



Managing the Budget: Stock-Flow Reasoning and the CO₂ Accumulation Problem

Ben R. Newell,^a Arthur Kary,^a Chris Moore,^a Cleotilde Gonzalez^b

^a*School of Psychology, University of New South Wales*

^b*Department of Social and Decision Sciences, Carnegie Mellon University*

Received 1 July 2014; received in revised form 13 July 2015; accepted 13 July 2015

Abstract

The majority of people show persistent poor performance in reasoning about “stock-flow problems” in the laboratory. An important example is the failure to understand the relationship between the “stock” of CO₂ in the atmosphere, the “inflow” via anthropogenic CO₂ emissions, and the “outflow” via natural CO₂ absorption. This study addresses potential causes of reasoning failures in the CO₂ accumulation problem and reports two experiments involving a simple re-framing of the task as managing an analogous *financial* (rather than CO₂) budget. In Experiment 1 a financial version of the task that required participants to think in terms of controlling *debt* demonstrated significant improvements compared to a standard CO₂ accumulation problem. Experiment 2, in which participants were invited to think about managing *savings*, suggested that this improvement was fortuitous and coincidental rather than due to a fundamental change in understanding the stock-flow relationships. The role of graphical information in aiding or abetting stock-flow reasoning was also explored in both experiments, with the results suggesting that graphs do not always assist understanding. The potential for leveraging the kind of reasoning exhibited in such tasks in an effort to change people’s willingness to reduce CO₂ emissions is briefly discussed.

Keywords: Stock-flow reasoning; CO₂ accumulation; Correlation heuristic; Climate change

1. Introduction

There is overwhelming agreement among climate scientists that the globe is warming up, due in large part to increases in the anthropogenic emissions of greenhouse gases (e.g., Anderegg, Prall, Harold, & Schneider, 2010; Cook et al., 2013). Despite this consensus in the scientific community, highly divergent opinions about the existence and

Correspondence should be sent to Ben R. Newell, School of Psychology, University of New South Wales, Sydney 2052, Australia. E-mail: ben.newell@unsw.edu.au

implications of global warming remain entrenched in the wider community in many countries (e.g., Leiserowitz, Smith, & Marlon, 2010; Leviston, Walker, & Morwinski, 2013; Weber & Stern, 2011).

The manifold explanations for this “disconnect” between the science and popular belief range from fundamental differences in the way people “view” the world (e.g., Kahan et al., 2012), to the attempts to manufacture doubt in the science (e.g., Oreskes & Conway, 2010), to deficits in lay understanding of the mechanisms of global warming (e.g., Bord, O’Connor, & Fisher, 2000; Ranney, Clark, Reinholz, & Cohen, 2012; Sterman, 2008). Here we focus on this last issue and examine how a change in framing the CO₂ accumulation problem might aid understanding of some of the basic science behind how global warming “works.”

Our research builds on studies that have examined people’s understanding of the relationship between the amount of CO₂ emitted into the atmosphere (“inflow”), the amount of CO₂ absorbed via natural processes (“outflow”), and the resulting “stock” or concentration of CO₂ that accumulates in the atmosphere (e.g., Dutt & Gonzalez, 2012a,b; Guy, Kashima, Walker, & O’Neill, 2013; Moxnes & Saisel, 2009; Sterman & Booth-Sweeney, 2007). The general consensus emerging from this and related work (e.g., Cronin, Gonzalez, & Sterman, 2009) is that lay people’s understanding of these relationships is poor. Moreover, some researchers have argued that misconceptions about the basic rules governing this aspect of the carbon cycle could lead to unwarranted complacency with regard to policies for addressing climate change (e.g., Sterman, 2008).

In the context of this growing literature, we examine two main questions: (1) Can a simple reframing of the relationships in terms of financial rather than CO₂ “budget” management aid understanding of stocks and flows? And (2) Do graphical representations of these relationships help or hinder understanding?

1.1. Reframing the message

Sterman and Booth-Sweeney (2007) provided evidence that despite the apparent simplicity of stock-flow dynamics, even highly educated adults have great difficulty in completing tasks that require understanding basic relationships. Their participants were asked to sketch a pattern of emission and removal trajectories across time that would lead to a gradual rise followed by the stabilization of atmospheric CO₂. In the scenario this required a rapid reduction in emissions to achieve convergence with the rate of removal. However, the vast majority of their participants sketched emissions trajectories that either maintained current levels or increased them—thereby preventing stabilization from occurring. A common explanation of such erroneous responding is over-reliance on the use of a “correlation heuristic” whereby participants reason that the output of a system should “look like” (be positively correlated with) its inputs (Cronin et al., 2009; Sterman, 2008). Thus, because the concentration (output) gradually rises and then stabilizes, people think emissions (inputs) should follow a similar pattern.

In an effort to improve performance on the CO₂ task some researchers have turned to analogies and different frames. Analogies can help people integrate the presumably

unfamiliar information about the climate system into existing knowledge structures (Dutt & Gonzalez, 2012a,b; Gonzalez & Wong, 2012; Guy et al., 2013; Moxnes & Saisel, 2009). This work shows some promise, but there is still considerable room for improvement. For example, Guy et al. (2013) provided a direct empirical test of the usefulness of a “bathtub analogy” in the CO₂ task (see Sterman & Booth-Sweeney, 2007) in which participants were invited to think about the stock of CO₂ as water accumulating in a bathtub due to in-flow from a tap (emissions) and out-flow via the plug-hole (absorptions). However, even when provided this analogical context only 17% (Experiment 1) and 11% (Experiment 2) of the participants completed the task correctly.

In recent work (Newell, Kary, Moore, & Gonzalez, 2013) we adopted a different analogical context in an effort to improve reasoning in the CO₂ task. We were inspired by discussions of the climate problem in terms of the management of a carbon budget (Rau-pach et al., 2011)—the idea that we have a limited amount of carbon to “spend” (burn) before we run the risk of dangerous climate change—and presented the problem in an analogous financial debt management context.

We hypothesized that many people may be able to intuit that if they spend more than they earn then they will get into debt, and that if they keep spending more than they earn that debt will continue to increase. In contrast, if they spend less or the same amount as they earn, then debt will reduce or remain constant (assuming no interest). This kind of debt management involves sequential monitoring of income and expenditure and is something that many of us grapple with across time (“I will pay off that credit card by the end of year!”). This situation contrasts with the bathtub in which we typically only have two interactions with the system—turning the tap on and then off—and in which the out-flow is usually zero (because we tend to run a bath with the plug in).¹ We thus predicted that performance on a financial “debt management” stock-flow task would be superior to the CO₂ task due in large part to the increased familiarity with the core concepts of the problem (though see Brunstein, Gonzalez, & Kanter, 2010). In addition, we predicted that participants who were given the CO₂ task but had an opportunity to think about the problem in the analogous financial management context would be more accurate than those just given the CO₂ task. This prediction was based on literature showing that the process of comparing analogous contexts in which concepts and relationships have clear one-to-one mappings (e.g., emissions across time = spending across time; CO₂ concentration = accumulated debt) improves reasoning (e.g., Gentner, Loewenstein, & Thompson, 2003).

Our preliminary findings using this reframing were promising with some evidence that inviting participants to think about the CO₂ problem in terms of financial *debt* management led to improvements in predicting the trajectory of future emissions (Newell et al., 2013). Here we seek to replicate and extend those findings in two ways. First, we examine an issue that has received some recent discussion in the literature, namely whether providing graphical representations of the relations between stocks and flows hinders or helps performance (e.g., Cronin et al., 2009; Guy et al., 2013; Sterman & Booth-Sweeney, 2007). Second, in Experiment 2, we examine whether the concepts involved in budget

management more generally confer advantages in stock-flow reasoning by exploring a financial *savings* management analogy.

