Climate Change and Conceptual Change

By

David J. Clark

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Michael A. Ranney, Co-chair
Professor Richard B. Ivry, Co-chair
Professor Sonia Bishop
Professor John Canny

Fall 2013

Climate Change and Conceptual Change

Abstract

Climate Change and Conceptual Change

by

David J. Clark

Doctor of Philosophy in Psychology

University of California, Berkeley

Professor Michael A. Ranney, Co-chair
Professor Richard B. Ivry, Co-chair

My work is awesome. Give me a Ph.D.

With love and respect for all that is good—in particular the unwavering support of my grandma (even if she doesn't fully accept climate change *or* evolution).

Contents

List of Figures

List of Tables

Acknowledgments

Thanks for all the help!

This dissertation was based on the ucastrothesis template. It was typeset with LuaLaTeX using fonts from the TeX Gyre project (Pagella for main text and Heros for sans serif) and Inconsolata for monospaced text. The research reported here was supported in part by the Institute of Education Sciences pre-doctoral training grant R305B090026 to the University of California, Berkeley and the University of California, Berkeley Committee on Research. The opinions expressed are those of the author and do not necessarily represent views of the Institute of Education Sciences, the U.S. Department of Education, or the University of California.

Chapter 1

Teaching the Mechanism of the Greenhouse Effect

As described above, American's lag much of the world regarding acceptance of anthropogenic climate change. Informally, Michael Ranney, then other members of our group started questioning whether people were able to mechanistically explain how human activities cause an increase in global mean temperature. Almost no one could provide a satisfactory explanation, including us! Prof. Ranney, Lloyd Goldwasser, and Daniel Reinholz (along with input from myself and Ronald Cohen) proceeded to develop a short 400-word description of the mechanism. This text is reproduced in full in Appendix D. Before reading the text yourself, I would encourage you to spend 10 minutes describing *your* understanding either aloud or on paper. [1]

Given that our informal investigations implied that almost no-one knew the basic concepts described in our 400 words, we initiated the line of research described below. In these experiments, we sought to formally ask:

1. Is this lack of understanding for the mechanism of global climate change as pervasive as it seemed to be?

2. Does instruction regarding the mechanism of global climate change increase individuals' acceptance of the reality of anthropogenic climate change?

Along the way, we additionally considered related aspects of learners' cognition (the details of which are described below).

The history of educational research would imply that it's quite difficult to arrive at definitive answers to big policy qeustions. For example, phonics vs. whole-word reading has been debated at least since the dawn of the Common Era, as discussed in **history-reading-instruction** Below, however, I report on a series of experiments that argue strongly (if not definitively!)

[1] If you personally doubt the veracity of anthropogenic climate change, then you may modify the exercise to describing the mechanism of the greenhouse effect, first described by Nobel laureate XXX Ahreneous (Sp?) in **Ahreneous** (and accepted as fact since that time!).

that instruction regarding the pysical mechanism of the greenhouse effect appears to have some positive effect on public acceptance of anthropogenic climate change. As discussed above, such public acceptance seems central to any truly democratic approach to the problem of climate change.

Others' studies noted above detract from the utility of approaching climate change as a science education problem. However, unlike those studies, the interventions in this chapter have focused on a fundamental, well-researched knowledge gap, and our assessment focused on acceptance/belief. Such contrasts may explain the difference between observing instructional benefits (as we have) or polarization (as others occasionally have; cf. Lundmark, 2007). We'll see further evidence below, however, that such interventions are applicable across a variety of settings, time-frames, and populations, and that global warming understandings and attitudes are far from static. Most importantly, such understandings seem to affect attitudes and beliefs in a meaningful way.

1.1 Introducing our survey methods

This chapter, as well as Chapters **??** and **??** all utilize similar survey methods to assess climate-related beliefs and attitudes, in addition to a number of related constructs relevant to Ranney's (2012) RTMD theory. For reference, the full list of survey items used in this body of research is included in Appendix B. Below, we'll examine the way our interventions are able to shift these beliefs and attitudes (primarily those related to climate change), as well as noting how these beliefs and attitudes relate to one another. And, I hasten to note, the number of potential relationships between the many variables we have measured would require an enormous amount of data to test fully. As such, we will restrict ourselves primarily to the exploration of *a priori* relationships of interest.

1.1.1 Clarifying "beliefs" and "attitudes"

Survey methods in the social sciences may use the terms "belief" and "attitude" in numerous ways. For example, an "attitude" may refer to a measured response, whereas a "belief" may refer to a latent variable that explains a number of such responses **attitude-latent-var-ref** Here, we take a more common-language approach. Specifically, in the text that follows, a "belief" is a measure of agreement with an objectively verifiable fact about the world (For example. For example, the reality of anthropogenic climate change (item `gw1_2`) may be difficult to ascertain, but in the end, it is something that could be settled by observation. An "attitude," on the other hand, is a measure of agreement with an emotional stance towards some aspect of the world. For example, worry about global warming (`gw2_3`).

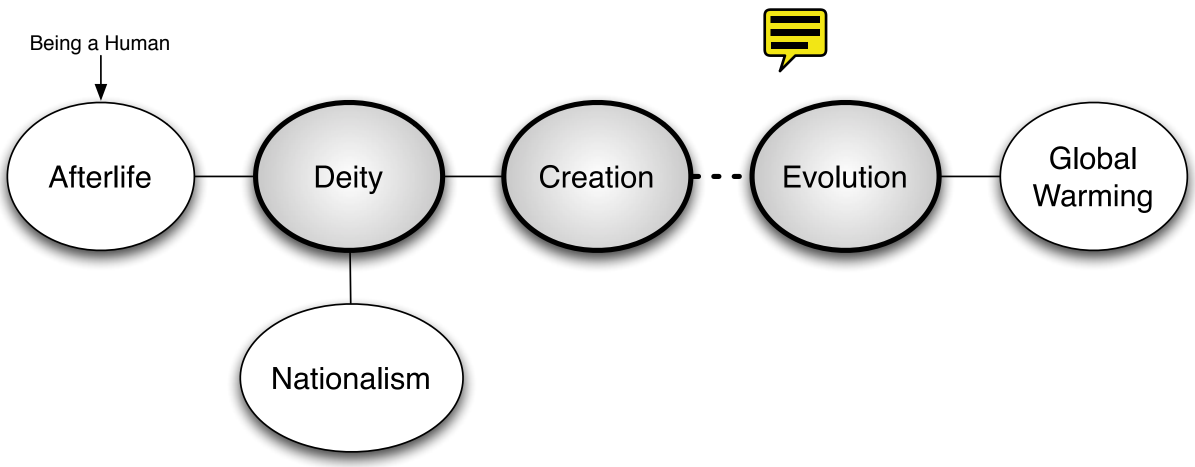


Figure 1.1: Core RTMD relationships (adapted from Ranney, 2012). Conceptual reinforcement (operationalized as positive correlation) is represented by a solid line, and conceptual incompatability by a dashed line. A coherent cluster representing devotion to "god and country" is on the left, and a science-accepting cluster is on the right. The most direct conflict is captured by the explanatory competition between creation and evolution.

1.1.2 An overview of survey items

Survey items were primarily sourced from Martinez (2009), and were chosen on the basis of both observed quality (from the results of Martinez, 2009) and conceptual fit for our interventions. The first page in particular (consisting of the 6 items with a "1" prior to the underscore) was selected as the most ideal set of 6 questions for targeting the 6 RTMD constructs, depicted in Figure 1.1 (Ranney, 2012). Notably, `gw1_2` was selected for its focus on acceptance of *anthropogenic* climate change.

1.1.3 Naïve survey results

Most of the climate-related interventions that follow include some measure of participant attitudes and beliefs prior to the intervention. In this dissertation, we are primarily seeking insight into different forms of conceptual *change* (and thus, one hopes, behavioral change). Given this, a detailed consideration of these naïve results is beyond our scope. However, some of the relationships obtained seem relevant to understanding the mind of our potential students. I thus note such results below. For a fuller treatment of survey material, please consult the relevant publications of the Reasoning group (notably, Cohen, 2012a; Ranney, Clark, Reinholz, and Cohen, 2012a).

Most importantly, we were able to observe some important relationships from Cohen's (2012a) data contrary to the "knowledge deficit" and polarization arguments referenced in Chapter **??**. First, we observe a robust correlation between mechanistic climate change knowledge and attitude toward climate change. This result was maintained even when taking political party into account. Specifically, mechanistic knowledge correlates with acceptance that global warming is occurring ($r = 0.22$, $p = 0.0002$) and is anthropogenic ($r = 0.17$, $p = 0.005$). Anthropogenic climate change acceptance also predicted financial "willingness to sacrifice" ($\chi^2(4) > 32$, $p < 0.001$ for each of four items), and one's knowledge score predicted two of these items ($\chi^2(1) > 3.8$, $p < 0.05$ for both). Further, acceptance of biological evolution was found to predict beliefs and attitudes toward climate change (as RTMD hypothesizes, and, e.g., Ranney, 2012 found). We might infer, then, that "acceptance of controversial science" is a problem above and beyond political ideology. These findings suggest that the effects of well-chosen aspects of education are both significant and somewhat independent of political affiliation. Indeed, evolution acceptance was a significant predictor of climate change acceptance even in a model including the two major political parties ($\chi^2(4) = 12.3$, $p < 0.02$; N.B., including other parties dramatically reduces quality of fit for any model, likely due to small bin sizes). Given these results with a representative sample of the american public, we considered ourselves justified in focusing primarily on attitude and belief questions in the interventions that follow.

Below, as appropriate, we will see how such relationships maintain in our various populations in the context of a number of interventions.

1.2 Study 1: Classroom interventions at UC and UT

This experiment was a fairly thick, exploratory observation of individuals' beliefs, attitudes and knowledge. Here, "thick" means that we explored the same phenomenon via multiple routes—for example, doing keyword searches of textual responses, and examining coded responses via a number of metrics. We sought to understand how a relatively brief 400-word mechanistic explanation might affect these measures, as well as how this might be modulated by prior commitment to one's own explanations and stated attitudes. The general flow of the experiment is given in Figure 1.2.

The primary goal here was a proof of concept. By assessing university students—some of our nation's most highly educated citizens—we provide a strong test of our belief that most Americans are ignorant of the mechanism of global climate change. An additional concern was that for maximal power, it is preferable to sample naïve beliefs prior to the intervention. In such a design, we are able to use repeated measures statistics and consequently have much greater power. On the downside, however, problems can arise from assessment prior to an intervention. For example, we were concerned that individuals might exhibit an increase in their stated belief in anthropogenic climate change merely by dint of experimental demand. This is evaluated via comparison between our sandwich and no-pretest groups.

As described in Section 1.1, pre-test responses can be used to asses naïve knowledge and attitudes in the general public.

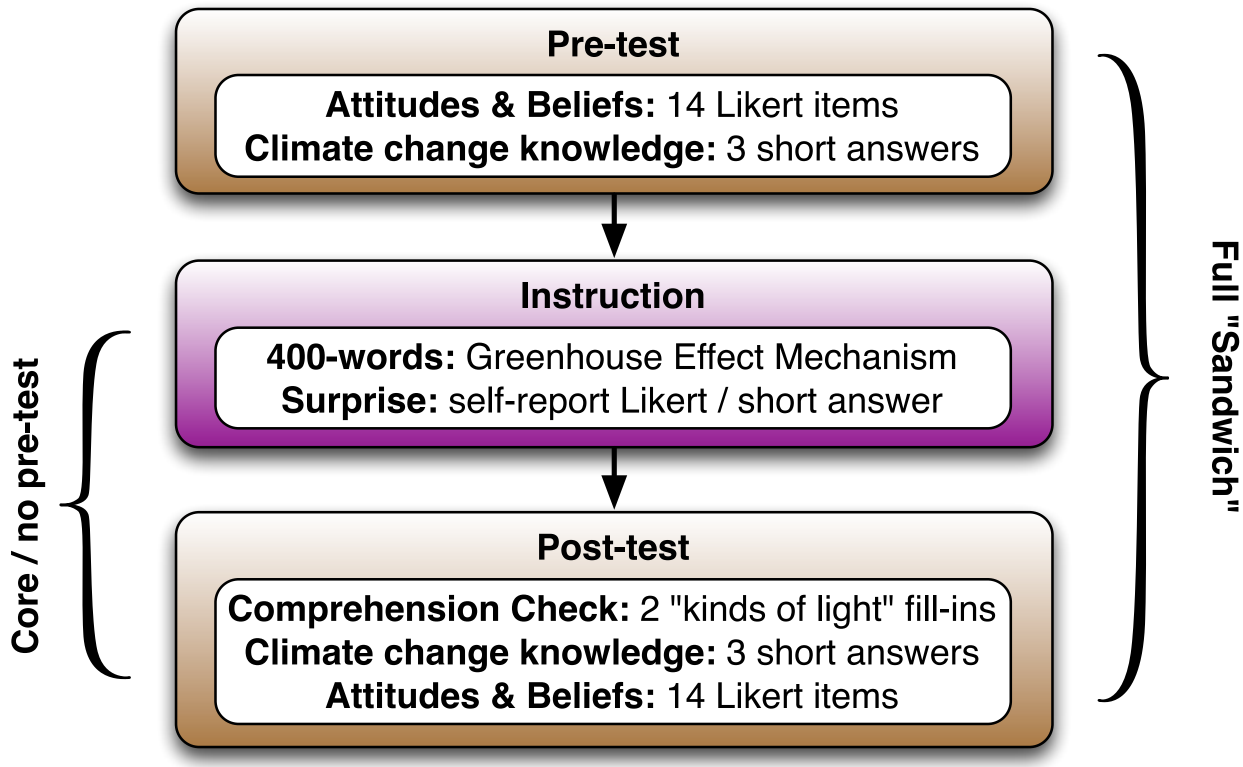


Figure 1.2: An overview of experimental flow for Section 1.2. Flow for other experiments in Chapter 1 was similar. The analogy to a sandwich takes the knowledge and attitutde tests to be slices of bread, and the educational intervention itself is the "jam."

1.2.1 Methods

Materials and Procedure

The general form of the intervention is given in Figure 1.2, and was collected anonymously. Participants were split into two groups, receiving either the "no-pretest" version of the intervention, or the full "sandwich" (filled with nourishing descriptions of climate change!).

The climate change knowledge portion of the pre- and post-tests consisted of the three questions described in Section E.1. For this experiment, the likert items (all on a 1-9 scale) consisted only of knwgbl followed by the first 13 items in Table B.2. Both groups read the educational text regarding the mechanism of greenhouse gasses (reproduced in full in Appendix D), and indicated any surprise they may have experienced (again, on a 1-9 scale). The "kinds of light" check consisted of two fill-in-the-blank questions regarding the kinds of light coming to earth from the sun and radiating away. Here, "sunlight" or "visible light" were considered correct for incoming, and "infrared" was considered correct for outgoing. Some participants wrote "ultraviolet" for incoming light, which one could

Table 1.1: Number of students from both sub-populations stating membership in a given political party. Note that the demographic survey was given after the intervention (and thus may have influence willingness to state party).

Party	Berkeley	Brownsville
decline to state	5	7
democrat	40	7
independent	4	2
libertarian	9	1
none	24	16
other	1	1
republican	2	3

charitably ascribe to a partially correct understanding.

After completing the intervention described above, participants also completed a demographic survey, detailed in Table B.1. The experiment began with a page of instructions, including the assertion that no tricks or deceptions were involved in this study. Lastly, given that their experimental intervention was shorter, individuals in the "no pre-test" group were asked to provide some feedback on the intervention, and also on Al Gore's "An Inconvenient Truth" (if they had seen it). These responses were used to refine our methods going forward, and **bem-future** notwithstanding, should have had no effect on the results.

Participants were run simultaneously for each of the two classes. Instructions were administered by the course instructor, and students received one of two packets—placing them into one of the two groups described above. After completing the consent form on the front of the packet, individuals proceeded to read and answer questions. The entire experiment required approximately 25 minutes to complete.

Participants

103 University of California, Berkeley, and 46 University of Texas, Brownsville, undergraduates were randomly assigned to one of our two groups: "sandwich" or "no-pretest." Below, we report data from the 85 Berkeley and 41 Brownsville students who completed the survey as intended and had been U.S. residents for ten years or more (because we expressly consider U.S. exceptionalism/nationalism). Of the Berkeley data, we analyzed 43 no-pretest surveys and the pretest part of 42 sandwich surveys—but due to anticipated time constraints, only 30 sandwich post-tests could be completed/obtained. Of the Brownsville data, we analyzed 22 no-pretest and 19 complete sandwich surveys.

