understand atmospheric CO₂ accumulation is not as simple as it seems. This mixed methods study of public understanding of atmospheric CO₂ accumulation and stated climate policy support extends previous research on SF failure by showing that:

- Seemingly similar SF tasks may assess different types of knowledge, and people perform significantly better on tasks assessing declarative and procedural knowledge compared with tasks assessing conditional knowledge
- When faced with a climate SF task, most people use one of three overarching ways of reasoning: system reasoning, pattern reasoning, and phenomenological reasoning
- System reasoning took on three different forms which we name conservation of mass, no accumulation, and historic debt balancing. These three different ways of reasoning suggest that the system was treated using three distinctly different mental representations

Taken together, our findings show that SF failure can be due to the use of inappropriate mental representations of SF tasks rather than a poor understanding of the principles of accumulation. This calls for both a more nuanced discussion on how to promote understanding of climate science and a more detailed exploration of the links between different (types) of climate science knowledge and climate policy support.

Acknowledgments The authors wish to thank Matthew A. Cronin and an anonymous reviewer for valuable feedback on the manuscript.

Funding information Funding from the Swedish Energy Agency, Adlerbertska, and Chalmers University of Technology undergraduate education is gratefully acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alexander PA, Schallert DL, Hare VC (1991) Coming to terms: how researchers in learning and literacy talk about knowledge. *Rev Educ Res* 61:315–343
- Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK (2010) *How learning works: seven research-based principles for smart teaching*. Wiley
- Archer D, Brovkin V (2008) The millennial atmospheric lifetime of anthropogenic CO₂. *Clim Chang* 90:283–297
- Biggs JB (2003) *Teaching for quality learning at university: what the student does*. McGraw-hill education (UK)
- Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qual Res Psychol* 3:77–101
- Chen X (2011) Why do people misunderstand climate change? Heuristics, mental models and ontological assumptions. *Clim Chang* 108:31–46
- Cronin MA, Gonzalez C (2007) Understanding the building blocks of dynamic systems. *Syst Dyn Rev* 23:1–17
- Cronin MA, Gonzalez C, Sterman JD (2009) Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organ Behav Hum Decis Process* 108:116–130