1.2. The role of graphs

In their original examination of stock-flow reasoning in the CO₂ task, Sterman and Booth-Sweeney (2007) included a condition that omitted graphical representations of the problem and simply required participants to respond to multiple-choice questions. They found similar levels of (poor) performance irrespective of response mode. In a similar vein, Cronin et al. (2009) found that stock-flow reasoning was poor and unaffected in a non-climate-related stock-flow task (the Department Store Task) regardless of whether information was presented in line graphs, bar graphs, tables of numbers, or as a narrative. In contrast to this backdrop of null results, Guy et al. (2013) argued that the adoption of the erroneous correlation heuristic response in the CO₂ task was *promoted* by showing participants potentially misleading (and complex) graphs. They found that presenting the same information minus the graph led to improved accuracy, especially when the absence of a graph was combined with the presence of the bathtub analogy. This finding echoes Cronin and Gonzalez's (2007) conclusion that salient features of a graph in accumulation problems similar to the CO₂ task may bias people's judgments of accumulation. These contrasting patterns of results imply that the jury is still out with regard to the role graphs play in aiding or abetting stock-flow reasoning.

The temptation to adopt a correlation response can be readily inferred by looking at the kinds of graphs used in the CO₂ task (e.g., Guy et al., 2013; Sterman & Booth-Sweeney, 2007). Fig. 1 displays a scenario (leftmost top graph) in which the stock of CO₂ accumulates across time and then stabilizes over the final two decades (far right of the leftmost top graph). The bottom left graph shows the absorptions remaining constant across time (green line) and the emissions rising up to the year 2000. The task facing participants is to complete the trajectory of this emissions line (Newell et al., 2013; Sterman & Booth-Sweeney, 2007) or estimate its endpoint (Guy et al., 2013; current experiments) so that the concentration depicted in the top graph is achieved; specifically, so that the concentration stabilizes by the final period. The two panels on the right of the figure show how, using the same numbers, the CO₂ task can be re-described as a financial debt management problem (see Fig. 1 caption for more details).

Guy et al. (2013) compared conditions in which participants were given either no graph or a version of the emissions graph (lower left panel of Fig. 1). They argued that performance was poorer in the emissions graph condition because the rising trajectory invites participants to extrapolate the trend, thereby giving an estimate that is higher than the correct answer of 40GtC (e.g., Lewandowsky, 2011). Note, however, that correlation heuristic responding is typically inferred in stock-flow tasks when participants complete the bottom left graph in Fig. 1 by *copying the pattern* depicted in the top left graph—a gradual rise and then flattening out of emissions—that is, when participants reason that the output of a system should *correlate* with its input (e.g., Cronin et al., 2009; Newell et al., 2013; Sterman & Booth-Sweeney, 2007).

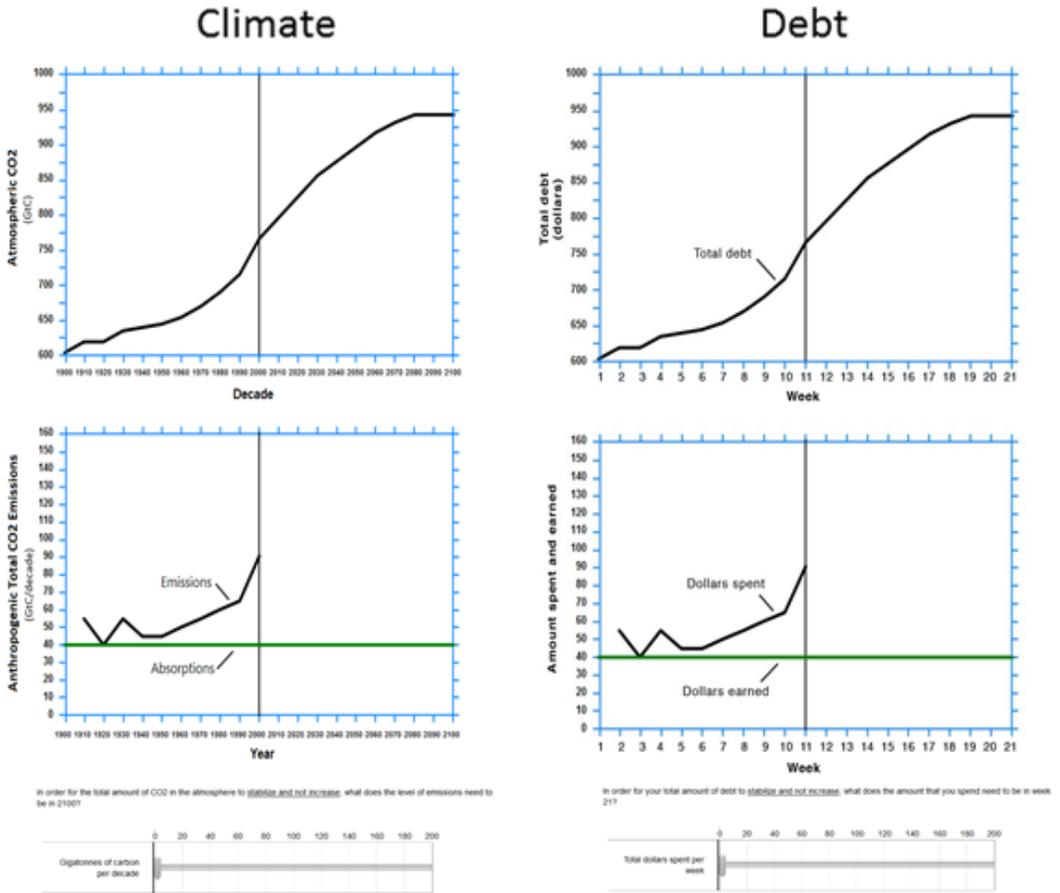


Fig. 1. Screenshots showing the graphs in the “Both Graphs” condition of the Climate (left panels) and Debt (right panels) conditions. The upper graphs show accumulated stock of CO₂ (left) or monetary debt (right) and the lower panels show the emissions and absorptions of CO₂ (left) and the spending and earning of money (right). At the bottom of each panel is the slider scale participants used to indicate the value that the emissions/spending line needs to be at by the last time period to achieve the stabilization depicted in the upper graph of each scenario. The correct answer for both tasks is 40 GtC/year (Climate) or \$40/week (Financial). This is the value at which the emissions and absorptions trajectories or the spending and earning trajectories converge, thereby leading to a stabilization in CO₂ concentration or debt. A correlation heuristic-consistent response would be ~140 (GtC or \$), which is where a participant may extrapolate the emissions/spending line to where it would be if they were copying the trajectory of the CO₂ stock/accumulated debt shown in the upper graph. Note that participants saw *either* the climate graphs (left panel) *or* the financial graphs (right panel), and participants in the Climate & Debt condition only answered the CO₂ task. In Experiment 2, the Debt task was converted to a Savings task by relabeling the inflows, outflows, and accumulation lines. (See text for further details and Appendix A for the instructions given to participants in the Climate & Debt condition of Experiment 1 and the Savings Frame of Experiment 2.)

In Experiment 1 we sought to clarify the role played by graphical representations in the CO₂ task by comparing conditions in which participants were given both graphs, only the accumulation graph (top panel), only the emissions graph (lower panel), or neither. If we find a pattern of results similar to those of Guy et al. (2013), then we will see the most accurate performance when neither graph is presented; what is less clear is whether presenting both graphs—in which arguably both the initial rising trajectory of the emissions line in the bottom graph and the shape of the accumulated stock in the top graph promote incorrect responding—will lead to poorer performance than when either of those graphs are presented alone. It is also possible that graphs will have different effects depending on the context of the stock-flow reasoning task. For example, if participants find the debt management task more familiar and thus easier to reason about (e.g., Newell et al., 2013), it may not matter whether any graphical relationships are presented.