Of the 85 Berkeley students analyzed, 2 did not complete the demographic test. 43 were female and 40 were male. Mean conservatism was 3.69 (9-point scale; 1.65 standard deviation). Of the 41 Brownsville students 21 were female, 20 were male. Here, mean conservativism was 4.95 (1.77 *sd*). Political affiliation is reported in Table 1.1.

Procedure

Analysis

Handwritten responses were coded and placed into a spreadsheet (for details, see Appendix E. Given the rich nature of these data, many analyses were employed. As such, please see the results section that follows for details of the analysis used for each question.

1.2.2 Results

Please note that all statistical tests are reported in full in the tables associated with this section. In the text, I primarily indicate only a basic value to give a sense for the strength of the result.

Learning the Global Warming Mechanism

Even our rather sophisticated samples initially exhibited incorrect or non-normative understandings of the greenhouse effect's mechanism (e.g., on the roles of ultraviolet light, the ozone layer's depletion, nongreenhouse-gas pollution, and the reflection of incoming light). Most notably, not a single pre-test explanation mentioned different light/radiation types or atmospheric retention time, despite an explicit prompt to explain any differences between the energy traveling toward and away from Earth. However, after reading the 400-word description, 61% of the Berkeley participants across both groups correctly answered that "infrared" light was emitted from Earth (in its fill-in-the-blank space), as did 55% of the Brownsville students who responded.

Beyond the blank-filling items, we statistically analyzed individuals' qualitative explanations—creating scoring rubrics for three central concepts:

Light Differentiating between the types of light entering and exiting the atmosphere

GHGs Atmospheric greenhouse gases' interactions with radiation

Energy The increased atmospheric retention time of energy

Inter-rater reliability was computed using a weighted modification of Cohen's κ, described in full in Appendix A. This reliability was high (weighted $\kappa = 0.71$ based on about one-third of the Berkeley data; $\kappa = 0.67$ across the full Brownsville dataset). Scores were generated based on three separate aspects of understanding captured in the coded texts: "Light," "GHGs," and "Energy." Significant improvements ($p < 0.05$ for all 6 improvement possibilities across the two groups and three conceptual categories). We had no specific hypotheses, however, regarding specific effects of a given concept. Therefore, the data reported below and in Figure 1.3 use combined knowledge scores (each sub-score contributes 3 of 9 total possible points).

Improvements in participant knowledge were readily obtained via different approaches to analysis. For example, our Berkeley students didn't tend to mention the mechanism in

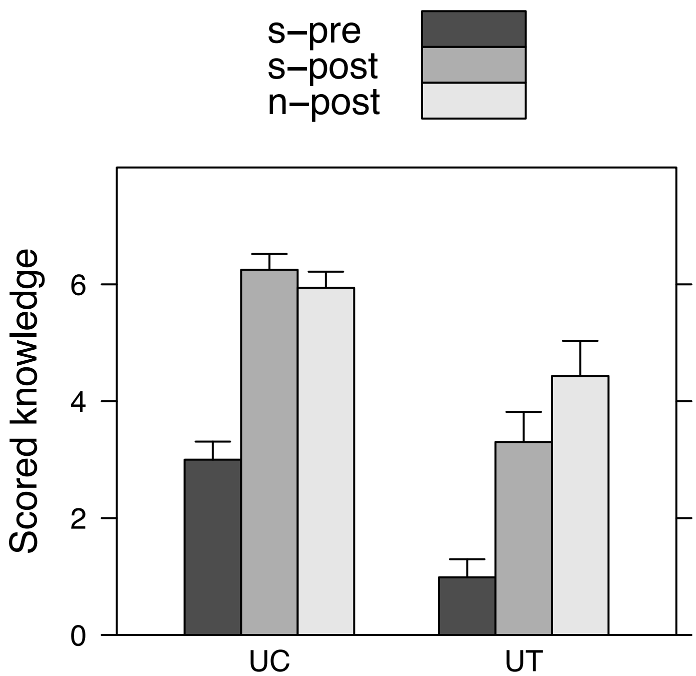


Figure 1.3: Combined scored knowledge for participants in our classroom interventions. All improvements were significant ($p < 0.01$)

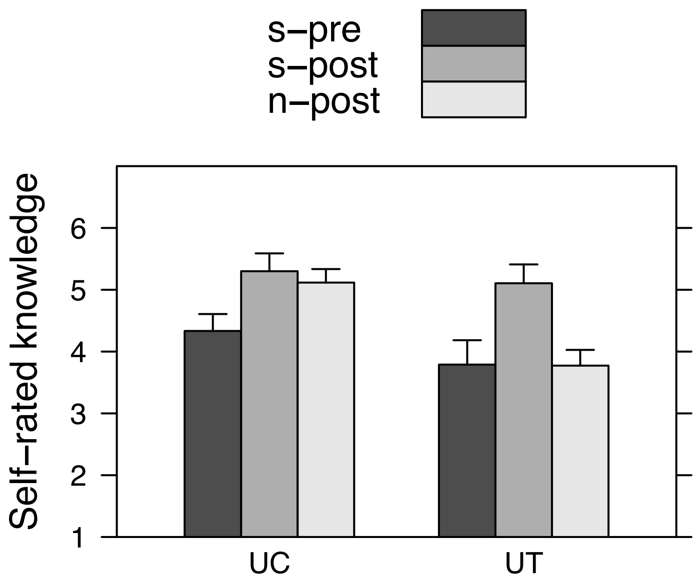


Figure 1.4: Self-rated knowledge scores for participants in our classroom interventions. Improvements in our sandwich groups were significant ($p < 0.01$), while improvements for our no-pretest groups were non-existent in the Brownsville group, and smaller (but still significant, $p < 0.05$ in the Berkeley group.

pre-test (11/42), but they did on the post-test (26/30 for the sandwich group, 39/43 for the no-pretest group).

Participants from both schools experienced significant gains in their self-rated knowledge after the intervention as well ($p < 0.01$). However, for our brownsville students in the no-pretest group, they reported a post-test self-rated knowledge rating almost numerically identical to pre-test ratings in the sandwich group. And while Berkeley students in the no-pretest group increased significantly ($p < 0.05$), that increase was numerically smaller than the sandwich group. These ratings are reported in Figure 1.4.

Global Warming Acceptance Via Mechanistic Learning

To arrive at an easily comparable measure of global warming acceptance, we averaged together all of the items starting with "gw" used in this study. lifsty was omitted due to some concerns regarding multiple interpretation. This concern was in fact unfounded—this construct shifted similarly to the others—but we retain this set of items throughout our statistical testing to maintain consistency and genuine *a priori* hypothesis testing.

It may seem quite remarkable, but participants' global warming acceptance increased dramatically after our brief intervention, as predicted. Proportionally, participants shifted on average 14% closer to "extreme" agreement with climate change items. To assess this,

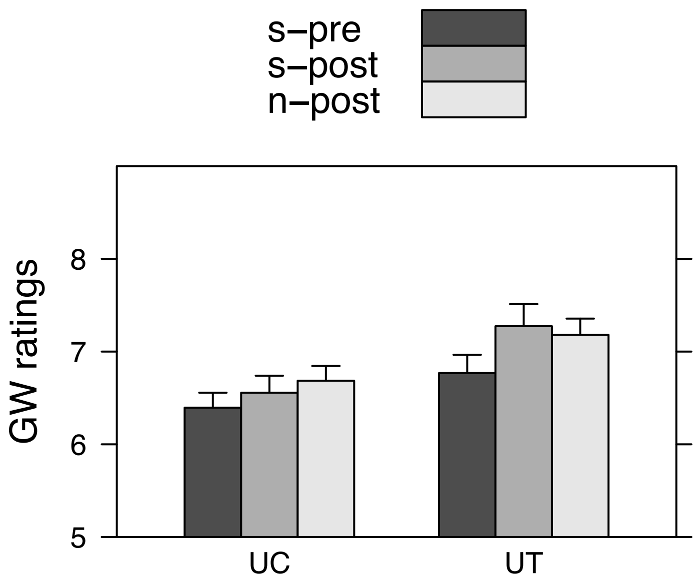


Figure 1.5: Changes in mean of Climate Change related beliefs and attitudes. Improvements from pre- to post-test were significant ($p < 0.05$) with Berkeley students using a combined t-test with imputation. Improvements for Brownsville students were also significant using imputation ($p < 0.01$), as well as looking only at the sandwich group ($p < 0.01$).

we used all of the 73 Berkeley post-test ratings in a paired t-test, and used imputation for pretest scores for the no-pretest group. We found a significant change in global warming acceptance on the posttests, as compared to pre-test measures ($t(72) = 2.28$, $p = .01$). This result was replicated with the Brownsville surveys ($t(39) = 4.24$, $p < .0001$). These ratings are given in Figure 1.5.

Predicting Naïve GW Beliefs and Attitudes

The relationship between knowledge and attitudes was also reflected in Berkeley students' naïve pre-test data, in which participants' self-perceived ratings of their own global warming knowledge correlated significantly with their global warming attitudes ($r = 0.39$, $p = 0.01$). This was not the case with Brownsville students ($r = 0.15$, $p = 0.55$). This may be reflective of overall lower self-perceived knowledge by Brownsville students. But consider the findings above, where we see an even more striking difference in terms of self-rated knowledge between our Berkeley and Brownsville populations. It seems that Brownsville students may simply have a much less grounded notion of their self-knowledge when they are not provided with any context on the matter. The relationships between self-rated knowledge and GW attitudes are depicted in Figure **??**.

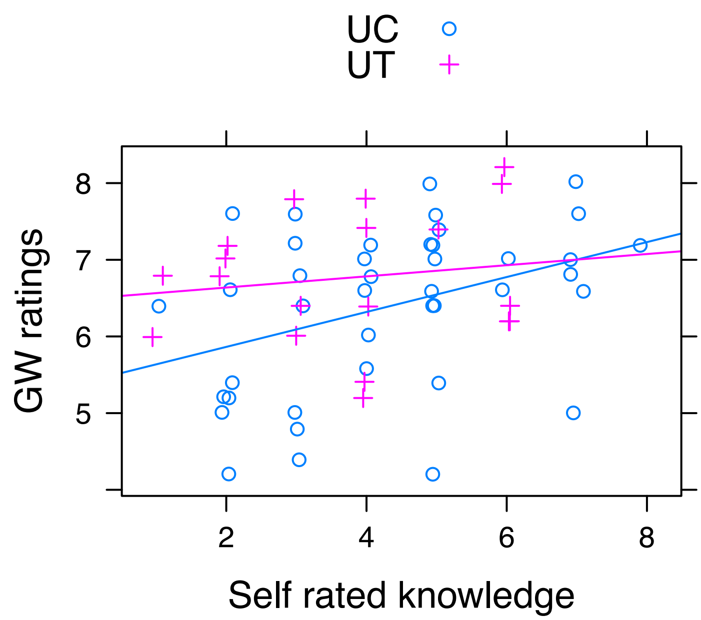


Figure 1.6: Relationship between naïve pre-test self-rated knowledge and mean GW beliefs and attitudes. This data was only available for participants who took the pre-test (i.e., the "sandwich" group). A significant relationship obtains for the Berkeley students ($p < 0.01$), while there is very little relationship in the Brownsville population. Note that no (significant) prediction of attitudes was possible based on scored knowledge.

Surprise

Please recall that we had also predicted a between-conditions difference in surprise ratings due to reduced hindsight biases among the sandwich participants. The difference for Berkeley students was at the significance border-line ($t(42.08) = 1.65$, $p = 0.05$). These surprise ratings increased from a mean of 2.3 to 3.0 on a 9-point scale. It is a bit of a surprise that their ratings are so low in general! the surprise ratings only reached "6" in the no-pretest condition (out of 9, with "5" being "somewhat surprising"), but were as high as "9" (i.e., "extremely surprising") in the sandwich condition. This difference in distribution is depicted in Figure 1.7. Among Brownsville students, surprise was uniformly higher, with a numerically similar difference between conditions, although this result was not significant ($t(38.1) = 0.92$, $p = 0.18$).

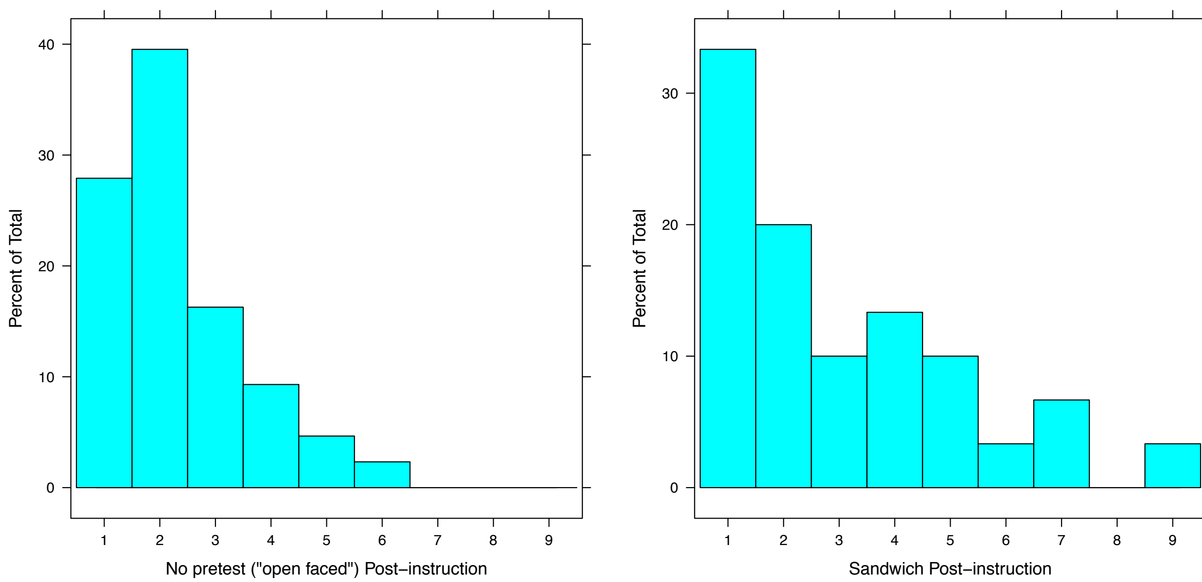


Figure 1.7: Distributions of surprise ratings for the sandwich and open-faced groups, note the slight increase in "1" ratings (which may indicate resistance to the intervention) co-occurs with an increase (from none) in ratings 7-9 in the sandwich group.

Note, I suspect it is unlikely that individuals experienced the same kind of "visceral" surprise from the blurb that can be obtained by, for example, statistics we've used regarding things like abortion and the death penalty. And, while it may be due to a limitation of imagination, I have even difficulty imagining an evolution item that would elicit that kind of surprise either.

Table 1.2: Summary of "improvement" results for Berkeley classroom interventions. All results were *a priori* unless description starts with *"post hoc"*.

Result	p-value	Statistic
People don't tend to mention the mechanism in pre-test (11/42), but they do in post-test (S post is 26/30, N is 39/43 – stat computed for S group pre- vs. post-, which has a lesser prevalence of mech than N)	3.2×10^{-7}	Fisher's exact "two-sided"
Misconceptions are common in the pretest but not the post test, total 0.38 pre- to 0.10/0.12 (Sandwich/No-pretest) post-test. Ozone .19 to .03/.02 (S/N), wrong GHG .24 to 0.07/0.09 (S/N). (test on total misconceptions, Sandwich pre- to group-specific post-test)	0.01	Fisher's exact "two-sided"
Participants don't mention energy leaving the earth until prompted. Specifically, of the four codes that deal with this topic, only 6 mention something about "trapped heat" in the pre-test on the first (i.e., the only unscaffolded) question.	0.0002	Fisher's exact "two-sided"
Use of infrared is greater post-test than pre-test. Goes from 0 to 16 / 22 in S / N groups.	3.5×10^{-8}	Fisher's exact "two-sided"
Sandwich: GHG Objecive knowledge scores improve after the blurb	5.08×10^{-5}	$t(29) = -4.75$ (paired)
No-pretest: GHG Objecive knowledge scores improve after the blurb	2.00×10^{-6}	$t(78.2) = -5.14$ (Welch)
Sandwich: Light Objecive knowledge scores improve after the blurb	3.94×10^{-7}	$t(29) = -6.51$ (paired)
No-pretest: Light Objecive knowledge scores improve after the blurb	1.20×10^{-4}	$t(79.02) = -4.06$ (Welch)
Sandwich: Energy Objecive knowledge scores improve after the blurb	0.04	$t(29) = -2.15$ (paired)
No-pretest: Energy Objecive knowledge scores improve after the blurb	4.60×10^{-4}	$t(80.82) = -3.6547$ (Welch)
Differences in GW attitudes are significant	0.013	$t(72) = -2.28$ (paired / imputed)
Sandwich: pre- to post-test: Increase in self rated knowledge is highly significant	1.40×10^{-5}	$t(29) = 4.96$ (paired)

Table 1.2: Improvements in Berkeley classroom interventions, continued

Result	p-value	Statistic
No-pretest: post-test (compared to S pre-test) increase in self rated knowledge is significant	0.014	$t(78.7) = 2.23$ (Welch)

Table 1.3: Summary of individual and group differences for Berkeley classroom interventions. All results were *a priori* unless description starts with *"post hoc"*.