- de Jong T, Ferguson-Hessler MG (1996) Types and qualities of knowledge. *Educ Psychol* 31:105–113
- Dryden R, Morgan MG, Bostrom A, de BWB (2018) Public perceptions of how long air pollution and carbon dioxide remain in the atmosphere. *Risk Anal* 38:525–534
- Dutt V, Gonzalez C (2009) Human “mis”-perceptions of climate change. In: Proceedings of the human factors and ergonomics society annual meeting. SAGE Publications, pp 384–388
- Dutt V, Gonzalez C (2012a) Decisions from experience reduce misconceptions about climate change. *J Environ Psychol* 32:19–29
- Dutt V, Gonzalez C (2012b) Human control of climate change. *Clim Chang* 111:497–518
- Dutt V, Gonzalez C (2013) Reducing the linear perception of nonlinearity: use of a physical representation. *J Behav Decis Mak* 26:51–67
- Eggert S, Nitsch A, Boone WJ, Nückles M, Bögeholz S (2017) Supporting students’ learning and socioscientific reasoning about climate change—the effect of computer-based concept mapping scaffolds. *Res Sci Educ* 47(1):137–159
- Fischer H, Degen C, Funke J (2015) Improving stock-flow reasoning with verbal formats. *Simul Gaming* 46:255–269
- Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP (2014) Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci* 111(23):8410–8415
- Gifford R (2011) The dragons of inaction: psychological barriers that limit climate change mitigation and adaptation. *Am Psychol* 66:290–302
- Gillis J (2017) Carbon in atmosphere is rising, even as emissions stabilize. *The New York Times*
- Gilovich T, Savitsky K (2002) Like goes with like: the role of representativeness in erroneous and pseudo-scientific beliefs
- Goodwin P, Katavouta A, Roussenov VM et al (2018) Pathways to 1.5 °C and 2 °C warming based on observational and geological constraints. *Nat Geosci* 11:102–107
- Guy S, Kashima Y, Walker I, O’Neill S (2013) Comparing the atmosphere to a bathtub: effectiveness of analogy for reasoning about accumulation. *Clim Chang* 121:579–594
- Guy S, Kashima Y, Walker I, O’Neill S (2014) Investigating the effects of knowledge and ideology on climate change beliefs. *Eur J Soc Psychol* 44:421–429
- Hamilton LC, Hartter J, Saito K (2015) Trust in scientists on climate change and vaccines. *SAGE Open* 5(3)
- Hornsey MJ, Harris EA, Bain PG, Fielding KS (2016) Meta-analyses of the determinants and outcomes of belief in climate change. *Nat Clim Chang* 6:622
- Houghton JT, Ding Y, Griggs DJ et al (2001) *Climate change 2001: the scientific basis*. The Press Syndicate of the University of Cambridge
- Jones N, Ross H, Lynam T et al (2011) Mental models: an interdisciplinary synthesis of theory and methods. *Ecol Soc* 16
- Kahneman D, Frederick S (2002) Representativeness revisited: attribute substitution in intuitive judgment. In: Heuristics and biases: the psychology of intuitive judgment, pp 49:49–49:81
- Korzilius H, Raaijmakers S, Rouwette E, Vennix J (2014) Thinking aloud while solving a stock-flow task: surfacing the correlation heuristic and other reasoning patterns. *Syst Res Behav Sci* 31(2):268–279
- Meinshausen M, Smith SJ, Calvin K et al (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim Chang* 109:213–241
- Morgan MG, Fischhoff B, Bostrom A, Atman CJ (2002) *Risk communication: a mental models approach*. Cambridge University Press
- Moxnes E, Saisel AK (2009) Misperceptions of global climate change: information policies. *Clim Chang* 93:15–37
- Newell BR, Kary A, Moore C, Gonzalez C (2013) Managing our debt: changing context reduces misunderstanding of global warming. In: 35th annual meeting of the Cognitive Science Society (CogSci 2013). pp 3139–3144
- Newell BR, Kary A, Moore C, Gonzalez C (2016) Managing the budget: stock-flow reasoning and the CO2 accumulation problem. *Top Cogn Sci* 8:138–159
- Pérez-Reverte A (1998) *The Club Dumas*. Vintage International, New York
- Peters GP, Andrew RM, Boden T et al (2012) The challenge to keep global warming below 2 °C. *Nat Clim Chang* 3:4–6
- Podolefsky NS, Finkelstein ND (2006) Use of analogy in learning physics: the role of representations. *Physical Review Special Topics-Physics Education Research* 2(2):020101
- Ranney MA, Clark D (2016) Climate change conceptual change: scientific information can transform attitudes. *Top Cogn Sci* 8:49–75
- Redish EF (2003) *Teaching physics with the physics suite*. Wiley, Hoboken
- Reichert C, Cervato C, Niederhauser D, Larsen MD (2015) Understanding atmospheric carbon budgets: teaching students conservation of mass. *J Geosci Educ* 63:222–232
- Savelsbergh ER, De Jong T, Ferguson-Hessler MG (2002) Situational knowledge in physics: the case of electrodynamics. *J Res Sci Teach* 39(10):928–951

- Shi J, Visschers VH, Siegrist M, Arvai J (2016) Knowledge as a driver of public perceptions about climate change reassessed. *Nat Clim Chang* 6:759
- Sterman JD (2008) Risk communication on climate: mental models and mass balance. *Science* 322:532–533
- Sterman JD (2011) Communicating climate change risks in a skeptical world. *Clim Chang* 108:811
- Sterman JD, Booth Sweeney L (2002) Cloudy skies: assessing public understanding of global warming. *Syst Dyn Rev* 18:207–240
- Sterman JD, Booth Sweeney L (2007) Understanding public complacency about climate change: adults' mental models of climate change violate conservation of matter. *Clim Chang* 80:213–238
- Sterner EO, Johansson DJ (2017) The effect of climate–carbon cycle feedbacks on emission metrics. *Environ Res Lett* 12:034019
- Turns SR, Van Meter PN (2011) Applying knowledge from educational psychology and cognitive science to a first course in thermodynamics. In: ASEE annual conference and exposition, conference proceedings
- Wibeck V (2014) Enhancing learning, communication and public engagement about climate change – some lessons from recent literature. *Environ Educ Res* 20:387–411

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Erik O. Sterner¹ · Tom Adawi² · U. Martin Persson¹ · Ulrika Lundqvist¹

Tom Adawi
tom.adawi@chalmers.se

U. Martin Persson
martin.persson@chalmers.se

Ulrika Lundqvist
ulrika.lundqvist@chalmers.se

¹ Department Space, Earth and Environment, Chalmers University of Technology, 412 96 Gothenburg, Sweden

² Department Communication and Learning in Science, Chalmers University of Technology, 412 96 Gothenburg, Sweden