1.3. Overview of Experiment 1

In summary, our aim in Experiment 1 was to examine the combined impact of providing an analogous financial context and the presence/absence of graphical information on reasoning in the CO₂ accumulation problem. To this end, we used a fullyfactorial, between-subjects design in which the context of the accumulation problem was crossed with the mode of information presentation. Specifically, participants were asked to solve one of the following: the CO₂ accumulation problem “alone” (closely modeled after Sterman & Booth-Sweeney, 2007), the debt management problem “alone” (modeled on Newell et al., 2013), or a CO₂ accumulation problem in which additional information about the stock-flow relationships was provided in terms of the debt management problem (e.g., “emissions are like spending,” “removal is like earning”—see Method for more details). Note that in this third condition participants still only solved the CO₂ task, but the financial context was provided as a kind of scaffold in the hope that it would assist reasoning (cf., Newell et al., 2013). As described above, these three conditions were crossed with the four graphical ones (both graphs, neither graph, emission graph only, accumulation graph only) to create 12 between-subject groups.

2. Experiment 1

2.1. Participants

A total of 481 U.S.-based participants were recruited through Amazon Mechanical Turk and randomly allocated to conditions. Twenty-three participants were excluded because they failed to answer at least one of two embedded attention check questions. This left 458 participants (177 females) with a mean age of 32.1 ($SD = 10.9$). Individual cell sizes ranged from 34 to 42 participants. Participants were paid \$0.50 for completing the experiment. The average completion time was 8.5 min.²

2.2. Design and materials

The experiment used a 3 (Climate, Debt, Climate & Debt) \times 4 (No graph, Both Graphs, Emissions only, Accumulation only) design. The graphical stimuli were adapted from Sterman and Booth-Sweeney (2007) and Newell et al. (2013).

In the No Graph conditions, participants were presented with verbal descriptions of their stock-flow scenario in accordance with their framing condition. These verbal descriptions contained the key information from the graphs, including the starting and ending points of the accumulation, emissions, and absorption trajectories. The preamble information about climate change and the instructions were based closely on Sterman and Booth-Sweeney (2007). In the Both Graphs conditions, participants were presented with the same verbal description but were also referred to the emissions and accumulation graphs as sources of the numbers. In the Emissions only and Accumulation only conditions, only the relevant graph was presented along with the full verbal description of the problem. In all conditions, below the description and/or graphs was a response slider, ranging from 0 to 200 GtC or 0 to 200 dollars. Participants were asked to respond with the value that was required for the CO₂ or the debt to “stabilize and not increase” by the end of the scenario. The correct value was 40GtC or \$40. (See Appendix A and Figure A1 for an example of the full set of instructions from the Both Graphs group of the Climate & Debt condition. Fig. 1 displays the graphs used in the Climate condition and the Debt condition, and provides an explanation of the correct response.)

2.3. Procedure

After filling in demographic information, participants in the Climate and the Climate & Debt conditions were presented with an introduction about climate change (based on material in Sterman & Booth-Sweeney, 2007), which explained how CO₂ accumulates when emissions into the atmosphere are greater than absorptions by natural processes. Those in the Climate & Debt condition were then given additional information inviting them to think about this CO₂ accumulation problem in the context of controlling one’s personal debt. Participants in both conditions were then presented with the particular experimental scenario and the response slider according to their condition. If appropriate (i.e., in those conditions that included graphs), relationships were explained with reference to those graphs. Following the explanation, participants in the Climate & Debt condition were presented with three additional multiple-choice mapping questions quizzing them on the relationship between the key variables in the debt and the CO₂ scenario (e.g., accumulation of CO₂ in the atmosphere is equivalent to: amount of debt, amount you spend, or amount you earn? see Appendix A for full wording) and were given the correct answers before being asked to respond to the CO₂ accumulation task. Participants in the Debt Condition were only given information about personal debt management and only asked to solve the debt-accumulation problem; climate change and CO₂ were not mentioned.³

2.4. Results

There are two potential dependent measures one could analyze: (a) mean estimates of emissions or spending values and (b) the proportion of participants in each condition who provided the exactly correct answer (40) (i.e., a binary correct/incorrect analysis, cf., Guy et al., 2013).

We chose to focus on (i), partly for brevity and to accommodate reviewers' requests, but largely because it provides an indication of the *level of (in)accuracy* across conditions. Arguably, being approximately correct in these tasks is sufficient to demonstrate some basic understanding of the underlying relationships. Analyses of the mean estimates allows conclusions to be drawn about which contexts are likely to be generally useful in communication settings—(i.e., by reducing wildly inaccurate estimates)—even if they do not point to a “magic bullet” solution that leads everyone to the correct answer. Such subtleties are potentially lost if one only examines exactly correct answers. Analysis of the binary data—which provides a broadly consistent qualitative pattern to the mean estimates—is available on request from the authors.

Fig. 2 displays a violin plot of the estimated emissions (Climate and Climate & Debt conditions) or spending values (Debt Condition) required for stabilization, for each cell

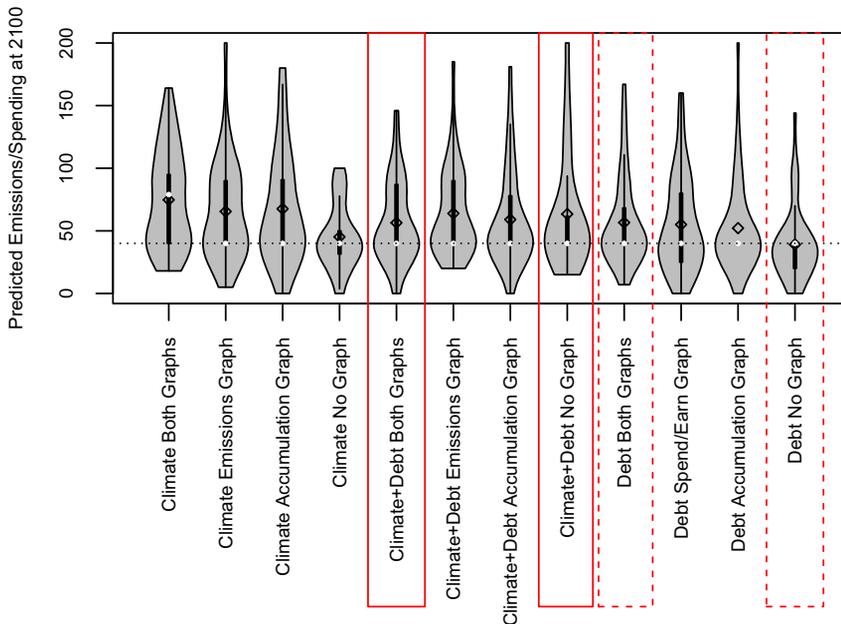


Fig. 2. A violin plot showing the predicted emissions/spending in each of the 12 cells of Experiment 1. The white dot is the median, and the open diamond is the mean; the box encompasses the 25%–75% interquartile range, and the gray areas either side of the center line of each “violin” are rotated kernel density estimation plots. The gray dotted line at 40 on the y-axis indicates the correct answer. (To facilitate comparison of common conditions across Experiments 1 and 2, the relevant violins are highlighted with red boxes—solid lines for climate and debt conditions, dashed lines for debt alone.)

of the design. Violin plots are similar to box plots, except that they also show the probability density of the data at different values. Thus, the gray shading on each side of the center line of each “violin” is a rotated kernel density estimation indicating the probability that the data lies at the values indicated on the *Y*-axis. The figure suggests that with the exception of participants given both graphs in the climate context (i.e., the “standard” Sterman & Booth-Sweeney, 2007 version), the median response (the white dot inside each “violin”) is very close to the correct value of 40 (indicated by the dashed line running horizontally across the figure). Although the uniformity in median scores might at first glance appear surprising, given that for every condition there is only one (same) correct answer, it only requires a subset of participants in each condition to “get it” for this pattern to arise.