Result	p-value	Statistic
Surprise is significantly greater in S group than N group	0.053	$t(42.08) = 1.65$ (Welch)
Post-test, slopes (i.e., correlations) between surprise and self-rated knowledge differ between N group (negative, significant) and S group (which was numerically positive).	0.036	$t(69) = 2.137$ (interaction term in a significant linear model)
The no post-test group had a significantly higher word count than the sandwich group's post-test answers for know1	2.5×10^{-4}	$t(82.91) = -383$ (paired)
No-pretest (post): Females are significantly more accepting of climate change than males	0.048	$t(40.19) = -1.71$ (Welch)
There is a significant positive correlation between number of times seeing an inconvenient truth and warming attitudes	0.022	$r(41) = 0.309$

Table 1.4: Summary of relationships between variables for Berkeley classroom interventions. All results were *a priori* unless description starts with *"post hoc"*.

Result	p-value	Statistic
Surprise is almost significantly positively correlated with change in total objectively scored knowledge	0.047	$r(28) = 0.355$

Table 1.4: Relationships between variables in Berkeley classroom interventions, continued

Result	p-value	Statistic
Post hoc: There is a significant correlation between self-rated knowledge and GW attitudes on the pre-test ONLY (differences in self-rated knowledge are also insigificant). NB: we predicted the opposite result!	0.012	$r(40) = 0.386$
Post hoc: Sandwich: Negative correlation between post-test self-rated knowledge and CHANGE in objective score	0.011	$r(28) = -0.458$
Post hoc: Sandwich: Interaction term – reversal in slope for the S group between scored and self-rated knowledge pre- to post-test	0.047	$t(68) = -0.324$ (interaction term in a significant linear model)
Carefulness is significantly correlated with posttest GW attitudes.	0.035	$r(41) = 0.279$
Rereading is significantly correlated with posttest GW attitudes.	0.055	$r(41) = 0.247$
Post hoc: Counter to our initial hypothesis,There is a negative correlation between re-reading and post-test objective knowledge scores. NB: we predicted the opposite result!	0.019	$r(41) = -0.356$

Table 1.5: Summary of "improvement" results for Brownsville classroom interventions. All results were *a priori* unless description starts with *"post hoc"*.

Result	p-value	Statistic
Differences in GW attitudes are significant	1.30×10^{-4}	$t(38 = -4.02$ (paired / imputed)
Use of infrared is greater post-test than pre-test. Goes from 0 to 7 / 6 in Sandwich / No-pretest groups respectively (tested only for Sandwich group)	8.00×10^{-3}	Fisher's exact "two-sided"
Sandwich: pre- to post-test: Increase in self rated knowledge is highly significant	0.001	$t(18) = 18$ (paired)
Sandwich: GHG Objecive knowledge scores improve after the blurb	0.0034	$t(18) = 3.38$ (paired)

Table 1.5: Improvements in Brownsville classroom interventions, continued

Result	p-value	Statistic
No-pretest: GHG Objecive knowledge scores improve after the blurb	5.8×10^{-4}	$t(33.2) = 3.81$ (Welch)
Sandwich: Light Objecive knowledge scores improve after the blurb	0.0095	$t(18) = 2.9$ (paired)
No-pretest: Light Objecive knowledge scores improve after the blurb	1.4×10^{-4}	$t(25.6) = 4.48$ (Welch)
Sandwich: Energy Objecive knowledge scores improve after the blurb	0.02	$t(18) = 2.5$ (paired)
No-pretest: Energy Objecive knowledge scores improve after the blurb	2.9×10^{-4}	$t(36.8) = 4$ (Welch)

Table 1.6: Summary of failures to replicate and associated results with Brownsville classroom interventions. All results were *a priori* unless description starts with *"post hoc"*.

Result	p-value	Statistic
No-pretest: post-test increase in self rated knowledge (compared to S pre-) is not significant	0.51	$t(31.4) = -0.036$ (Welch)
Post hoc: Self-rated knowledge on post-test is significantly lower for No-pretest than Sandwich group	0.0019	$t(36.6) = -3.36$ (Welch)
There is no correlation between self-rated knowledge and GW attitudes on the pre-test	0.55	$r(17) = 0.15$
Surprise is not significantly greater in S group than N group	0.18	$t(38.1) = 0.92$ (Welch)

1.2.3 Discussion

This experiment replicates and extends findings from prior interviews and Cohen's (2012b) survey, such that even rather well-educated people initially held mostly non-normative understandings of global warming's mechanism. Only 400 words later, though (roughly the duration of a TV commercial break), dramatic increases were observed in (1) mechanistic knowledge and (2) global warming acceptance. Further, the increases were found in divergent U.S. states and colleges. Certainly, this suggests that this educational

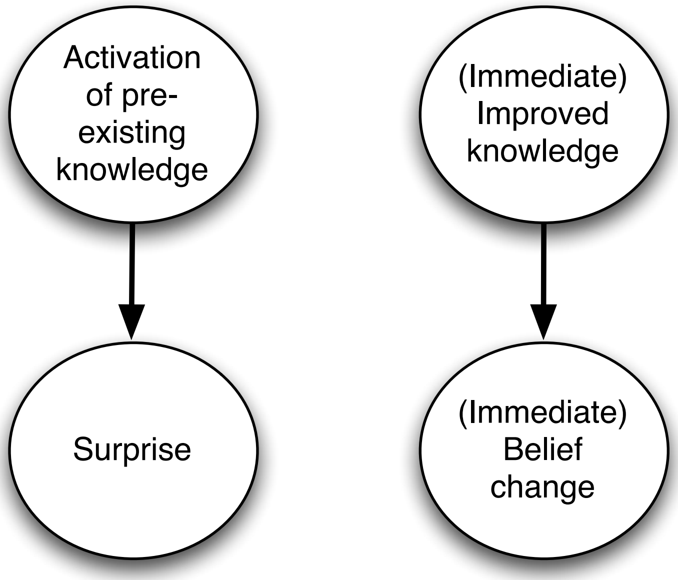


Figure 1.8: A graphical model representing the relationship between forms of psychological processing of factual information observed in chapter 1. Here, the relationship between pre-existing information and surprise was the result of an experimental manipulation and is assumed to be causal. NB: results in the Brownsville data were numerically consistent with the significant result found with the Berkeley data. Namely, there appears that activating pre-existing knowledge boosts surprise.

intervention is a reasonable object of study! Differences in surprise ratings between the sandwich and no-pretest ("open-faced") groups further support the notion that eliciting an explanation or theory prior to offering information increases surprise and reduces post-hoc rationalization and hindsight bias. (On surprise, see Chapter **??**; Munnich, Ranney, & Song, 2007.) A graphical depiction of the probabilistic relationships we've established so far are in Figure **??**.

In addition, we may note that there is scant difference between our sandwich and no pre-test groups in terms of post-test attitudes (with differences going in opposite directions across our two groups). Thus, it seems unlikely that a pre-test incurs a greater burden of experimental demand over the core intervention (400 words followed by a post-test). Moreover, in some populations, the pre-test may help to anchor self assessment (as with our UT data). And finally, the sandwich intervention appears to increase reported feelings of surprise, and likely decreases post-hoc bias. Given these many benefits, in the work that followed, we standardized on using a sandwich style intervention.

1.3 Study 2: A web-based intervention with UC Undergrads

Given the replicated demonstrations of significant attitude changes described above, we proceeded to assess whether the mechanism-explanation effects we had obtained were durable rather than transient. This study extended prior work by employing a followup test several days after our post-test. We also wondered whether any "experimental demand" from the classroom setting might have driven our prior results, so we provided the intervention on-line; that is, we assessed whether our materials would elicit significant attitude change even though students participated via their own computers, without experimenter observation. Thus we concurrently explored the longevity (via delayed followup) and format (online) aspects of our phenomenon. We also extended our prompts to incorporate more demographic and introspection queries.

1.3.1 Methods

Materials and Procedure

The survey and instructional materials were largely analogous to those reported in Section 1.2. The primary difference was that administration was conducted entirely online, via the Qualtrics Inc. (Provo, UT) system. Eight items were added to pre- and post-test attitude surveys to add reliability to the related RTMD metrics (specifically national and religious affinities; these metrics will be reported elsewhere). The engage item was added to provide a clearer assessment of behavioral intentions. Five further questions were introduced immediately following the instructional material to elicit introspection (about embarrassment, disagreement, etc.). These questions were distilled from the result of guided interviews with participants in a pilot study quite similar to this one, but carried out with participants using a computer in our lab testing space.

RPP recruitment allowed us to administer a pre-test to about half of the students ($n = 36$, one was removed due to technical problems) between 8 and 26 days ($\mu = 18.5$ days) before any of the 80 participated in the study, which may have allayed test-retest effects (although we found little evidence for them in the experiments reported above). Thus, as with the above, some participants received the full survey testing "sandwich" while others lacked the demographics and attitude portion of the pre-test. Note that this differs from previous "no-pretest" groups, in that this group still provided their naïve description of the mechanism of global warming). Also note that demographics were *also* collected following the primary intervention, as we did not have them for all participants. Thus, for some participants, we collected some demographic information twice (some of which is reported on below). A delayed post-test was given to all participants between 1 and 8 days later ($\mu = 4$ days). This followup test had the same format as the immediate post-test. This range of delays prior to followup was used to assess the timecourse of retention in planning subsequent studies. We lack the power to test forgetting over time here, and numerically we did not observe any.

Participants

Undergraduates (N = 80) were recruited via the Research Participation Program (RPP), administered by the University of California, Berkeley (UCB) psychology department. For this study, I specifically recruited conservative individuals based on data from the RPP pre-screening phase (completed 324 times, though some students complete the form more than once—even though they get no additional credit!). One participant was excluded due to technical problems with his data. (this was one one of the participants who took the RPP pre-test).

Of the analyzed participants, 46 were female, 33 male. In addition, this study reveals that (at least for Berkeley undergraduates), reported political party affiliation is highly unstable! Of the 36 students who *did* take the RPP pre-screening survey, we obtained two reports of political party. 12 of the 36 students reported something different before and after, though they did not shift in a coherent way, as shown in Table 1.7. The complete breakdown of stated party affiliation is given in Table 1.8.

Table 1.7: Stated party affiliation for the 12 of 36 individuals for whom we have two responses for political party. It is difficult to discern a clear pattern here, beyond a general instability in stated party affiliation. For example, an equal number of "none" responses shifted to "democrat" and "republican." Similary, three "democrats" shifted away, while four individuals shifted to become "democrats." Notably, no individuals shifted to "green."

RPP pre-test	After Intervention
democrat	libertarian
democrat	none
democrat	independent
independent	democrat
independent	libertarian
none	democrat
none	democrat
none	republican
none	republican
none	independent
decline to state	none
decline to state	democrat

A final demographic measure of interest is conservativism. Given the large number of students who completed the RPP pre-test survey, we were able to selectively contact

Table 1.8: Total number of individuals reporting affiliation with a given party. Note that we aggressively recruited conservatives to participate from the RPP pre-screening population, yet still ended up with more republicans coming from the remainder of our population!

	RPP pre-test (fraction)	Study post-test (fraction)
decline to state	3.00 (0.08)	1.00 (0.01)
democrat	12.00 (0.33)	26.00 (0.33)
green	0 (0)	1.00 (0.01)
independent	2.00 (0.06)	3.00 (0.04)
libertarian	1.00 (0.03)	4.00 (0.05)
none	17.00 (0.47)	37.00 (0.47)
republican	1.00 (0.03)	7.00 (0.09)
total	36.00 (1.00)	79.00 (1.00)

those that we identified as the most conservative of the 300+ students who took the pre-test. The RPP survey is standardized and shared with all experimenters. In this case, the RPP survey includes 2 questions for economic and social conservativism, while our survey only includes a question for "conservativism." A full treatment of the relationship between these variables is not justified here. Briefly, social conservativism is quite healthily correlated with our later measure of non-specific conservativism ($r = 0.84$; economic conservativism was correlated at 0.68). The addition of economic conservativism as a second predictor or as a way to generate a mean conservativism score on the pre-test *does* increase goodness of fit, but the reductions in error are small. Moreover, the Bayesian Information Criterion (BIC) weighs heavily in favor of a model using only social conservativism. However we slice things, though, participants reported higher values for non-specific conservativism after the intervention as compared to economic *or* social conservativism on the RPP pre-test, as shown in Table 1.9. Certainly, there is no evidence here that folks have moved in a less conservative direction! It does seem that recruitment of conservatives may have been somewhat successful, based on the 0.5-point higher score for participants who did the RPP pre-test (and were thus the subject of targeted conservative recruitment).

Analysis

As we have now collected 3 observations of the same measure within subjects, for many of the tests in this study, we shifted from using t-tests to using robust mixed-effects regression. For this, we used the lmer function from the lme4 package in R, which handles a variety of unbalanced designs. Goodness of fit for these regression models was assessed utilizing type-II sums of squares via the Anova function from the car package. Once a null

Table 1.9: Conservativism scores from the RPP pre-test and intervention post-test. Note that scores are not directly comparable, as the differing levels of specificity entailed somewhat different wording on the questions. That said, individuals were certainly close to the middle of the scale, and if anything moved closer to the middle on average

	Reported conservativism		
	Social (pre-test)	Economic (pre-test)	Non-specific (post-test)
Full Sandwich	3.86	3.25	4.31
No (RPP) pre-test			3.80

model was rejected, Tukey-style contrasts were computed as simultaneous comparisons with glht from the multcomp package. Unless otherwise stated, all tests were *a priori*.

1.3.2 Results

As before, a full report of statistical tests for this study is given at the end of this section in Table 1.15. Statistics are provided in-text when appropriate.

Learning the Global Warming Mechanism

In general and as anticipated, we replicated results from Study 1 and extended them by finding that shifts were retained over the mean (four-day) delay. Scored knowledge was comparable to previously tested UC students, rising from 3.8 on pre-test to 6.5 post-test and 6.3 on delayed test (gains from pre-test were significant at $p < 0.0001$ for both subsequent scores). These are plotted in Figure 1.9.

Self-rated knowledge means also increased markedly from pre- to post-test (4.5 to 5.6 on a 9-point scale, $t(79) = 8.5$, $p < 0.001$). Retention of this increase, gratifyingly, was also noted on the delayed post-test (5.2, $t(79) = 6.2$, $p < 0.001$). The immediate increase in self-rated knowledge, replicates results from the "sandwich" interventions in Study 1. Scores are shown in Figure 1.10.

Global Warming Acceptance Via Mechanistic Learning

GW belief ratings (with higher ratings being more in concert with science's consensus) increased from a 6.20 pre-test mean to a 6.54 post-test mean ($t(79) = 2.5$, $p = 0.006$; a healthy improvement on our 1–9 Likert scales!). Some of this improvement diminished over the following days, but most was retained: the mean score on the delayed post-test was 6.44 ($t(79) = 1.7$, $p = 0.05$).