A 3×4 ANOVA on the mean estimates (open diamonds in the figure) found a main effect of analogical context, $F(2, 446) = 5.09$, $p = .0065$, $\eta_p^2 = 0.022$, and a main effect of graph type, $F(3, 446) = 2.89$, $p = .0101$, $\eta_p^2 = .019$. A post hoc Scheffé contrast found that participants in the Climate ($M = 63.55$, $SD = 38.68$) and Climate & Debt conditions ($M = 60.75$, $SD = 37.84$) provided estimates of CO₂ emissions required for stabilization that were, on average, higher (and thus further from the correct answer of 40GtC) than those in the Debt condition who were asked to predict the level of spending needed to stabilize financial debt ($M = 50.78$, $SD = 34.88$), $F(2, 446) = 9.96$, $p < .001$. However, the comparison between emissions estimates given in the Climate & Debt and the Climate groups was not significant, $F(2, 446) = 0.365$, $p = .306$.

A post hoc Scheffé contrast following up on the graph main effect found that participants presented with Both Graphs ($M = 62.58$, $SD = 36.21$), the Accumulation Graph ($M = 59.17$, $SD = 39.38$), or the Emissions Graph ($M = 61.50$, $SD = 35.90$) all provided significantly higher estimates than those in the No Graph condition ($M = 49.82$, $SD = 36.43$), $F(3, 446) = 9.136$, $p < .0001$. The interaction between graph type and analogical context was not significant.

2.5. Discussion

The results of Experiment 1 suggest support for our contention that participants find the concepts of debt management more familiar and easier to reason about and thus when given the debt task alone they fare better than participants given the climate task alone. However, there was no evidence for the idea that supplementing the climate task with the debt context “scaffold” improved performance relative to the climate task alone. This latter finding contrasts with the preliminary evidence reported in Newell et al. (2013), suggesting that perhaps the earlier finding was not robust. The results also suggest that providing any type of graph leads to less accurate responses than providing no graphs at all (cf. Guy et al., 2013). There is no suggestion of differences in the extent of correlation heuristic-consistent responding as a function of *different* types of graph—rather just an effect of the presence or absence of *any* graph(s).

Although perhaps rather sobering in terms of the potential for improving climate change communication, this pattern of results does raise a question about why the debt

framing *alone* aids reasoning. We conjectured that the advantage comes from mere familiarity of the concepts in the financial context and expectations this created. But familiarity with the context does not always lead to improvements (e.g., Brunstein et al., 2010), so perhaps there is something else at play. One possibility is that the debt management “works” simply because people think that “debt is bad” and that to reduce debt, spending needs to be reduced, too. Thus, the fact that participants provide lower estimates for spending may be simply coincidental given this reasoning; it does not necessarily imply that they understand the structural nature of the stock-flow relationships, and perhaps the mere “surface similarity” guides responding (cf. Gonzalez & Wong, 2012). There is some support for this notion in the results of Experiment 1. Note that in the rightmost violin of Fig. 2—the Debt No Graphs condition—the probability mass is shifted downwards, as is the box encompassing the inter-quartile range, suggesting a concentration of estimates *below* the correct response of \$40. This pattern suggests that when participants were not given a graphical reminder of the *stabilized* “stock” of debt, several of them reasoned that they should spend *less* than they were earning to stabilize their debt (rather than equalizing spending and earning). Experiment 2 was therefore designed to examine whether a financial frame in which the naïve response is to think that “up” or “more is good”—namely *savings*—would yield improvements similar to those seen with a debt frame.

3. Experiment 2

Experiment 2 was a factorial design that crossed the presence and absence of graphs (this time just with the two most extreme levels of Both Graphs and No Graphs), with the presence or absence of the CO₂ climate task and the type of financial frame—debt or savings. This yielded eight between-subjects conditions. Four of these—Climate & Debt Both Graphs, Climate & Debt No Graphs, Debt Both Graphs, and Debt No Graphs—were replications of Experiment 1. The remaining four—Climate & Savings Both Graphs, Climate & Savings No Graphs, Savings Both Graphs, and Savings No Graphs—were novel. We chose not to run the climate task alone conditions because there is little doubt (given our own and others’ prior work—e.g., Guy et al., 2013; Newell et al., 2013; Sterman & Booth-Sweeney, 2007) that we would have replicated the poor performance seen in this condition in Experiment 1. Instead, we used the data from the relevant Experiment 1 conditions to conduct planned cross-experimental comparisons (see Results).

The savings frame of the financial task was created by relabeling the variables involved in the debt task (see Fig. 1, right panel). Thus, the total accumulated debt depicted in the top, accumulation graph became “total savings,” the green “dollars earned” line in the lower graph became “dollars spent,” and the black “dollars spent” line was relabeled as “dollars earned.” The task was then described as one in which a particular savings goal needed to be reached, and that the level of earnings to achieve this goal by the final week of the timeframe had to be predicted (see Appendix A2). The basic

stock-flow relationships remained the same as in the debt frame but with a different goal: stabilizing savings, rather than debt. Thus, participants needed to realize that to stabilize savings at a given level they needed to *reduce* their earnings and match them with spending by the final period.

The comparison of the debt alone and savings alone frames allows us to draw inferences about stock-flow understanding: If the improvements we observed in the debt frame are fortuitous and due to participants predicting lower levels of spending purely because they think “debt is bad,” then we might see erroneous predictions of earnings needing to be higher than spending at the final time period because they think that “up/more is good” for savings. Examining the combination of these frames with the climate task allows us to test two further questions. First, to again examine the robustness (or lack thereof) of the improvements seen in the climate and debt analogy condition reported by Newell et al. (2013); and second to test whether even if such effects are found they are coincidental rather than due to deeper understanding.

3.1. Participants

A total of 364 U.S. participants who had not completed Experiment 1 were recruited through Amazon mechanical Turk and randomly allocated to conditions. All participants were paid US\$0.50 for participating. Participants were excluded from analysis if they failed at least 1 of 2 attention check questions, leaving 336 valid participants (132 female). Individual cell sizes varied from 34 to 47. The average completion time was 7.13 min.⁴ The mean age of the sample was 32.09 with a standard deviation of 10.59.

3.2. Design and materials

The experiment used a 2 (Climate Task Present vs. Absent) \times 2 (Debt vs. Savings Frame) \times 2 (Graphs vs. No Graphs) design. The climate task and debt material were as per Experiment 1 and we created new materials for the savings Frame (see Appendix A2 for an example of the savings frame). The procedure, including the mapping questions for those in the combined conditions, was the same as in Experiment 1.

3.3. Results

Fig. 3 shows a violin plot of participants’ estimates by condition. Visual inspection of the figure leads to one immediate stand out result: Participants in the Savings Frame Both Graphs condition (second violin from the right) show clear patterns of *overestimation*. The probability density (gray shaded area) is higher up the Y-axis relative to all other groups, as are the mean and median responses illustrating that many of the participants in this condition reasoned that earnings still needed to be much higher than spending *even when a savings goal had been reached* (i.e., savings had stabilized).