A more recent analysis, however, implies that these results may not be so clear. In particular, using the lmer package yeilded insignificant improvements even from pre- to

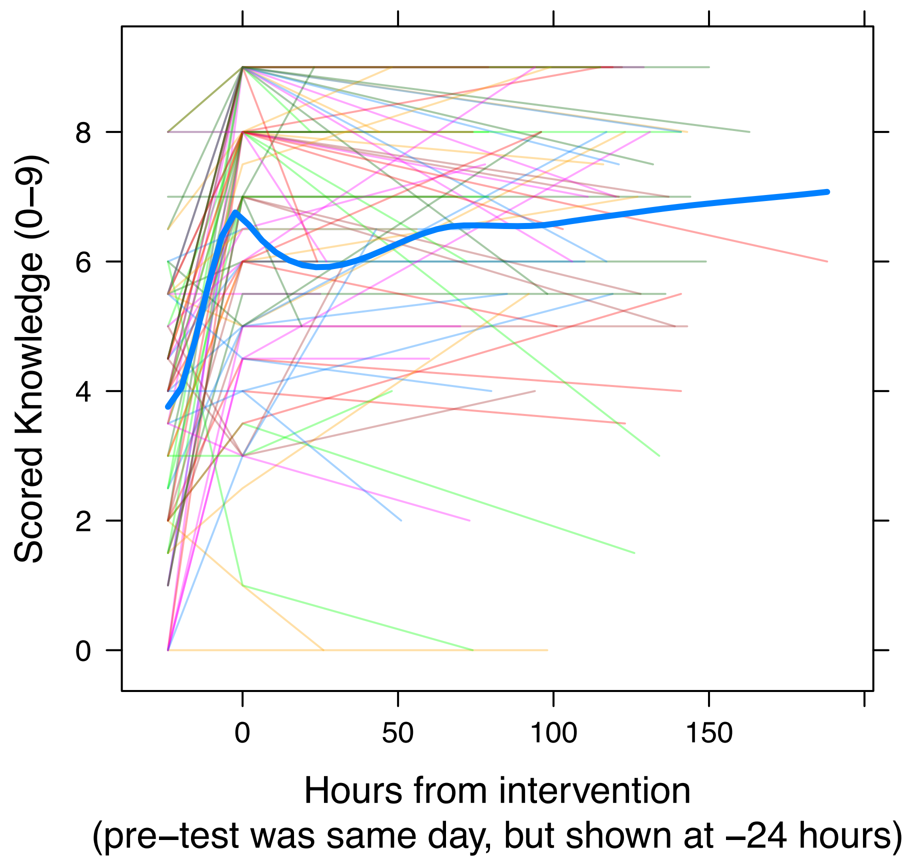


Figure 1.9: Scored knowledge before and after our intervention. Faint lines represent individual performance, while the strong line is a LOESS robust, smooth regression line. Note that the pre-test here was immediately prior to instruction. Participants overwhelmingly increased their scored knowledge following the intervention. While this study was not designed to assess forgetting in an individual over time, it is interesting to note that earlier respondents tended to be scored lower than later respondents. Most importantly, it is easy to spot individuals who improved or worsened over the delay.

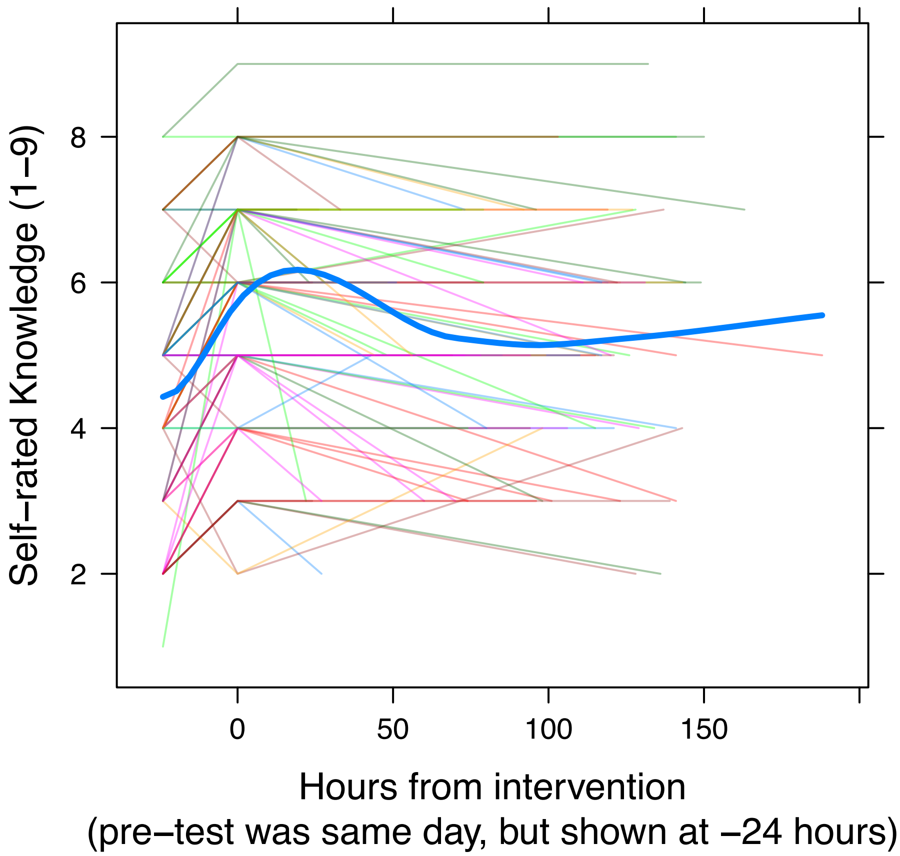


Figure 1.10: Self-rated knowledge before and after our intervention. Note that the pre-test here was immediately prior to instruction.

post-test. Upon closer inspection, the breakdown for GW attitudes is quite different across sandwich and no pre-test groups, as seen in Table 1.10. Unlike in previous studies, the Sandwich group contained a good number of participants who were recruited for their conservativism (and there is a clear group difference on this measure). Thus, we are not as justified in assuming the pre-test is representative of *all* participants. Individual scores are depicted in Figure 1.11.

BUT we have over 300 folks from RPP. So, we should be able to address this at least partially that way. Throw everyone into one big model?

Table 1.10: Mean GW ratings for Sandwich and No-pretest groups. Note that the increase from pre-test to post-test is only 0.16 for individuals in the Sandwich group, and delayed test results are lower than where they started. The No pre-test group, however, starts much higher and stays higher.

	pre	pst	fol
No (RPP) pre-test	NA	6.68	6.65
Full Sandwich	6.20	6.36	6.18

Surprise

As explained in the Methods, some participants recieved a full "Sandwich" intervention, including beliefs and attitudes administered during RPP pre-screening. All participants, however, completed a knowledge pre-test. Thus, we were able to informally compare these two partially novel conditions with our previous profiles for reported surprise (i.e., with Figure 1.7). The surprise results from our current conditions are provided in Figure 1.12, and provide further support for the centrality of the initial *knowledge* pre-test.

In addition to repeating the surprise question from Study 1, on the basis of informal interviews, we had included another related question asking if people were "surprised or embarrassed about their own lack of knowledge." This item does seem to have elicited somewhat higher scores, with a mean of 3.6 vs. 2.9 for the original surprise question. The distributions for these responses are shown in Figure 1.13.

Micro-analysis of GW ratings

Table 1.11 reports the mean rating across participants for agreement with individual items (For the full text of these items, see Appendix B). The largest gains were found in agreeing with "Human activities are largely responsible for the climate changes…" (a 0.25 gain) and certainty that global warming is occurring (a 0.19 gain). In general, gains were fairly consistent across all GW measures, ranging only down to 0.11 at the lowest (the über-greenie "humans are severely abusing the environment"). Interestingly importance

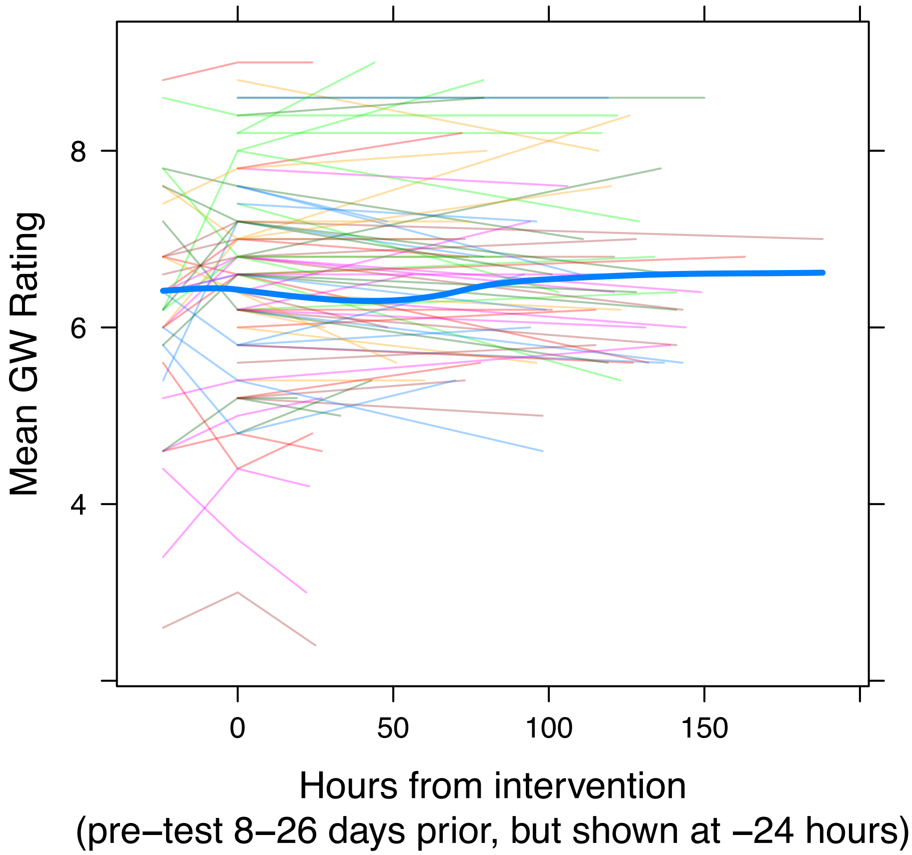


Figure 1.11: Reported global warming (GW) beliefs and attitudes before and after our intervention. Note that the pre-test here was administered on average 18.5 days prior to instruction as a part of UC Berkeley's undergraduate participant pool (RPP) pre-screening. The fact that over half of our participants are missing pre-test data pulls the initial value of our regressor up markedly from the true pre-test mean of 6.2. We can also notice here that all of the lowest-rating individuals took the delayed test quite early. We should certainly be taking those initial dips with a grain of salt. We cannot assume that's an actual retention timecourse!

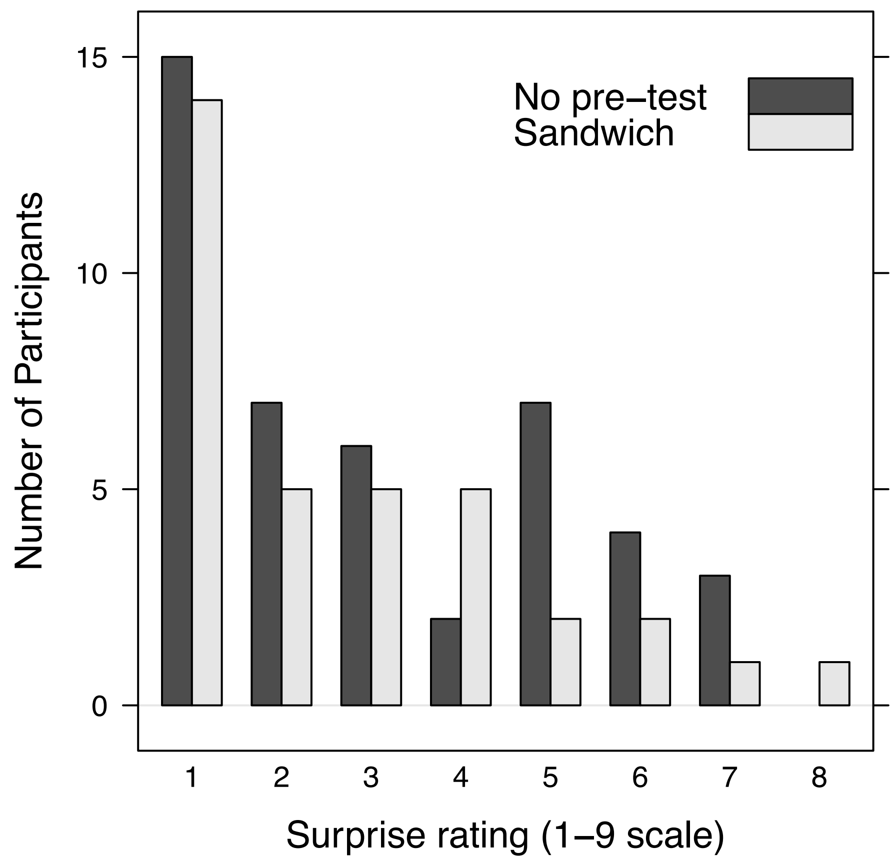


Figure 1.12: Surprise ratings for individuals in Study 2. Note that there is very little difference between subjects that had recently taken the knowledge and attitude pre-test or not. Moreover, these distributions are analagous to "sandwich" participants in previous studies. Thus, it seems to matter little if and when participants recieve the beliefs and attitutdes surveys—the higher surprise values appear to be linked to asking participants to "put their cards on the table" with a knowledge pre-test.

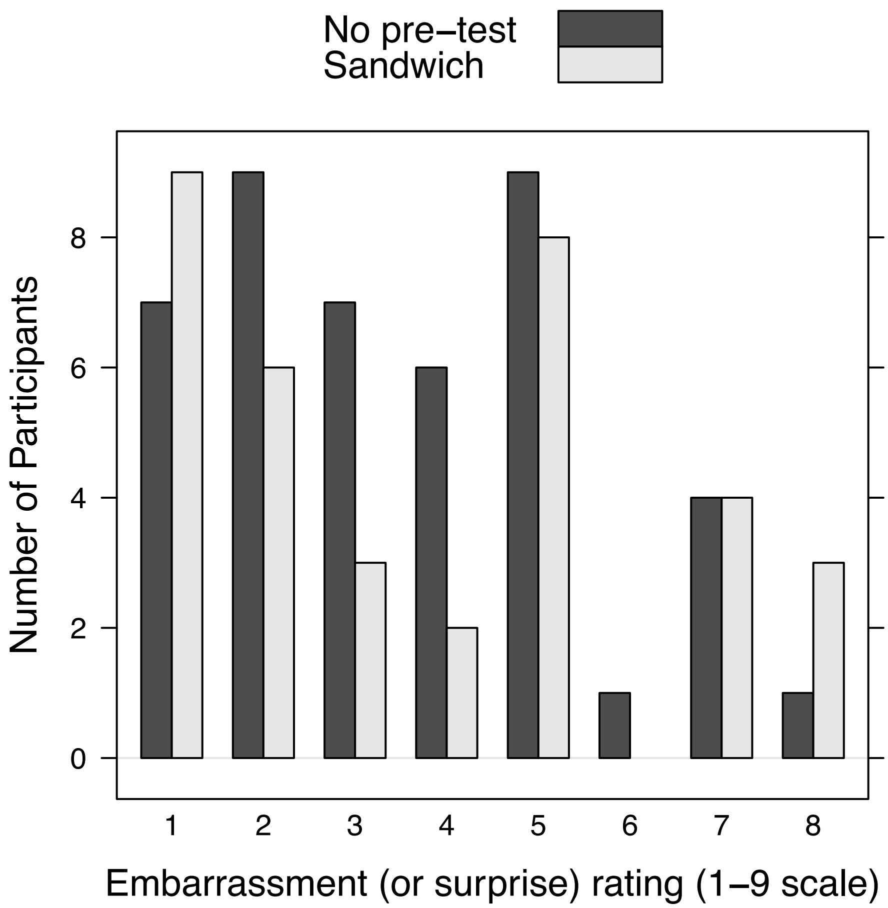


Figure 1.13: Ratings of embarrassment or surprise at lack of knowledge for individuals in Study 2. Distributions are again not obviously different between groups. However, there appears to be less of a floor effect than we obtain with our original phrasing of surprise.

Table 1.11: Mean GW ratings, online with UC undergrads. For means, please refer to Table 1.10

	Pre-test	Post-test	Delayed test
gw1_2	6.61	6.86	6.36
gw2_1	5.19	5.31	5.25
gw2_2	6.61	6.81	6.67
gw2_3	5.81	5.97	5.97
gw2_4	6.78	6.86	6.67
engage	5.91	5.86	6.11
lifesty	4.83	5.11	4.94

of lifestyle changed the most (0.27, though this was not included in the tested average gw variable). Expectation of engagement, dishearteningly, clocks in at a 0.05 *drop*!

Correlations

Scored knowledge and self-rated knowledge are significantly correlated pre-test, so participants have decent meta-cognition here. However, unlike with our Berkeley students in Study 1, there was only a marginally significant (and smaller) correlation between naïve pre-test self-rated knowledge and GW attitudes ($r(34) = 0.27$, $p = 0.11$). Thus, while the metacognitive aspect of self-rated knowledge appears robust across Berkeley student populations, the relationship with GW attitudes is weak, or perhaps even spurious.

1.3.3 Discussion

In sum, this study extends the finding that well-considered information, even received online, increases anthropogenic global warming acceptance and behaviorally relevant attitudes; the conceptual changes that result from reading even 400 words have notable longevity. These effects have been replicated with members of the general public as well (unpublished data). Computer-based interventions often scale well, enhance reliability, and prove cost-effective; given our results, we recommend the online distribution of mechanistic explanations, especially about climate change.

Followup tests occurred over a range of delays, and retention here was used in determining intervals for future studies. Given the almost total lack of mean forgetting over the observed interval, we would start our subsequent study's delayed test after a delay of a full week.

Table 1.12: Summary of results from Study 2.

Result	p-value	Statistic
Scored knowledge is different across tests in a repeated-measures Anova.	2.2×10^{-16}	$\chi^2(2) = 128.39$
Scored knowledge is higher in post-test than in pre-test.	5×10^{-3}	$z = 10.09$
Scored knowledge is higher in delayed test than in pre-test.	5×10^{-3}	$z = 9.52$
Self-rated knowledge is different across tests in a repeated-measures Anova.	2.2×10^{-16}	$\chi^2(2) = 110.25$
Self-rated knowledge is higher in post-test than in pre-test.	1×10^{-5}	$z = 10.47$
Self-rated knowledge is higher in the delayed test than in pre-test.	1×10^{-5}	$z = 5.922$
Self-rated knowledge is lower in the delayed test than in the post-test.	1.5×10^{-5}	$z = -4.55$
GW Attitudes (multiple tests).		
Self-rated knowledge is significantely correlated with scored knowledge on the pre-test.	4.7×10^{-6}	$r(76) = 0.49$
Self-rated knowledge is *not* significantly correlated with GW attitudes on the pre-test.	0.11	$r(34) = 0.27$

1.4 Study 3: An intervention on Amazon Mechanical Turk

1.4.1 Methods

Experimental methods in this study were nearly identical to Study 2, above (detailed in Section 1.3.1). The primary difference here was that participants were recruited via Amazon's Mechanical Turk (discussed further in Section 1.4.1. In general, the remainder of the methods focuses only on the differences with those reported in Section 1.3.1.

Materials and Procedure

On the basis of our previous two studies, the most convenient design that still elicits the highest surprise was an immediate pre-test including all measures—eliciting descriptions prior to instruction maximized surprise, and the timing of other attitudinal measures seemed to have little effect (or may have actually reduced our power to observe shifts in attitudes). Thus, here, all participants completed a pre-test immediately prior to the main intervention and post-test, most closely approximating the sandwich intervention from Study 1 with the addition of a delayed test some days later.

The most interesting difference in procedure between this and Study 2 is the utilization of Mechanical Turk. A minor change is that the retention interval was extended to begin 4 days later, with the longest interval being 10.8 days ($\mu = 5.5$ days). A reminder was sent throught the Mechanical Turk bonus system, where participants were paid 5 cents along with a message that the delayed test was open for completion. The majority of payment was provided at the time of completion of the delayed test. Perhaps unsurprisingly, most participants responded shortly after the first reminder, and the majority of the remainder responded shortly after the final reminder.

Materials were identical, apart from the addition of a question requesting the individual's "worker ID"—a unique identifier used to track and pay individuals, akin to a social security number assigned by Amazon.

Data quality on Amazon Mechanical Turk

The use of an anonymous, on-line labor pool raises a number of additional concerns about data quality. For example, people may try to take the survey again, they may lie about their demographics (i.e., claiming they are U.S. residents so that they may gain the credit), and (bizarrely, as this does not reduce time required, or increase payment) they may copy and paste from online sources.

Re-taking is one of the most easily guarded against concerns on Amazon Mechanical Turk, as Amazon will attempt to enforce this if requested, as was done in this study. However, in addition, no IP addresses were repeated in the participant group for this study.

Verification of location is fairly straightforward using "geo IP" databases. In this case, geographical locations were retrieved from the freely available **maxmind-database** On the primary survey, participants indicated the state they were in. Participant IP addresses from both intervention and followup surveys were subsequently checked against this reported location. Here, two individuals IP's appeared to be located in India and Turkey, and thus these participants were excluded. Most other participants had IP addresses that resolved to the state they claimed to be from. One participant's IP address was not listed in the MaxMind database, and was traced to Hughes Net, a U.S.-only satellite internet provider.

Checking for plagiarism is a well-trod topic in academia today, though there is no one clear approach. For the purposes of this survey, we relied on a combination of coder judgement, and automated checking using Google's Custom Search API. A complete description of the process I used is provided in Appendix **??**. In summary, I used a combination of my own and others' judgement to identify clear cases of plagiarism. I then developed an approach using google search to identify texts that already existed on the internet, being sure that this method caught the clear cases we'd identified via manual inspection. Ultimately, 2 individuals clearly copied and pasted materials from the internet. These could have both been identified by the unusual presence of extended unicode characters in the text of their answers, as well as the presence of newlines.

Note that one individual was *both* foreign *and* copied text directly form the web. Thus, we exclude only 3 subjects total.

While survey consistency seemed reasonable for all remaining participants, it did highlight two individuals who answered almost exclusively 1 or 9 to all items. These individuals were retained in the analyses below, as they still provided information on their beliefs.

Participants

"Workers" on the Amazon Mechanical Turk platform (N = 41) were recruited to complete a two-part survey. Approximately 75% of these individuals ($n = 30$) completed the delayed test. After removal of problematic participants above, we were left with 38 participants in the primary intervention, with 28 in the followup. 17 of the total retained participants were female. Two of our participants report as being born outside the United States, but residing here for at least 22 years. Stated party affiliations are listed in Table 1.13. Mean conservativism was 3.9. Compared with our college students above, conservativism was in the same ballpark and participants were far more likely to declare Democrat or Republican than with our college students above. While the sample is still clearly biased towards the Democratic / liberal end of the political spectrum, the ratio between Republicans and Democrats is far less extreme than previous samples.

Table 1.13: Stated party affiliations for participants in Study 3 (via Mechanical Turk).

Party	Number (percentage)
democrat	22 (58%)
republican	8 (21%)
independent	6 (16%)
none	1 (3%)
other	1 (3%)
total	38.00 (100%)

A final interesting feature of our population is that all participants reporting their faith as Christian self-identified as Catholic. That is, we had no "protestant" participants.

1.4.2 Results

Overall, we replicate our central results from above. As before, full statistics are reported in a table at the end of this section. Only information directly supporting the narrative is included in the textual exposition.

Problems and Data Quality

While the text of the followup survey was like the pre-test (i.e., the first question made no affordance for "if you would add anything") three individuals still put "nothing to add" for all three questions. These individuals were retained as, if anything, they should weaken followup retention effects. Moreover, they all still answered the likert survey questions. While not definitive, this is certainly one clue that individuals recruited via Mechanical Turk are perhaps rushing a bit more than our previous college samples.

Learning and Global Warming Mechanism

Again, we replicated the above results by finding that shifts were retained over the mean 5.5-day delay. Scored knowledge was comparable to previously tested students, rising from 1.9 on pre-test to 4.8 post-test and 3.9 on delayed test (gains from pre-test were significant at $p < 0.002$ for both subsequent scores). These are plotted in Figure 1.14.

Self-rated knowledge means also increased markedly from pre- to post-test (4.2 to 4.7 on a 9-point scale, $p < 0.01$). Retention of this increase, however, failed to exceed our significance threshold (4.5, $p = 0.21$). These scores are reported in Figure 1.15. It seems not unlikely, given our previous results, that more subjects would yield a significant difference here. Even so, it should be noted that studies with college students tended to yield increases of over a point on self-rated knowledge (at least with "sandwich" interventions such as this one). Thus, even our significant pre- to post-test gains are relatively small.

Global Warming Acceptance Via Mechanistic Learning

Happily, while introspection regarding knowledge seemed to yeild somewhat smaller gains than in Studies 1 and 2, GW belief ratings increased significantly from a 6.34 pre-test mean to a 6.64 post-test mean ($p = 0.001$). Some of this improvement diminished over the following days, but most was retained: the mean score on the delayed post-test was 6.58 ($p = 0.006$). Note that these values are very much in line with those obtained in the previous studies. Individual attitude breakdowns are reported in Table 1.14. Individual data are depicted in Figure 1.16.

Surprise

As in Study 2, individuals ranked their embarrassment or surprise at their own lack of knowledge higher than straight surprise. Mean values were 2.9 and 4.1 respectively, quite comparable to Study 2. Distributions for both questions are depicted in Figure 1.17).

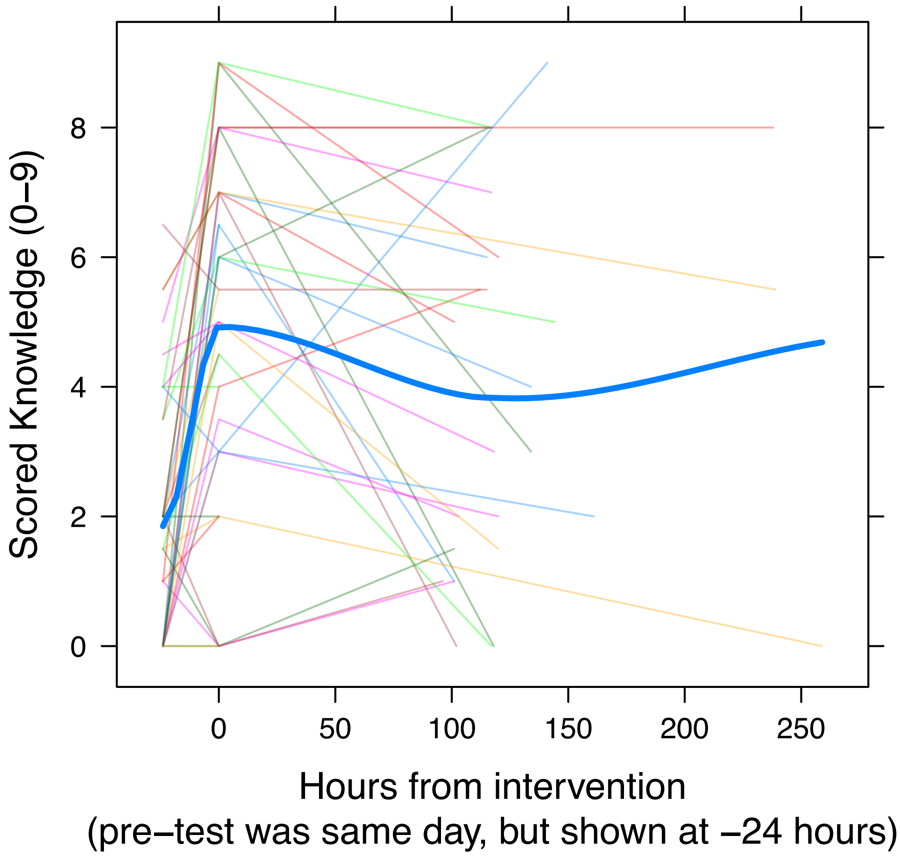


Figure 1.14: Scored knowledge before and after our intervention. Faint lines represent individual performance, while the strong line is a LOESS robust, smooth regression line. Participants overwhelmingly increased their scored knowledge following the intervention. As before, this study was not designed to assess forgetting in an individual over time, but it is interesting to note that we again obtain relatively high scores later in our delayed testing.

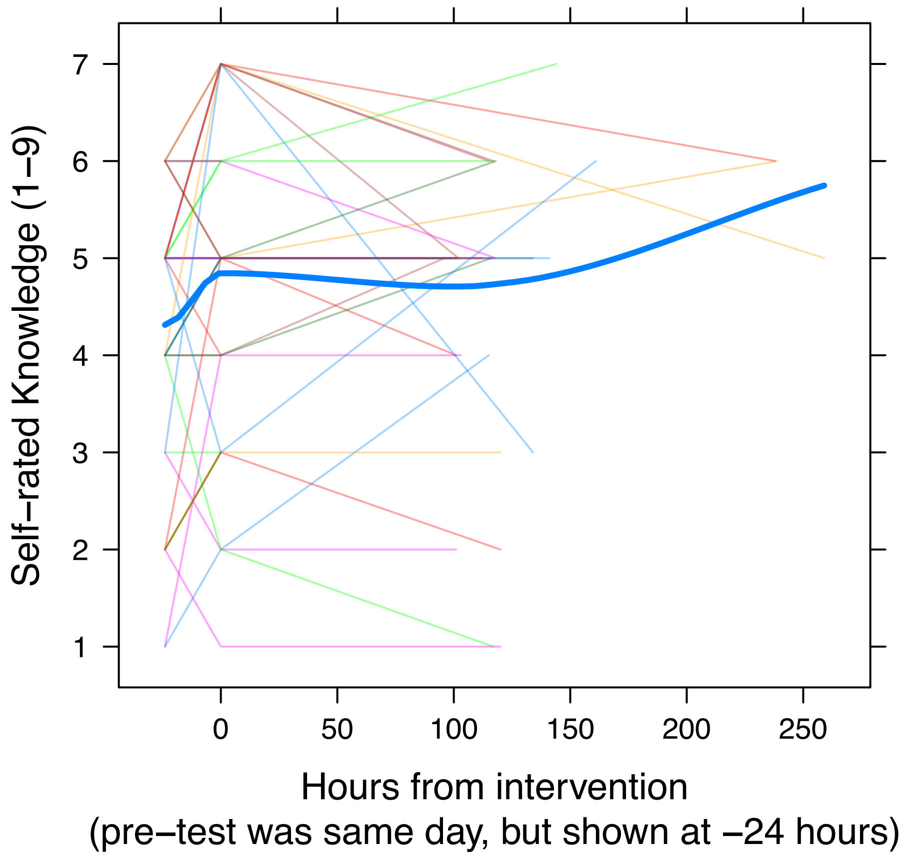


Figure 1.15: Self-rated knowledge before and after our intervention. Note that the pre-test here was immediately prior to instruction.

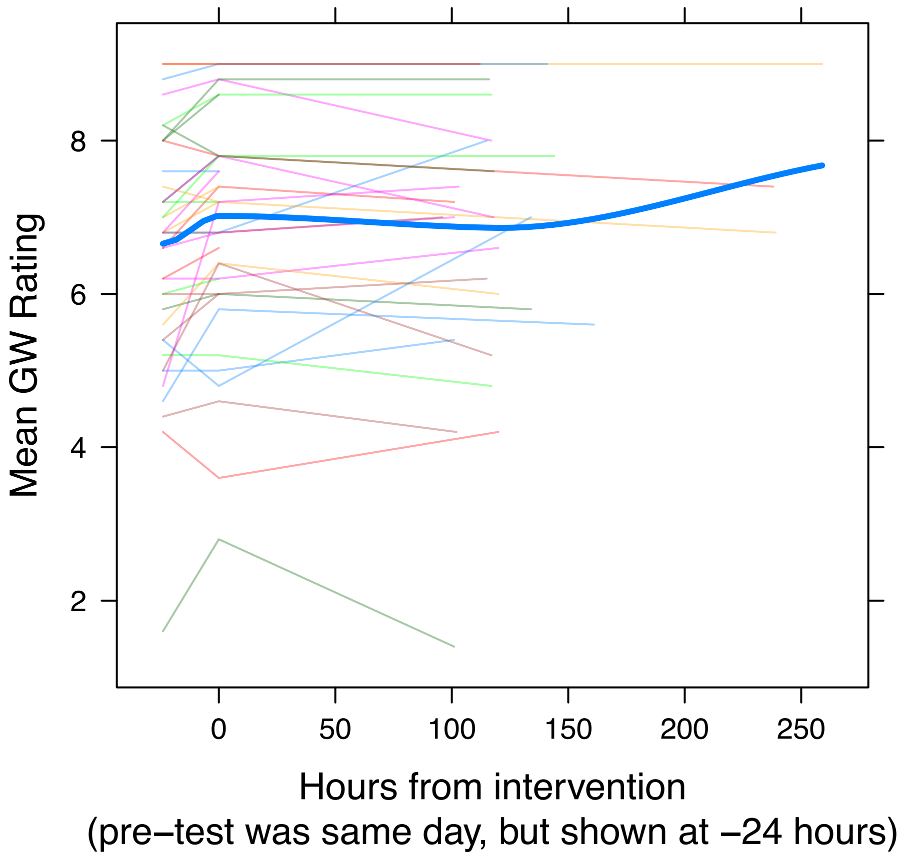


Figure 1.16: Reported global warming (GW) beliefs and attitudes before and after our intervention on Mechanical Turk.