A 2 (Climate Task: Present vs. Absent) \times 2 (Financial Frame: Debt vs Savings) \times 2 (Graphs: Present vs. Absent) found a significant effect of financial framing, with

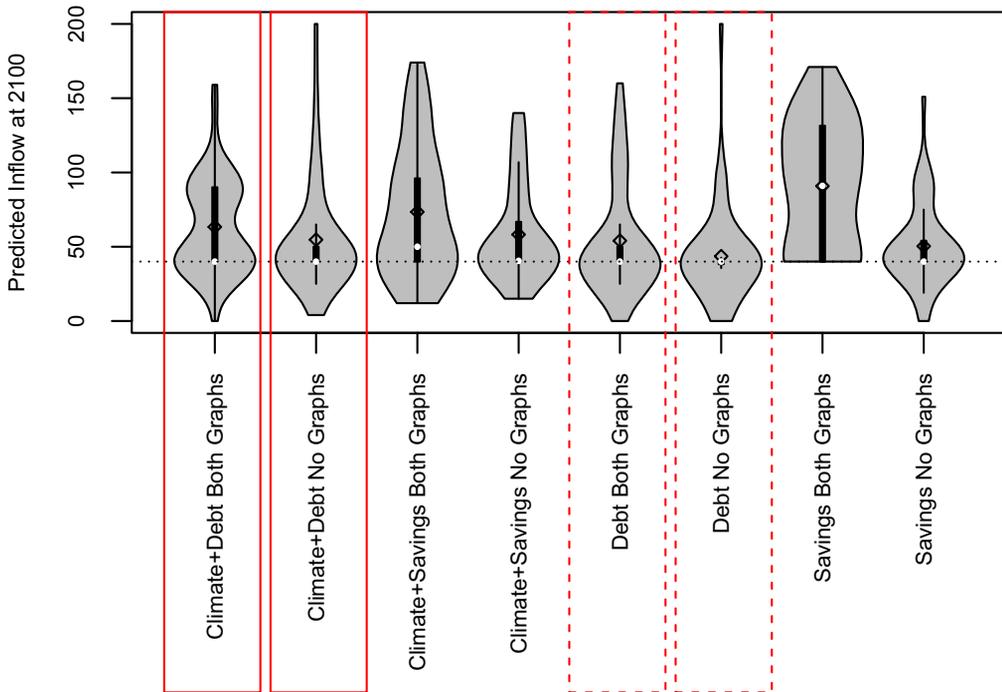


Fig. 3. A violin plot showing the predicted emissions/spending in each of the eight cells of Experiment 2. The white dot is the median, and the open diamond is the mean; the box encompasses the 25%–75% interquartile range, and the gray areas either side of the center line of each “violin” are rotated kernel density estimation plots. The gray dotted line at 40 on the y-axis indicates the correct answer. (To facilitate comparison of common conditions across Experiments 1 and 2, the relevant violins are highlighted with red boxes—solid lines for climate and debt conditions, dashed lines for debt alone.)

estimates higher (and thus further from the correct answer) in the Savings conditions ($M = 67.16$, $SD = 39.41$) than the Debt conditions ($M = 53.97$, $SD = 34.75$), $F(1, 328) = 21.02$, $p < .001$, $\eta_p^2 = 0.039$. A significant effect of Graphs was also found, with estimates higher with graphs ($M = 68.49$, $SD = 40.05$) than without graphs ($M = 51.37$, $SD = 32.61$), $F(1, 328) = 21.395$, $p < 0.001$, $\eta_p^2 = 0.061$. There was also a significant interaction effect between Financial Frame and Graph, $F(1, 328) = 5.92$, $p = .015$, $\eta_p^2 = 0.017$, with the difference between debt and savings greater when graphs were present than when graphs were absent. The ANOVA found no significant main effect of the presence of the Climate task, $F(1, 328) = 1.28$, $p = .26$, $\eta_p^2 = 0.003$; nor did that factor interact with the other two.

To examine whether the advantage for the debt analogy over the climate task alone (observed in Experiment 1) was replicated, a planned cross-experiment analysis was conducted comparing the relevant Climate task alone conditions from Experiment 1 with the Debt Alone conditions from Experiment 2. This 2(Graphs vs. No Graphs) \times (Climate vs. Debt) revealed a significant effect of the analogical context, $F(1, 165) = 4.48$, $p = .036$, $\eta_p^2 = 0.027$, replicating the finding that estimates were higher (and thus further

from the correct answer) in the climate ($M = 60.74$, $SD = 36.06$) than in the debt context ($M = 48.98$, $SD = 34.73$). There was also a significant main effect of graph, $F(1,165) = 14.45$, $p < .001$, $\eta_p^2 = 0.088$, with estimates higher when both graphs were present ($M = 63.20$, $SD = 38.60$) than when graphs were absent ($M = 44.27$, $SD = 29.41$). The interaction was not significant, $F(1, 165) = 3.29$, $p = .074$, $\eta_p^2 = 0.02$.

A similar cross-experimental analysis compared the Climate Task alone conditions from Experiment 1 with the Climate & Debt conditions from Experiment 2. This revealed, again consistent with Experiment 1, no advantage of including the debt scaffolding $F(1, 163) = 0.03$, $p = .86$, $\eta_p^2 = 0.0002$, but a clear effect of graph $F(1, 163) = 13.21$, $p < .001$, $\eta_p^2 = 0.08$, with estimates higher when both graphs were present ($M = 68.49$, $SD = 35.10$) than when graphs were absent ($M = 50.59$, $SD = 32.51$). There was however, a significant interaction effect, $F(1, 163) = 3.98$, $p = .048$, $\eta_p^2 = 0.024$, because the difference between the graph and no graph conditions was greater in the Climate Task alone condition than in the Climate & Debt condition. Finally, a comparable cross-experiment analysis of the Savings & Climate conditions from Experiment 2 with the Climate Task alone conditions of Experiment 1 also revealed no effect of the savings analogy ($F < 1$), and a similar detrimental effect of the presence of graphs ($F = 14.53$).

3.4. Discussion

The standout result from Experiment 2 is that mere familiarity with concepts is *not* enough on its own to improve stock-flow reasoning (cf., Brunstein et al., 2010; Cronin et al., 2009; Gonzalez & Wong, 2012). Participants given the savings frame alone were, on average, more inaccurate in their estimates than participants in the other conditions—especially when graphs were present. This second finding of a deleterious effects of graphs replicates the finding in Experiment 1 and echoes the conclusions of Guy et al. (2013) and Cronin and Gonzalez (2007), that graphs may not always help participants solve stock-flow problems. The fact that the savings frame alone did not help reasoning lends weight to the notion that the more accurate estimates observed in the debt conditions of both experiments (and in Newell et al., 2013) are due in part to the coincidental similarity in the “valence” of the debt problem: “debt is bad, so spending should come down.” In contrast, the opposing valence in the savings problem—“savings are good, so earnings should go up” led to more inaccurate responding. The cross-experiment analyses confirmed that providing a financial context (either debt or savings) as a “scaffold” for reasoning did not confer an advantage over performance in the standard climate task.

4. General discussion

To avoid dangerous global warming, society needs to adopt rapid and considerable cuts to CO₂ emissions. A full appreciation of why this is the case requires some basic understanding of the relationship between accumulated and emitted CO₂. In two experiments we tested whether reframing the problem in terms of financial, rather than CO₂,

“budget” management aided understanding of stocks and flows, and whether graphical representations of these relationships helped or hindered understanding.