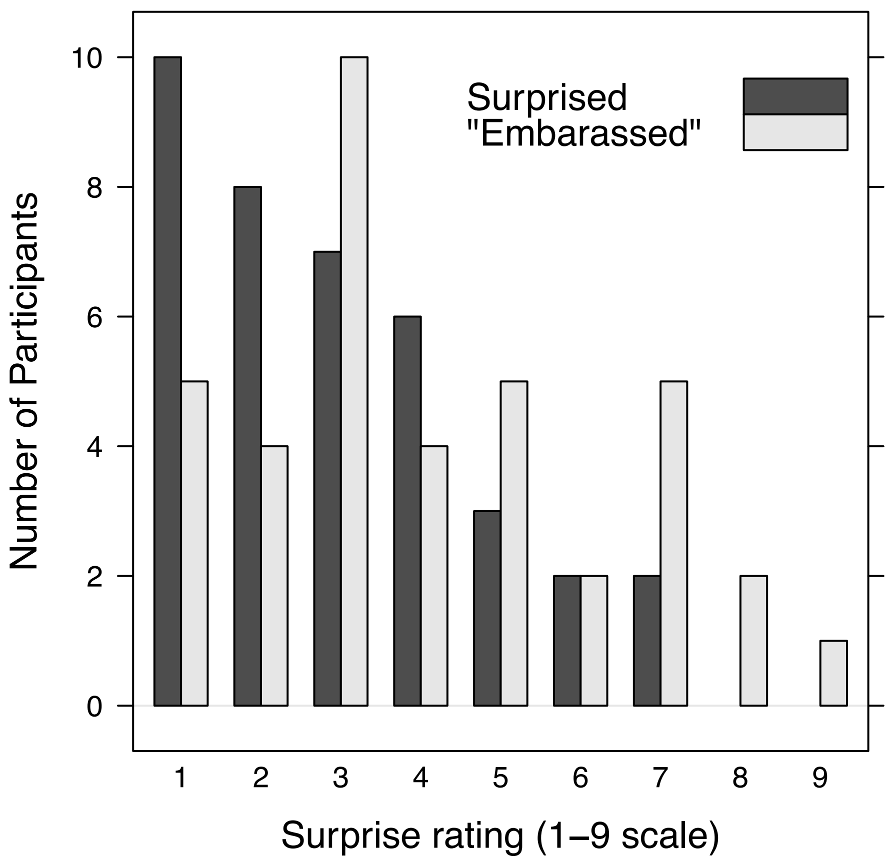


Figure 1.17: Surprise and “embarassment” ratings for individuals in Study 3.

Table 1.14: Mean GW ratings for the Mechanical Turk mechanism intervention. Note that means are only computed over the items with codes starting with "gw".

	Pre-test	Post-test	Delayed test
gw1_2	6.87	7.34	7.07
gw2_1	6.03	6.55	6.25
gw2_2	6.82	7.05	7.04
gw2_3	6.05	6.42	6.36
gw2_4	6.89	7.05	6.86
engage	6.37	6.42	6.46
lifsty	5.37	5.63	6.04
mean	6.34	6.64	6.58

Correlations

As with our UC students, participants demonstrated a significant correlation between self-rated knowledge and actual knowledge ($r(36) = 0.49$). And, as in Study 2, we again failed to replicate a significant relationship between self-rated knowledge and GW attitudes on the pre-test. Given these repeated failrures to replicate, it seems prudent to abandon the naïve relationship between self-rated knowledge and GW attitudes (at least in this general sense).

1.4.3 Discussion

This study provides an evaluation of our Global Warming mechanism intervention in a much more true-to-life scenario. Specifically, we should be concerned with evaluating a population that is representative of individuals that might engage with our materials were they to be made generally available on line. The participants we obtained here on Mechanical Turk are likely much closer to this ideal than our college students above.

Table 1.15: Summary of results from Study 2.

Result	p-value	Statistic
Scored knowledge is different across tests in a repeated-measures Anova.	2.7×10^{-8}	$\chi^2(2) = 34.87$
Scored knowledge is higher in post-test than in pre-test.	5×10^{-5}	$z = -5.86$

Table 1.15: Results from Study 2, continued.

Result	p-value	Statistic
Scored knowledge is higher in delayed test than in pre-test.	0.0014	$z = 3.30$
Self-rated knowledge is different across tests in a repeated-measures Anova.	0.21	$\chi^2(2) = 7.69$
Self-rated knowledge is higher in post-test than in pre-test.	0.01	$z = 2.72$
Self-rated knowledge is *not* significantly higher in the delayed test than in pre-test.	0.21	$z = 1.69$
Mean GW attitude is different across tests in a repeated-measures Anova.	0.000 91	$\chi^2(2) = 14.01$
Mean GW attitude is higher in post-test than in pre-test.	0.000 85	$z = 3.445$
Mean GW attitude is higher in delayed test than in pre-test.	0.0062	$z = 2.84$
Self-rated knowledge is significantely correlated with scored knowledge at pre-test.	0.0017	$r(36) = 0.49$
Self-rated knowledge is *not* significantly correlated with GW attitudes at pre-test.	0.32	$r(36) = 0.16$

1.5 Summary and Conclusions

We've shown across a number of populations that ignorance of the basic physical/chemical mechanism of the greenhouse effect is nearly universal. In addition to this research, Felipe (2012) describes successes with a curriculum involving the mechanism with 11th graders (and see Chapter **??** for more on the numerical estimation aspects of that study). Across a variety of intervention styles, we have shown that individuals are able to markedly increase their abilty to describe the greenhouse effect, and that such an intervention additionally shifts climate-related beliefs and attitudes.

As can be seen from the work described above, the act of educating the American public about the basic mechanism of climate change is a daunting, multi-faceted challenge. And even this is not sufficient to know if such an endeavor will truly effect some positive action towards the larger problem of climate change. For that, we will need to examine some connection to behavior. This is the clearest lack in the current research. Such an endeavor will be even more challenging than the above, but it is of such critical importance!

Overall, however, we have seen evidence that materials such as those exhibited in Appendix D are likely to be effective both in college classrooms as well as on-line. Evidence from related studies has provided additional support for effective application in high school classrooms.

Acknowledgements

The work reported in this chapter has been previously published, in part, in Ranney et al. (2012a, 2012b), and Clark, Ranney, and Felipe (in press). All such material is re-used here with the permission of my co-authors, the publishers, and the Graduate Division at the University of California, Berkeley.

Bibliography

Clark, D., Ranney, M. A., & Felipe, J. (in press). Knowledge helps: mechanistic information and numeric evidence as cognitive levers to overcome stasis and build public consensus on climate change. In *Proceedings of the 35th annual meeting of the cognitive science society*. Austin, TX: Cognitive Science Society.

Cohen, S. (2012a). *San Diego's IPCCC: the implications and politics of climate change confusion* (Undergraduate thesis, University of California, Berkeley).

Cohen, S. (2012b). *San diego's IPCCC: the implications and politics of climate change confusion* (Undergraduate thesis, University of California, Berkeley).

Felipe, J. (2012). *Numerical reasoning, knowledge, and environmental behavior regarding climate change* (Masters thesis, University of California, Berkeley).

Martinez, C. (2009). *Factors possibly related to evolution acceptance in the U.S.* (Master's thesis, University of California, Berkeley).

Ranney, M. A. (2012). Why don't americans accept evolution as much as people in peer nations do? a theory (reinforced theistic manifest destiny) and some pertinent evidence. In K. Rosengren, M. Evans, G. Sinatra, & S. Brem (Eds.), *Evolution challenges: integrating research and practice in teaching and learning about evolution* (pp. 233–269). Oxford: Oxford University Press.

Ranney, M. A., Clark, D., Reinholz, D. L., & Cohen, S. (2012a). 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.

Ranney, M. A., Clark, D., Reinholz, D. L., & Cohen, S. (2012b). Improving americans' modest global warming knowledge in the light of RTMD (reinforced theistic manifest destiny) theory. In K. van Aalst, K. Thompson, M. M. Jacobson, & P. Reimann (Eds.), *The future of learning: proceedings of the tenth international conference of the learning sciences* (Vol. 2, pp. 2-481 to 2-482). International Society of the Learning Sciences, Inc.

Appendix A

Extension to Cohen's κ for multiple scores

Here, I'll describe exactly how I computed those κ's for the mechanism studies.

A note on using κ in the text above: it seems completely admissable to continue to use κ for the measure described above, given it's similarity in spirit to Cohen's original measure. This seems doubly admissable considering that formal distributional considerations appear to be considered rarely, if ever, in the psychological literature. Thus, any deviations in the specifics of formal properties are unlikely to be a concern for the casual reader.

Appendix B

Survey Items Used in Chapters 1, ??, & ??

Table B.1: Demographic questions. Numeric codes assigned by survey software are given in parentheses and not shown to participants. If no list of choices is given for a question, a blank space was provided.

Question ID	Wording / Answers
state	In what U.S. state or territory to you reside? Please use the 2-letter abbreviation (e.g., MD, DC, PR)
gender	What is your gender?
	Male (1), Female (2)
citizen	Are you a U.S. citizen or permanent resident?
	Yes (1), No (2)
us_born	Were you born in the U.S.?
	Yes (1), No (2)
english	Is English your first language?
	Yes (1), No (2)
party	What is your strongest political party affiliation?
	None (1), Democrat (2), Green (3), Independent (4), Libertarian (5), Republican (6), Other (7), Decline to State (8)
conserv	On the following scale, indicate the extent to which you consider yourself to be liberal or conservative on most political and social issues:

Table B.1: Demographic questions, continued

Question ID	Wording / Answers
	1 Extremely Liberal, 2, 3 Somewhat Liberal, 4, 5 Moderate, 6, 7 Somewhat Conservative, 8, 9 Extremely Conservative (numerical scale given to participants)
faith	What is your main religious faith?
	Atheist (1), Agnostic (2), Buddhist (3), Christian (4), Hindu (5), Jewish (6), Muslim (7), Spiritual but not religious (8), Other (9), Decline to state (10)
ed_level	What is the highest level of education you have completed?
	No higher than 8th grade (1), Some high school (9-12th grade) (2), High school diploma / GED (3), Some college (4), Bachelor's degree (5), Master's degree (6), Professional degree (7), Doctoral degree (8)
cc_inst	Have you received any instruction regarding Global Warming (Climate Change) in the last 2 years? If so, when was the most recent?
	No (1), Fall 2012 (2), Summer 2012 (3), Spring 2012 (4), Winter 2011-2012 (5), Fall 2011 (6), Summer 2011 (7), Spring 2011 (8), Winter 2010-2011 (9)
cci_desc	Please describe the instruction you've received regarding Global Warming (Climate Change)
us_years	You indicated that you were born outside the U.S. How many years have you been living in the U.S.? (Please round up to the nearest whole number.)
where_born	Where were you born?
eng_years	How many years have you been speaking English?
first_lang	What is your first language?

Table B.2: Belief / attitude ("RTMD") questions. Most items on a 1–9 scale (Noted below the table). Scales for items that deviate are reported below the question.

Question ID	Wording / Answers
evo1_1	Evolution accurately explains how plants, animals, and humans came to be as they are.

1–9 scale usually consisted of: Extremely Disagree 1, Strongly Disagree 2, Disagree 3, Mildly Disagree 4, Neither Agree Nor Disagree 5, Mildly Agree 6, Agree 7, Strongly Agree 8, Extremely Agree 9. Where different, the scale is noted in the table next to the question above.

Table B.2: Belief / attitude ("RTMD") questions.

Question ID	Wording / Answers
gw1_2	Human activities are largely responsible for the climate change (global warming) that is going on now.
nat1_3	The United States is one of the very best countries on our planet (e.g., "in the top three").
dty1_4	There exists a supernatural being/deity (e.g., God) or set of beings/deities (gods).
aft1_5	After a person dies, that person experiences an afterlife of some sort (for instance, heaven/hell, reincarnation, enlightenment, nirvana, etc.).
cre1_6	Biblical creation accurately explains how plants, animals, and humans came to be as they are.
gw2_1	Global warming or climate changes, when they happen at all, are just parts of a natural cycle.
gw2_2	I am certain that global warming is actually occurring.
gw2_3	I am worried about global warming.
gw2_4	Humans are severely abusing the environment.
evo2_5	Evolution is unable to explain much of the physical evidence regarding the origins and development of life on Earth.
evo2_6	Other living things may have evolved, but humans have not.
lifsty	Overall, how important is it to change your current lifestyle to reduce your carbon footprint (i.e., to decrease the amount of greenhouse gases you emit both directly and indirectly)? Not Important 1, 2, Slightly Important 3, 4, Somewhat Important 5, 6, Very Important 7, 8, Extremely Important 9
engage	I intend to personally engage in more environmentally-friendly (e.g., sustainable, recycling, and/or resource-minimizing) activities in the future, compared to what I do now.
aft2	After a person dies, that person lives on in some way.
aft3	I don't believe that heaven exists.
dty2	God is created by human imagination.

1–9 scale usually consisted of: Extremely Disagree 1, Strongly Disagree 2, Disagree 3, Mildly Disagree 4, Neither Agree Nor Disagree 5, Mildly Agree 6, Agree 7, Strongly Agree 8, Extremely Agree 9. Where different, the scale is noted in the table next to the question above.

Table B.2: Belief / attitude ("RTMD") questions.

Question ID	Wording / Answers
dty3	The only true God is that of my religion.
cre2	A supreme being has never played any role in the origin or development of life on earth.
cre3	To me, creation gives a more satisfying explanation of life on Earth than does evolution.
nat2	The United States can fix just about any problem it might unintentionally create.
natmil	How many countries could defeat the U.S. militarily without assistance from other countries? ______________ Countries
knwgbl	Please indicate how knowledgeable you think you are about climate change—by choosing a number on the 1 (not knowledgeable at all) to 9 (extremely knowledgeable) scale below. Not knowledgeable at all about Climate Change 1, 2, 3, 4, Moderately knowledgeable about Climate Change 5, 6, 7, 8, Extremely knowledgeable about Climate Change 9

1–9 scale usually consisted of: Extremely Disagree 1, Strongly Disagree 2, Disagree 3, Mildly Disagree 4, Neither Agree Nor Disagree 5, Mildly Agree 6, Agree 7, Strongly Agree 8, Extremely Agree 9. Where different, the scale is noted in the table next to the question above.

Appendix C

A Graphical Method for checking Survey Quality

Below, I explain a graphical method for identifying individuals who fall outside of the normal variation in responses across survey items.

Acknowledgements

Particular thanks to Luke Miratrix for his input in developing these methods.

Appendix D

The 400 Words

How does climate change ("global warming") work? The mechanism of the greenhouse effect

[Or: "Why do some gases concern scientists—like carbon dioxide (CO_2)—but not others, like oxygen?"]

Scientists tell us that human activities are changing Earth's atmosphere and increasing Earth's average temperature. What causes these climate changes?

First, let's understand Earth's "normal" temperature: When Earth absorbs sunlight, which is mostly visible light, it heats up. Like the sun, Earth emits energy—but because it is cooler than the sun, Earth emits lower-energy infrared wavelengths. Greenhouse gases in the atmosphere (methane, carbon dioxide, etc.) let visible light pass through, but absorb infrared light—causing the atmosphere to heat up. The warmer atmosphere emits more infrared light, which tends to be re-absorbed—perhaps many times—before the energy eventually returns to space. The extra time this energy hangs around has helped keep Earth warm enough to support life as we know it. (In contrast, the moon has no atmosphere, and it is colder than Earth, on average.)

Since the industrial age began around the year 1750, atmospheric carbon dioxide has increased by 40% and methane has increased by 150%. Such increases cause *extra* infrared light absorption, further heating Earth above its typical temperature range (even as energy from the sun stays basically the same). In other words, energy that gets to Earth has an even harder time leaving it, causing Earth's average temperature to increase—producing global climate change.

[In molecular detail, greenhouse gases absorb infrared light because their molecules can vibrate to produce asymmetric distributions of electric charge, which match the energy levels of various infrared wavelengths. In contrast, non-greenhouse gases (such as oxygen and nitrogen—that is, O_2 and N_2) don't absorb infrared light, because they have symmetric charge distributions even when vibrating.]

Summary: (a) Earth absorbs most of the sunlight it receives; (b) Earth then emits the absorbed light's energy as infrared light; (c) greenhouse gases absorb a lot of the infrared light before it can leave our atmosphere; (d) being absorbed slows the rate at which energy escapes to space; and (e) the slower passage of energy heats up the atmosphere, water, and ground. By increasing the amount of greenhouse gases in the atmosphere, humans are increasing the atmosphere's absorption of infrared light, thereby warming Earth and disrupting global climate patterns.