Experiment 1 found that participants were better at reasoning about a financial debt than reasoning about our CO₂ “debt.” Experiment 2 showed that this improvement did not hold when the financial task required reasoning about savings. The results of both experiments indicate that providing graphical representations of stock-flow relationships can hinder reasoning. Neither experiment found evidence for an improvement in solving the CO₂ accumulation problem when a financial context (debt or savings) was used as a scaffold for reasoning.

4.1. Are analogical contexts and frames useful?

One account that can reconcile the findings across both experiments is the following: There is a core group of participants who understand the climate task irrespective of the presence or absence of the different financial analogies, and would also, presumably perform equally well whether they were given the debt or the savings frame, or the combination of either with the climate task. The analysis of the binary (correct/incorrect) data (available on request from the authors) suggests that approximately 30%–40% of participants in both experiments provided the correct answer (i.e., 40Gt/C or \$40) irrespective of the particular context or framing used. This is somewhat higher than found in other similar experiments (e.g., Guy et al., 2013; Sterman & Booth-Sweeney, 2007), but differences in the materials, task, and population preclude drawing strong inferences about the reasons for this level of accurate responders.

There are also participants who never appear to understand the problem and never hit upon the correct answer regardless of the way in which the information is presented. The binary (correct/incorrect) data suggest that this is an approximately similar proportion, around 40% (i.e., we never found more than 60% of participants getting the answer exactly correct regardless of the condition).

There are, however, other participants who find the task difficult, but not impossible, and thus are affected by the analogical context (cf. Gonzalez & Wong, 2012). Specifically, a debt frame in which the “valence” of the problem coincides with the correct response (“debt is bad, therefore reduce”) leads to improvements in accuracy. However, the results of Experiment 2 strongly suggest that this improvement is fortuitous and not because the familiarity of financial management enables a deep understanding of the stock-flow relationships. When *savings* are used instead of debt and the “valence” opposes the correct response (“savings are good, therefore increase”), accuracy is very poor. Moreover, even though our earlier work had indicated an advantage for using the debt frame as a scaffold for understanding the climate task (Newell et al., 2013), no clear benefit was observed here. The reason for this discrepancy remains unclear—and awaits further research—but may be driven in part by task requirements. In Newell et al. (2013), participants made estimates for each decade of a future emissions trajectory, whereas in Experiment 1 and 2 they only predicted the end point of the trajectory. Perhaps focusing on the endpoint alone in some way weakens the usefulness of the analogous context.

4.2. *Do graphs help or hinder?*

The presence of graphical representations of the stock-flow relationships also appear to affect the majority of participants who do not understand the task (cf. Guy et al., 2013). When the graphs invite erroneous and, at times, correlation heuristic-consistent responding; or perhaps reinforce the “up is good” interpretation in the savings task, then inaccurate estimates are made. The poor performance in the savings alone task is informative: It suggests that participants found the idea that one had to reduce earnings to achieve a savings goal to be rather counter-intuitive.⁵ This is not surprising given how we typically think about our savings (e.g., we need to earn more to save more) but again highlights the difficulty participants have in engaging the more abstract logical reasoning required to solve stock-flow tasks—especially when potentially misleading graphs are present (Guy et al., 2013).

An additional point regarding the role of graphs is that their usefulness may depend on whether participants are predicting only end-points or complete trajectories. In our experiments and those of Guy et al. (2013) participants predicted end-points and graphs were found to be a hindrance; in previous research that has found no detrimental impact of graphs (compared to text or tables, for instance), participants predicted trajectories (or multiple time-points) (e.g., Cronin et al., 2009; Sterman & Booth-Sweeney, 2007). Future research should systematically explore the effect of graphs in situations where participants predict either end-points or trajectories.

4.3. *Promoting deeper understanding*

A very general take-home message from this area of research is that ours and others’ (e.g., Guy et al., 2013) limited success in re-framing the CO₂ accumulation problem points to the need for more theoretically driven design of analogies to be used in future work. Our instructions in the conditions that combined the financial context with the climate task were quite directive, for example, telling participants to think of CO₂ accumulation in terms of debt and clearly stating, “If you spend more than you earn you get into debt, and if you keep spending more your debt grows.” However, given the absence of any clear advantages of the combined condition relative to the climate alone version, it seems they were insufficiently directive.

Perhaps, in hindsight, this is not so surprising given that in comparison to other analogy studies these instructions are relatively minimal. For example, we did not give participants an explicit comparison process (e.g., by first solving the debt task, then solving the climate task) nor did we invite them to list similarities and differences between the two scenarios (see e.g., Gentner et al., 2003; Gonzalez & Wong, 2012; Smith & Genter, 2012). Indeed, future studies might employ this more explicitly comparative technique to encourage participants to consider both the surface *and* the deeper structural or behavioral similarity across different stock-flow analogical contexts. Work by Gonzalez and Wong (2012) suggests that both the surface and behavioral similarity between systems/problems are important for improving stock-flow reasoning, but that behavioral similarity is more

crucial. Behavioral similarity (or deep structure) refers to the underlying functional relationships within a system. Gonzalez and Wong (2012) employed explicit analogical comparisons using versions of the “department store” stock-flow task and found that when participants were able to compare two problems sharing behavioral similarities, accuracy increased markedly.

5. Conclusions

We began this work with the aim of finding an analogy and a method of presentation that would help people understand the basic stock-flow processes involved in the dangerous build-up of CO₂ in the atmosphere. Our ultimate goal is to develop methods of communication that can overcome misunderstandings about CO₂ accumulation and thereby, potentially, shift attitudes and encourage behavior that will mitigate dangerous climate change (Guy et al., 2013; Sterman, 2008). The current results suggest that although people do find reasoning about financial debt easier than reasoning about the carbon “debt,” our earlier preliminary conclusions about the transferable positive influence of a debt frame on *deep understanding* of CO₂ accumulation were premature (Newell et al., 2013). A question for future research is whether the fortuitous *surface similarity* in financial *debt* management and managing our climate problem can be harnessed to make the important message about our effect on the climate both heard and heeded.

Acknowledgments

This work was supported by an Australian Research Council Future Fellowship (FT110100151) to the first author and a Linkage Project Grant (LP120100224). The support of the Australian Research Council Centre of Excellence for Climate System Science is also acknowledged. We thank Bethany Grows for research assistance. Cleotilde Gonzalez was supported by the National Science Foundation Award number 1154012.

Notes

1. In an attempt to address this issue, Guy et al. presented a cover story in which a toxic spill had occurred in the vicinity, thus forcing people to run water *through* their bath (i.e., with the plug out) to ensure that all the toxic water had left the system. However, the modest improvement in accuracy (see main text) suggests that even with this cover story many participants were not able to benefit from the mapping between the bathtub and the CO₂ task.
2. A 3×4 ANOVA on completion time detected a main effect of analogical context, $F(2, 446) = 3.286, p = .04, \eta_p^2 = 0.015$, with a post hoc Scheffé contrast finding that participants who received the Climate & Debt tasks took longer to complete the

study ($M = 10.30$, $SD = 20.77$) than participants in either the Debt ($M = 6.77$, $SD = 3.36$) or the Climate conditions ($M = 7.74$, $SD = 4.11$), $F(2, 446) = 6.14$, $p = .002$. The longer time taken by the Climate & Debt condition participants is unsurprising, given that they had to complete additional mapping questions.

3. After completing the stock-flow task, all participants answered questions about political orientation, agreement with regulation-based climate policy, belief in the influence of human activity on global warming, and attitudes toward the seriousness of anthropogenic global warming. The attitude measures were included because we wanted to examine whether accuracy on the CO₂ accumulation task was related to participants' attitudes toward climate change, and their willingness to endorse climate-related regulatory policy. We also wanted to examine whether accuracy was predicted by level of education or political orientation (cf. Guy et al., 2013). Despite some relatively weak, suggestive effects in the attitude measures in Experiment 1—(e.g., participants exposed to the climate task were somewhat more willing to endorse climate regulation-policy than those given the financial task), these effects did not replicate in Experiment 2. In addition, we found no support in either experiment for a relationship between accuracy on the accumulation task and willingness to act, nor did education level or political orientation predict task accuracy. For these reasons we chose not to report these data in full, but the analyses are available from the authors on request.
4. A $2 \times 2 \times 2$ ANOVA on completion time found a main effect of the presence of the Climate Task, $F(1, 328) = 11.721$, $p < .001$, $\eta_p^2 = 0.034$, with participants who solved either the Climate and Debt or Climate and Savings versions ($M = 8.51$, $SD = 9.66$) taking longer than those who solved either the debt or savings problems alone ($M = 5.84$, $SD = 3.27$). This additional time taken is consistent with the pattern in Experiment 1 (see note 2).
5. This interpretation is supported by an examination of the answers to the mapping questions quizzing participants about the relations between the concepts in the climate and the financial task (see Methods). In the Climate & Debt conditions of Experiment 1 and Experiment 2, 88% and 82%, respectively, of participants answered at least two of the three mapping questions correctly; in the Climate & Savings condition of Experiment 2 this proportion dropped to 50%. Note that the correct answers to these questions were provided before participants completed the estimation task.

References

- Anderegg, W. R. L., Prall, J. W., Harold, J., & Schneider, S. H. (2010). Expert credibility in climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 107(27), 12107–12109.
- Bord, R. J., O'Connor, R. E., & Fisher, A. (2000). In what sense does the public need to understand global climate change? *Public Understanding of Science*, 9(3), 205–218.

- Brunstein, A., Gonzalez, C., & Kanter, S. (2010). Effects of domain experience in the stock-flow failure. *System Dynamics Review*, 26(4), 347–354.
- Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., & Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8(2), 024024.
- Cronin, M. A., & Gonzalez, C. (2007). Understanding the building blocks of systems dynamics. *Systems Dynamics Review*, 23(1), 1–17.
- 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.
- Dutt, V., & Gonzalez, C. (2012a). Decisions from experience reduce misconceptions about climate change. *Journal of Environmental Psychology*, 32(1), 19–29.
- Dutt, V., & Gonzalez, C. (2012b). Human control of climate change. *Climatic Change*, 111(3–4), 497–518.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393–408.
- Gonzalez, C., & Wong, H. (2012). Understanding stocks and flows through analogy. *System Dynamics Review*, 28(1), 3–27.
- 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.
- Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, 2, 732–735.
- Leiserowitz, A. A., Smith, N., & Marlon, J. R. (2010). Americans' knowledge of climate change. Yale Project on Climate Change Communication.
- Leviston, Z., Walker, I. A., & Morwinski, S. (2013). Your opinion on climate change might not be as common as you think. *Nature Climate Change*, 3, 334–337.
- Lewandowsky, S. (2011). Popular consensus: Climate change set to continue. *Psychological Science*, 22(4), 460–463.
- Moxnes, E., & Saisel, A. K. (2009). Misperceptions of global climate change: Information policies. *Climatic Change*, 93(1–2), 15–37.
- Newell, B. R., Kary, A., Moore, C., & Gonzalez, C. (2013). Managing our debt: Changing Context Reduces Misunderstanding of Global Warming. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3139–3144). Austin, TX: Cognitive Science Society.
- Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt*. London: Bloomsbury.
- Ranney, M. A., Clark, D., Reinholz, D., & Cohen, S. (2012). Changing global warming beliefs with scientific information: Knowledge, attitudes, and RTMD (Reinforced Theistic Manifest Destiny theory). In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Meeting of the Cognitive Science Society* (pp. 2228–2233). Austin, TX: Cognitive Science Society.
- Raupach, M. R., Canadell, J. G., Ciais, P., Friedlingstein, P., Rayner, P. J., & Trudinger, C. M. (2011). The relationship between peak warming and cumulative CO₂ emissions, and its use to quantify vulnerabilities in the carbon-climate-human system. *Tellus Series B*, 63(2), 145–164.
- Smith, L. A., & Genter, D. (2012). Using spatial analogy to facilitate graph learning. In C. Stachniss, K. Schill, & D. Uttall (Eds.), *Spatial cognition VIII* (pp. 196–209). Berlin, Germany: Springer.
- Sterman, J. D. (2008). Risk communication on climate: Mental models and mass balance. *Science*, 322(5901), 532–533.
- Sterman, J. D., & Booth-Sweeney, L. (2007). Understanding public complacency about climate change: Adults' mental models of climate change violate conservation of matter. *Climatic Change*, 80(3–4), 213–238.
- Weber, E., & Stern, P. (2011). Public understanding of climate change in the United States. *American Psychologist*, 66(4), 315–328.

Appendix A1: Example Wording from the Climate & Debt Condition, Both Graphs group

Consider the issue of global warming. In 2001, the Intergovernmental Panel on Climate Change (IPCC), a scientific panel organized by the United Nations, concluded that the carbon dioxide (CO₂) and other greenhouse gas emissions were contributing to global warming. The panel stated that “most of the warming observed over the last 50 years is attributable to human activities.”

The amount of CO₂ in the atmosphere is affected by natural processes and human activity. Anthropogenic CO₂ emissions (emissions resulting from human activity, including combustion of fossil fuels and changes in land use, especially deforestation), have been growing since the start of the industrial revolution. Natural processes gradually absorb CO₂ from the atmosphere (e.g., as it is used by plant life and dissolves in the ocean).

Before humans started emitting large amounts of CO₂ into the atmosphere, emissions were largely the result of these natural processes, and the amount emitted into the atmosphere was roughly equal to the amount that was absorbed. This kept the CO₂ concentration in the atmosphere balanced. While there has been a sharp increase in CO₂ being emitted into the atmosphere, the amount being absorbed has remained constant. The consequence of this is that CO₂ is rapidly accumulating in the atmosphere and contributing to climate change.

You might like to think of controlling the level of CO₂ in the atmosphere as similar to controlling your personal finances. If you spend more than you earn, you get into debt, and if you keep spending more, your debt grows.

Atmospheric CO₂: “Debt Level”

The amount of CO₂ in the atmosphere is like our current “debt” level and just as you would not want your debt to get bigger, we do not want atmospheric CO₂ levels to keep rising. The top graph shows a scenario in which this has been achieved. On the graph, the amount of CO₂ in the atmosphere rises from an initial value of just over 600 gigatonnes of carbon (GtC) (in 1900) but then over a period of 200 years, it stabilizes (does not increase anymore) to be 945 GtC (by 2100).

Emissions and absorption of CO₂: “Spending & Earning”

CO₂ is emitted via human activities and absorbed by natural processes. You can think of CO₂ emissions as like money you spend and CO₂ absorption like money you earn. On the bottom graph, the green line shows how much CO₂ is absorbed per decade via natural processes—it remains constant across this same time period (1900–2100) at 40GtC. The black line shows how much CO₂ is emitted via human activity per decade across the same time period.

Your task is to predict the amount of CO₂ emissions in the year 2100 in order for the amount of CO₂ in the atmosphere to stabilize (stop going up) by that date. That is, to achieve the scenario shown in the graph, in which atmospheric CO₂ has stabilized.

In other words you need to think about how you would change your rate of spending (emissions) to ensure that your debt (atmospheric CO₂) stabilized (stopped going up).

Please think about how the processes of personal debt relate to the processes of climate change and correctly match the following items.

Accumulation of CO₂ in the atmosphere.