Shorter summary: Earth transforms sunlight's <u>visible</u> light energy into <u>infrared</u> light energy, which leaves Earth slowly because it is absorbed by greenhouse gases. When people produce greenhouse gases, energy leaves Earth <u>even more slowly</u>—<u>raising</u> Earth's temperature.

Appendix E

Mechanism questions and coding scheme for answers

E.1 Materials used by all coders

Development of the coding scheme was a multi-step process. Initially, two members of our group, Sarah Cohen and Roxana Farjadi, saught to identify conceptions that occurred across multiple surveys. These conceptions were assigned numerical codes, and these codes were arranged into general categories. Following this, I developed a more complete progression, describing relationships between the various categories, as well as grouping them into "misconceptions," "ignorance," and "mechanistic description." This allowed the beginnings of a scoring rubric to be developed. We then iterated the process with a larger group of coders to arrive at the final product reproduced below. What follows is the full text of the coding packet developed by Ms. Cohen and Ms. Farjadi, which also contains the text for the mechanism questions we asked in our interventions.

Note that Cohen (2012a) also reports on a coding schema, though that scheme exhibits differences with the one described here. Following this are a diagram representing relationships between the codes. A section containing a set of notes provided by Myles Crain are included in the next section. They provides a set of criteria for chosing between notes, and was used by the final set of coders. See chapter 1 for details.

Instructions

1. Responses can be classified in three categories at most. Give them as many codes as possible.
2. If the respondent talks about the differentiation of energy, refer to the "definition of differentiation of energy" table for additional help in categorizing.
3. If the respondent talks about *how greenhouse gases work*, refer to the "definition of greenhouse gases" table for additional help in categorizing.
4. If the respondent mentions greenhouse gases, refer to the "says/mentions greenhouse gases" table for additional help in categorizing.
5. If the respondent talks about *any type of mechanism for climate change*, refer to the "mechanism of climate change table." This table is broken into the sub-categories of energy, source, general chemical reactions, and respondent confusion. Please note that sometimes *a response can fit into more than one subcategory* under the overarching mechanism category.
6. If the respondent leaves a question blank, writes "do not know," or "same as above," refer to the last table, "Don't Know."
7. If the response prompts categorization ambiguities, first look at the response as a whole to look for phrases that might provide a clearer indication of what they mean. If the ambiguity can be clarified without coder inferences or assumptions, categorize the response into the code that provides the most possible credit (i.e., "be charitable within reason"). If the coder cannot clear up the ambiguity or must make assumptions, code the response into the category which best describes what the respondent actually says and not what the coder might think they are trying to say (i.e., "don't infer extra credit"). Also, note whether the respondent is defining something, explaining how climate change works, or *both*. To be doing both, the ideas must be clearly a definition and a mechanism. For instance, to say "greenhouse gases do X and thus trap heat on earth" would be both a definition and a mechanism. Even if a definition is embedded in a phrase that describes the mechanism, give them credit for both the mechanism and the definition.
8. Unless otherwise noted, all the categories listed can be applied to Know_1, Know_2, or Know_3.
9. See example column for examples of each code. Please note that for each example, the response may have been coded into more categories than just the category in which the example is placed (e.g., the example for MCCS2 was coded into SGHG1 as well as MCCS2).

<u>Definition of Terms</u>

Know_1: Please write 1-3 sentences (about 30 words or less) that you could use to explain <u>how climate change occurs</u> to a senior in high school.

Know_2: Please explain any differences regarding how energy (i.e., heat, light) travels to the Earth from the sun compared to how energy travels away from the Earth.

Know_3: Are all gases "greenhouse gases?" If not, what makes something a greenhouse gas?

Categories (Listed) – Please see tables for cutoffs, discussions, and comparisons between categories.

DD: Definition of the Differentiation of Light/Energy

DD1: Respondent differentiates between visible sunlight entering the atmosphere and infrared radiation/heat being emitted by the Earth.

DD2: Partial credit for differentiation: Respondent attempts to explain how energy differs when it enters the atmosphere and when it leaves, but does so in such a way that is either too incomplete or incorrect to fit into category DD1. Category DD2 is therefore "partial credit" for DD1. As long as the participant references some kind of asymmetry in how light is reflected, bounced, changed, etc. (even if mostly wrong), they fall in category DD2 and not DD3.

DD3: Completely incorrect attempt to differentiate kinds of light/energy – This only applies to when there is absolutely NO asymmetry referenced.

DGHG: Definition of Greenhouse Gases

DGHG1: Greenhouse Gas "right *definition*" – Respondent may or may not mention the exact phrase "greenhouse gas", but at least defines them in the right context. Respondent defines greenhouse gases as molecules that *absorb* energy, not as molecules that trap, stop, block, or reflect energy. Respondent may use the terms light, heat, radiation, or infrared radiation instead of energy in their definition.

DGHG2: Greenhouse Gas "partial credit *definition*" – Respondent may have demonstrated an understanding of some of the elements outlined in category DGHG1 but their answer is either too grammatically vague to pass judgment on correctness or contains elements of *incorrect* content ("partial credit"). To get a definition code, the respondent has to mention or allude to energy. Remember that responses in this category do *not* describe greenhouse gases as molecules that *"absorb energy."*

DGHG3: Not all gases are greenhouse gases: Respondent directly answers the question in Know_3 by stating in some way that not all gases are greenhouse gases.

DGHG4: Wrong concept of greenhouse gas: The participant holds obvious misconceptions about what a greenhouse gas is or how it works.

SGHG: Says/mentions greenhouse gases - If they give at least some definition or statement as to *what* greenhouse gases do or *how* they work, refer to DGHG categories.

SGHG1: *In know_1*: Simple mention of greenhouse gases (no explanation) –Participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* explanation of climate change.

In know_2: Simple mention of greenhouse gases - participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* explanation or strongly implied understanding of the concept of how energy functions in the atmosphere

In know_3: Simple mention of greenhouse gases - participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* explanation or strongly implied understanding of the concept of a greenhouse gas.

SGHG2 : *In know_1*: Simple mention of greenhouse gases –Respondent uses the term "greenhouse gas," or provides a specific example of one, like carbon dioxide, in the context of a mostly *incorrect* explanation of climate change.

In know_2: Simple mention of greenhouse gases - participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a mostly *incorrect* explanation or strongly implied understanding of the concept of how energy functions in the atmosphere

In know_3: Simple mention of greenhouse gases - participant uses the term “greenhouse gas,” or provides a specific example, like carbon dioxide, in the context of a mostly *incorrect* explanation or strongly implied understanding of the concept of a greenhouse gas.

SGHG3: Mentions greenhouse effect – Respondent explicitly uses the phrase “greenhouse effect,” or some variation thereof. The respondent may or may not offer an explanation of what the greenhouse effect is or how it works.

MCC: Mechanism of Climate Change, broken up by concept

MCCE: Mechanism of climate change, energy

MCCE1: Atmosphere Retention time: Respondent describes how long it takes for heat to leave the atmosphere in depth. They reference that there are “more” greenhouse gases now than there were before, which causes heat to stay in the atmosphere longer OR causes *more* heat to stay in the atmosphere (either time or amount are permissible in this category). The explanation must be in the context of comparing a previous instance when greenhouse gases existed to the presence of greenhouse gases in the atmosphere today.

MCCE2: Trapped heat as a *mechanism* for climate change: Respondent describes heat/energy/radiation as being trapped. They may describe energy changes but lack a comparison from our time to a previous time with greenhouse gases. For inclusion in this category, the respondent must use the idea of “trapping” or “stopping” heat from leaving and must NOT attempt to use the concept of energy being “trapped” as a definition of greenhouse gases– that would fall into category DGHG2. However, there are responses that may be coded as both categories MCCE2 and DGHG2 if the respondent separately defines greenhouse gases, as guided by the definition of category DGHG2, and describes the mechanism of climate change as trapping heat.

MCCE3: Input rate/amount of energy does not equal output rate/amount of energy – Respondent demonstrated some knowledge that rate/amount of energy input is different from the rate/amount of energy output, and so energy is “stuck” somewhere OR energy is “slowed down.” If the person does NOT reference a previous time with less GHGs, but does talk about heat being slowed or hindered from leaving the atmosphere, this category applies. Also, this category classifies responses that are vaguer than those in category MCCE2 or MCCE1.

MCCE4: Radiation from the sun directly heats the atmosphere – Respondent explicitly states or strongly implies that the atmosphere is heated by radiation from the sun. Respondent does not mention that Earth absorbs/reemits energy (i.e., the respondent skips differentiating energy).

MCCS: Mechanism of climate change, source

MCCS1: Human element: Respondent states or heavily implies that human emissions of greenhouse gases cause or contribute to global warming. This category includes references to fossil fuels and technology as causes of climate change.

MCCS2: Natural variation/weather patterns as an explanation for climate change: Respondent references natural variation in weather patterns as a cause of climate change thereby implying that anthropogenic emissions (“the human element”) are not the only causes of climate change.

MCCS3: Pollution: Respondent explicitly states or strongly implies that pollution causes global warming, with no explicit reference to energy’s function in the warming of the earth. This category also includes responses where the respondent seems to think that pollution physically “thickens the atmosphere” and thus causes warming. If the person references pollution (as opposed to greenhouse gases) as causing global warming, the response fits in this category.

MCCS4: Ozone: Respondent talked about the *depletion* of the ozone layer causing global warming.

MCCR: Mechanism of Climate Change, General Chemical Reactions

MCCR: Chemical Reactions and/or molecular properties explanations: participant attempts to explain the difference between energy entering Earth’s atmosphere and energy exiting Earth’s atmosphere from a strictly chemical perspective. Response does not include explicit differentiation between energies but rather uses chemical reactions in themselves as the cause of warming. A molecular perspective involving vibrations or other molecular properties may be used instead of chemical reactions or in addition to them. Response is too general to be given credit for categories DD1 or DGHG1.

MCCQ: Mechanism of Climate Change, Confused Respondent

MCCQ1: General Weather Confusion: Respondent thought we were asking about the seasons. The respondent may describe weather patterns, Earth's rotations, or the tilt of the Earth's axis.

MCCQ2: Did not understand: Respondent supplies a completely irrelevant answer (i.e. talks about high school perspectives).

DNK: Don't know or blank

DNK1: Don't know or N/A

DNK2: Code here if the participant uses a phrase similar to "I wouldn't add anything" or same as above.

Categories (organized by keyword)

Name of Category	Definition of Differentiation of Energy: DD In descending order from most thorough to least thorough	Distinctions:	Examples:
DD1	Respondent differentiates between visible sun light entering the atmosphere and infrared radiation/heat being emitted by the earth.	This category is fairly easy to find; if respondent say "reflected" IR (instead of absorbed and reemitted) that still fits here, provided that they made **some distinction** between light coming in and light going out.	"higher frequency radiation from the sun enters easily, but the lower frequency radiation reemitted by the cooler earth" (1Post) "the sun emits energy ans the earth absorbs that energy and then infared light comes back" (25Post)
DD2	Partial credit for differentiation: Respondent attempts to explain how energy differs when it enters the atmosphere and when it leaves, but does so in such a way that is either too incomplete or incorrect to fit into category DD1. Category DD2 is therefore "partial credit" for DD1. As long as the participant references some kind of asymmetry in how light is reflected, bounced, changed, etc. (even if mostly wrong), they fall in category DD2 and not DD3.	For example participant responses **may** include: -Failure to say how visible light becomes infrared -Failure to mention visible light AND infrared light (or heat) -Other **partially incorrect** attempts at differentiation	"Energy traveling to earth is converted to infared, [this energy can be absorbed by greenhouse gases]" (4Post). "The earth emits shorter wavelengths of energy whereas the sun emits longer ones." (6 Post)
DD3	Completely incorrect attempt to	Fails to understand that there is a difference in	"No difference on how energy travels."

	differentiate kinds of light/energy; this only applies to when there is absolutely NO asymmetry referenced	incoming and outgoing energy.	(27 Pre)

Name of category	Definition of Greenhouse Gas : DGHG In descending order from most thorough to least thorough	Distinctions:	Examples:
DGHG1	Greenhouse Gas "right *definition*" – Respondent may or may not mention the exact phrase "greenhouse gas", but at least defines them in the right context. Respondent defines greenhouse gases as molecules that *absorb* energy, not as molecules that trap, stop, block, or reflect energy. Respondent may use the terms light, heat, radiation, or infrared radiation instead of energy in their definition.	If you are having trouble deciding between DGHG1 and DGHG2, look at the context in which the definition of a greenhouse gas is given. Furthermore, if you really cannot tell what they are saying (because of grammar or vagueness) pick DGHG2. To be qualified in DGHG1, the respondent has to give some indication that they know *how* greenhouse work, not just that they cause something to happen, resulting in warming. (If respondent uses the concepts of trapping, stopping, blocking, or reflecting energy the response belongs in category DGHG2.) It doesn't matter for this category where the respondent thinks the energy comes from.	"Greenhouse gases absorb the reflected light…" (2Post) "Only the ones that can absorb infared light, like CO2 are considered greenhouse gases..."(3 Post)
DGHG2	Greenhouse Gas "partial credit *definition*" – Respondent may have demonstrated an understanding of some of the elements outlined in category DGHG1 but their answer is either too grammatically vague to pass judgment on correctness or contains elements of *incorrect* content ("partial credit"). To get a definition code, the respondent has to mention or allude to energy. Remember that responses in this category do *not* describe greenhouse gases as molecules that "*absorb energy.*"	Remember, this is the "Partial Credit" category. Cut-off: When respondent tries to explain the function of a greenhouse gas the response fits in DGHG2 when they do **not** say absorb.	"Climate change occurs due to the abundance of greenhouse gases in the atmosphere. Greenhouse gases, like co2, are slowly emitted into the atmosphere as energy, but as the abundance of this gas increases, it slowly warms up the earth, b/c greenhouse gases are created at a faster rate than they absorb infared light" (14 Post) "Carbon gases are released into the air that trap extra light" (16 Post)

DGHG3	Not all gases are greenhouse gases: Respondent directly answers the question in Know_3 by stating in some way that not all gases are greenhouse gases.	Just have to say "no" in some way, but do **not** have to understand why. Can also give counterexample to count in this category (e.g. saying, "N2 is not a greenhouse gas").	"No, a greenhouse gas is referring to…" (21Pre) "not all gases are greenhouse gases. No clue what makes a greenhouse gas a greenhouse gas" (24Pre)
DGHG4	Wrong concept of greenhouse gas: The participant holds obvious misconceptions about what a greenhouse gas is or how it works.	If there is some modicum of correctness do **not** put the response here. Give them the credit for what they know.	"Greenhouse gases are the gases that remain in the earth's atmosphere. They are unable to leave" (30Pre)

Name of category	Says/Mentions Greenhouse Gases: SGHG In descending order from most thorough to least thorough	Distinctions:	Examples:
SGHG1	-*In know_1*: Simple mention of greenhouse gases (no explanation) – Participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* mostly *correct* explanation of climate change. -*In know_2:* Simple mention of greenhouse gases - participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* mostly *correct* explanation or strongly implied understanding of the concept of how energy functions in the atmosphere -*In know_3*: Simple mention of greenhouse gases - participant uses the term "greenhouse gas," or provides a specific example, like carbon dioxide, in the context of a moderately or mostly *correct* mostly *correct* explanation or	If they do **not** describe the behavior of greenhouse gases, examine the context. If they mention it in a moderately or mostly correct context, then the response fits in SGHG1. Parts of the response can be wrong or irrelevant, but if they use the term greenhouse gases in a mostly correct context, SGHG1 is appropriate. This response does **not** fit into category DGHG1 because it does **not** say that greenhouse gases trap heat. Saying that GHGs cause warming does **not** give enough indication of understanding of *how* GHGs interact with energy. This response also does **not** fit into category MCCS3 because it does **not** specify that GHGs intrinsically cause warming.	"Climate change … can also be induced unnaturally by greenhouse gas buildup from carbon emissions" (13 Pre)

	strongly implied understanding of the concept of a greenhouse gas.		
SGHG2	-*In know_1*: Simple mention of greenhouse gases –Respondent uses the term “greenhouse gas,” or provides a specific example of one, like carbon dioxide, in the context of a mostly *incorrect* explanation of climate change. -*In know_2:* Simple mention of greenhouse gases - participant uses the term “greenhouse gas,” or provides a specific example, like carbon dioxide, in the context of a mostly *incorrect* explanation or strongly implied understanding of the concept of how energy functions in the atmosphere -*In know_3*: Simple mention of greenhouse gases - participant uses the term “greenhouse gas,” or provides a specific example, like carbon dioxide, in the context of a mostly *incorrect* explanation or strongly implied understanding of the concept of a greenhouse gas.	Responses fit into category SGHG2 when they mention GHGs (or a type of GHGs) but do so in a mostly **incorrect** explanation. When participants refer to ozone depletion as the main cause of global warming, for example, it is **incorrect**. Because this response does **not** explain how GHGs work, and the context is incorrect, it fits into SGHG2.	“Climate change occurs when the weather patterns abruptuly change and are abnormal. It occurs because of greenhouse gases such as carbon dioxide released into the atmosphere.” (35 Pre)
SGHG3	Mentions greenhouse effect – Respondent explicitly uses the phrase “greenhouse effect,” or some variation thereof. The respondent may or may not offer an explanation of what the greenhouse effect is or how it works.	If respondent defines GHGs correctly and then mentions the greenhouse effect separately, SGHG3 and DGHG1 can be used to categorize the same response. However, usually SGHG3 is used in place of DGHG1.	“climate change occurs due to an increase of trapped infared light in our atmosphere which is caused by the greenhouse effect.” (21 Post)