- a) Total amount of debt
- b) Amount that you earn
- c) Amount that you spend

Natural absorption of CO₂ from the atmosphere.

- a) Total amount of debt
- b) Amount that you earn
- c) Amount that you spend

Human emissions of CO₂ into the atmosphere.

- a) Total amount of debt
- b) Amount that you earn
- c) Amount that you spend

Remember

The “stock of CO₂” line shows the total amount of CO₂ in the atmosphere in each decade.

You can think of the atmospheric CO₂ in the top graph as how much you are in debt.

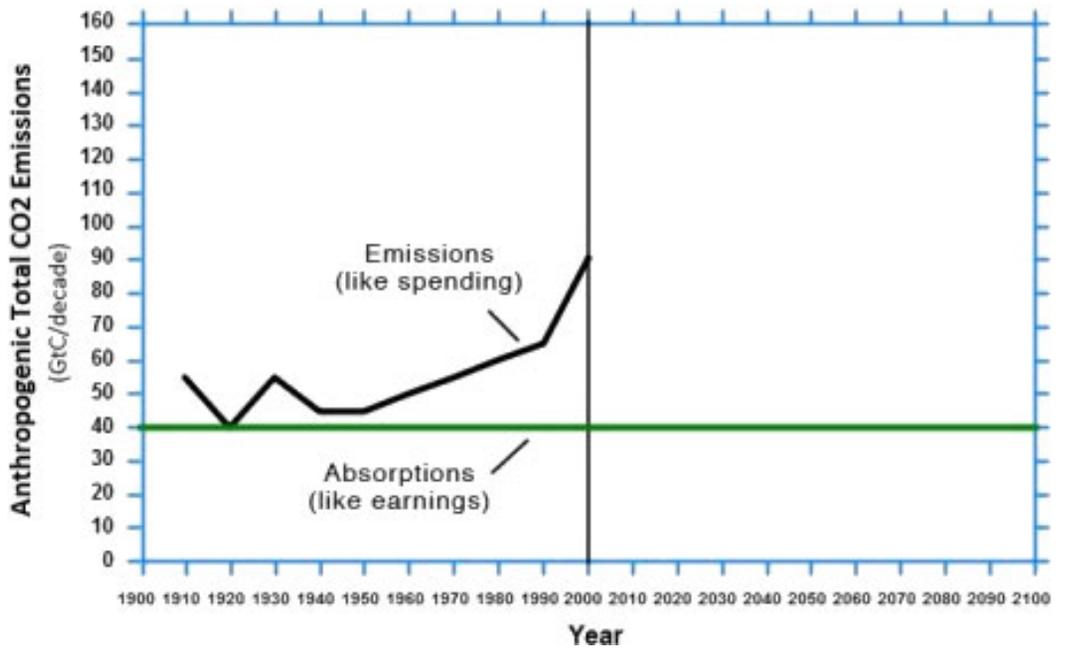
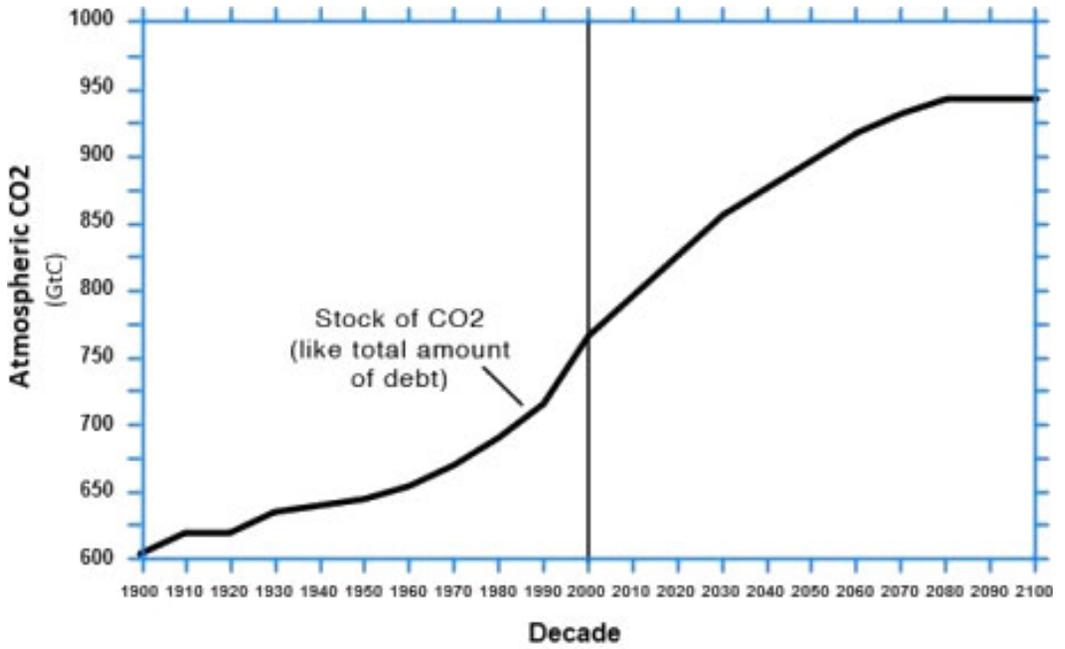
The emissions line in the second graph shows how much CO₂ was emitted via human activity in that decade.

You can think of the emissions in the second graph as how much money you are spending.

The absorptions line in the second graph shows how much CO₂ was absorbed via natural processes by in that decade.

You can think of the absorptions in the second graph as how much money you are earning.

In order for the total amount of CO₂ in the atmosphere to stabilize and not increase, what does the rate of emissions need to be in 2100?



Appendix A2: Example Wording from the Climate & Savings Condition, No Graphs group (Experiment 2).

You might like to think of controlling the level of CO₂ in the atmosphere as similar to controlling your personal finances. If you earn more than you spend, you make savings, and if you keep earning more than you spend, your savings grow.

Atmospheric CO₂: “Savings Level”

The amount of CO₂ in the atmosphere is like our current “savings” level and just as you might not like to keep saving after you have reached a particular goal (e.g., enough money for a new TV), we also do not want CO₂ emissions to rise beyond a particular level.

Imagine a scenario in which this has been achieved: The amount of CO₂ in the atmosphere rises from an initial value of just over 600 gigatonnes of carbon (GtC) (in 1900) but then over a period of 200 years, it stabilizes (does not increase anymore) to be 945 GtC (by 2100).

Emissions and absorption of CO₂: ‘Earning & Spending

CO₂ is emitted via human activities and absorbed by natural processes. You can think of CO₂ emissions as like money you earn and CO₂ absorption like money you spend. Imagine a scenario in which every decade, the rate of CO₂ being absorbed remains constant at 40 GtC per decade. Imagine also that by the year 2000, the rate of CO₂ being emitted into the atmosphere via human activities had reached 90 GtC per decade.

Your task is to predict the rate of CO₂ emissions in the year 2100 in order for the amount of CO₂ in the atmosphere to stabilize (stop going up) by that date. That is, to achieve the scenario described above, in which atmospheric CO₂ has stabilized.

In other words, you need to think about how you would change your rate of earnings (emissions) to ensure that your savings (atmospheric CO₂) stabilized (stopped going up).

The following task involves thinking about the rate of emissions needed to stabilize the level of CO₂ in the atmosphere.

Remember

The amount of CO₂ in the atmosphere stabilizes by 2100.

You can think of the atmospheric CO₂ as how much you have saved.

The emission of CO₂ via human activity into the atmosphere is 90 GtC in the year 2000.

You can think of the emissions as how much money you are earning.

The absorption of CO₂ via natural processes remains constant at 40 GtC every decade.

You can think of the absorptions as how much money you are spending.