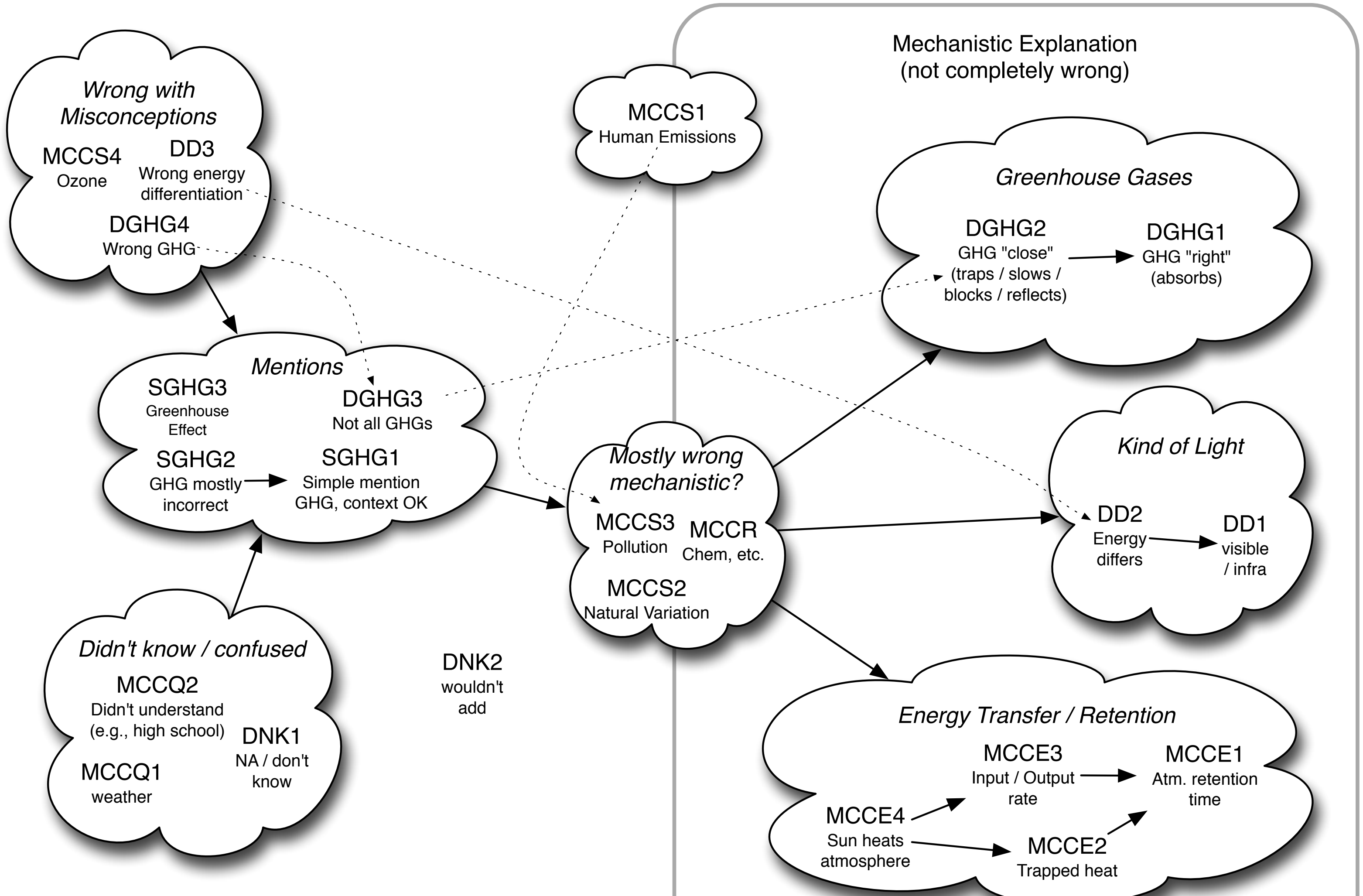
Name of category	Mechanism of Climate Change: MCC	Distinctions:	Examples:
ENERGY, Mechanism of Climate Change: MCCE			
MCCE1	Atmosphere Retention time: Respondent describes how long it takes for heat to leave the atmosphere in depth. They reference that there are "more" greenhouse gases now than there were before, which causes heat to stay in the atmosphere longer OR causes *more* heat to stay in the atmosphere (either time or amount are permissible in this category). The explanation must be in the context of comparing a previous instance when greenhouse gases existed to the presence of greenhouse gases in the atmosphere today	MCCE1 needs to have some sort of comparison to another time when there were not as many GHGs in the atmosphere. If they do not, then the response likely fits into MCCE2 or MCCE3. MCCE1 is the most specific category. Often there will be reference to "slowing" or "preventing" the escape of heat from the atmosphere	"Greenhouse gases absorb the reflected light and cause the earth to heat up (when more gases, slower rate of expulsion + therefore more heat" (2post) "but currently too much carbon gases are released into the air that trap extra light (heating earth up more than usual" (16 Post)
MCCE2	Trapped heat as a *mechanism* for climate change: Respondent describes heat/energy/radiation as being trapped. They may describe energy changes but lack a comparison from our time to a previous time with greenhouse gases. For inclusion in this category, the respondent must use the idea of "trapping" or "stopping" heat from leaving and must NOT attempt to use the concept of energy being "trapped" as a definition of greenhouse gases– that would fall into category DGHG2. However, there are responses that may be coded as both categories MCCE2 and DGHG2 if the respondent separately defines greenhouse gases, as guided by the definition of category DGHG2, and describes the mechanism of climate	This response fits into MCCE2 and **not** MCCE1 because it does **not** say that the more greenhouse gases there are in the atmosphere, the longer the energy stays in the atmosphere. Rather, it implies that there is a threshold beyond which energy "lingers" in the atmosphere. MCCE2 is almost MCCE1, but there is either a slight misunderstanding or miscommunication in the wording of the response (i.e., this category is partial credit). If energy being "trapped" is used to *define* a GHG, the response is coded in DGHG2 so as to avoid giving credit twice.	"Climate change is a gradual heating of the Earth's atmosphere due to trapped heat" (30 post) "co2. that creates a layer in our planet's atmosphere which traps sunlight and warms up the earth." (12 Pre)

	change as trapping heat.		
MCCE3	Input rate/amount of energy does not equal output rate/amount of energy – Respondent demonstrated some knowledge that rate/amount of energy input is different from the rate/amount of energy output, and so energy is “stuck” somewhere OR energy is “slowed down.” If the person does NOT reference a previous time with less GHGs, but does talk about heat being slowed or hindered from leaving the atmosphere, this category applies. Also, this category classifies responses that are vaguer than those in category MCCE2 or MCCE1.	If trying to decide between MCCE1, MCCE2, and MCCE3, first ascertain if there is a comparison to a different time with a different level of greenhouse gases. If yes, then MCCE1. Otherwise, look at the clarity: if they say heat is being STOPPED or TRAPPED, the response goes in MCCE2; if the response talks about how energy is slowed or hindered, then MCCE3.	“Cimate change is the heating up of the earth - above its normal temperature. It is caused by waves of heat leaving the earth's atmosphere, but certain greenhouse gases has caused the waves to leave even more slowly, causing the earth to be at a higher temperature.” (6 post) “it releases infared light which gets absorbed by the greenhouse gases in our atmosphere causing the earth to heat up” (15 Post) – This is a good example of both a definition and a mechanism.
MCCE4	Radiation from the sun directly heats the atmosphere – Respondent explicitly states or strongly implies that the atmosphere is heated by radiation from the sun. Respondent does not mention that Earth absorbs/reemits energy (i.e., the respondent skips differentiating energy).	If the respondent only refers to radiation **from the sun** heating greenhouse gases, then it fits in MCCE4. In other words, it will **not** fit into category DD1 because it fails to explain differentiation. Additionally, if the mechanism by which energy from the sun reaches the Earth is ambiguous and there are no clear indications in the rest of the response to suggest that the energy reaches the Earth’s surface, then the response should be classified in MCCE3.	“The atmosphere traps energy traveling from the sun.” (49 Pre)
SOURCE, Mechanism of Climate Change: MCCS			
MCCS1	Human element: Respondent states or heavily implies that human emissions of greenhouse gases cause or contribute to global warming. This category includes references to fossil fuels and technology as causes of	This category will include any reference to how humans cause climate change, e.g. the Industrial Revolution, cars, oil combustion, etc.	“Greenhouse gases emited by our cars, and industrial process and other human activity involving the burning of fossil fues or other combustables” (18 Pre)

	climate change.		
MCCS2	Natural variation/weather patterns as an explanation for climate change: Respondent references natural variation in weather patterns as a cause of climate change thereby implying that anthropogenic emissions ("the human element") are not the only causes of climate change.		"Climate change is a natural process (ice age - el nino) an can also be induced unnaturally by greenhouse gas buildup from carbon emissions" (13 Pre)
MCCS3	*MCCS3*: Pollution: Respondent explicitly states or strongly implies that pollution causes global warming, with no explicit reference to energy's function in the warming of the earth. This category also includes responses where the respondent seems to think that pollution physically "thickens the atmosphere" and thus causes warming. If the person references pollution (as opposed to greenhouse gases) as causing global warming, the response fits in this category.	This category needs some sort of implication that humans or "waste" emissions warm up the atmosphere by themselves, with no regard for energy's role.	"We produce too much carbon as waste. It ends up in the atmosphere. Heats up." (31 Pre)
MCCS4	*MCCS4*: Ozone: Respondent talked about the *depletion* of the ozone layer causing global warming.	If the respondent claims that ozone depletion causes climate change, it goes into MCCS4.	"ozone depletion also affect how the sun's heat and light is absorbed in our atmosphere and cause climate change." (28 Pre)
GENERAL CHEMICAL REACTIONS, Mechanism of Climate Change: MCCR			
MCCR	Chemical Reactions and/or molecular properties explanations: participant attempts to explain the difference between energy entering Earth's atmosphere and energy exiting Earth's atmosphere from a strictly chemical	Responses fit into this category if they provide a very general attempt to describe heat in the atmosphere. Often the respondent has misconceptions about the role of chemicals in the atmosphere and therefore their response **cannot** fit into categories DD1 or DGHG1 as	"The sun directly enters the earth causing many chemical reactions. The earths byproducts of these chemical reactions let out either heat or molecules. Some molecules reabsorb the heat and create global warming" (5

	perspective. Response does not include explicit differentiation between energies but rather uses chemical reactions in themselves as the cause of warming. A molecular perspective involving vibrations or other molecular properties may be used instead of chemical reactions or in addition to them. Response is too general to be given credit for categories DD1 or DGHG1.	well as this one.	post) "Energy travels to the earth from the sun in the rays of heat of the sun in the form on molecules in constant motion. Energy travels away from earth by the same force of interacting and fast moving molecules" (6 pre)
RESPONDENT CONFUSION, Mechanism of Climate Change: MCCQ			
MCCQ1	General Weather Confusion: Respondent thought we were asking about the seasons. The respondent may describe weather patterns, Earth's rotations, or the tilt of the Earth's axis.	Respondent could talk about seasons in conjugation with actual explanation of global warming. Read the whole response before coding.	"Climate change occurs when the sun is hitting the earth from a different angle. When it is winter, the sun's rays are less direct. In the summer, there are longer days w/ more direct sunlight" (21 Pre)
MCCQ2	Did not understand: Respondent supplies a completely irrelevant answer (i.e. talks about high school perspectives).		"It is senior year that students begin to get tired of the hgh school environment and are anxious to open a new chapter of their lives: colege. This is called senioritis. Therefore a climate change occurs to a senior in highschool when he/she is ready to leave high school and move on" (6 Pre)

Number of Category	Don't know: DNK	Distinctions:	Examples:
DNK1	N/A: maybe ran out of time.		"I do not know how climate change occurs I was never taught." (24Pre)
DNK2	Code here if the participant uses a phrase similar to "I wouldn't add anything" or "same as above."		"I wouldn't add anything." (3 Post)



Solid lines represent orderings. Dashed arrows indicate progressions across "clouds." In the absence of arrows, left to right is meant to indicate some sense of order

So, the best responders will have DGHG1, DD1 and MCCE1 codes somewhere as well as DGHG3 in know_3. MCCS1 is also great, but not as targeted by our intervention.

E.2 Notes on choosing codes

This "crib sheet" was generated by Myles Crain to identify a single defining characteristic and/or unique distinction within each code. Here are a few notes on how it was used:

- The crib sheet is NOT self-contained. Its meant to jog memory without having to constantly flip through the coding packet. The sheet is only useful if you are generally familiar with the coding scheme already.
- Assigning a code should be defensible with explicit references to the definitions and explanations of that code as provided in the packet.
- I've separated DGHG3 from the other DGHG codes intentionally (that is, DGHG3 coming after DGHG4 is NOT a typo).
- SGHG codes are only supposed to be used in the complete absence of a definition of GHGs. The SGHG category is primarily useful in coding for whether an explanation of climate change includes explicit reference to GHGs.
- Use MCCE codes to identify how a participant refers to energy within an explanation of climate change.
- Enquoted things are things that must appear in a response in order to apply the code (except when there are other options–for example, SGHG1 requires using the phrase ″GHG″ *or* citing specific examples of GHGs).

Following is the "crib sheet" itself:

DD1 visible incoming & infrared outgoing

DD2 asymmetry/difference reference

DD3 wrong, no asymmetry/difference

DGHG1 GHGs ″absorb″ energy

DGHG2 part correct, no ″absorb″

DGHG4 wrong

DGHG3 ″not all″, cite >1

SGHG1 ″GHG″/e.g., mostly accurate

SGHG2 ″GHG″/e.g., mostly wrong

SGHG3 ”greenhouse effect”

MCCE1 more gas/heat than before

MCCE2 heat/energy ”trapped”

MCCE3 different input/output rates/amounts

MCCE4 sun’s radiation heats atmosphere

MCCS1 humans/tech/fossil fuels

MCCS2 natural variation

MCCS3 pollution

MCCS4 O3 layer

MCCR chemical/molecular exclusively

MCCQ1 weather, confusion

MCCQ2 irrelevant

DNK1 ”don’t know”, n/a

DNK2 nothing added

E.3 Assigning scores to coded responses

Detail here the procedure for assigning scores to codes. Perhaps even include the R or python code?

Appendix F

Using imputation to combine participants with and without a pre-test

Imputation is a well-established approach to dealing with missing data (ref - probably Keppel and Wickens). In a number of the experiments in this dissertation, multiple groups received a similar intervention, but one group may have been missing a pre-test where we obtained their naïve baseline score (e.g., for a climate-relevant attitude). The approach we used in these cases was to use the participants for which we *did* have a pre-test score (i.e., our sandwich group), and use the average of those as an approximation to our other groups pre-test score. To be explicit, following is the exact R code used to compute this test for Study 1 in Chapter 1:

```
# Here, dfs is pre-populated with the measured values. We assign the mean of
# the sandwich (s) group scores to the pre-test scores for the no-pretest
# group (n). We then append the sandwich group scores unmodified.

imputed.df <- data.frame(pre.gw=mean(dfs$s.pre$total.gw),
                         total.gw=dfs$n.post$total.gw)

imputed.df <- rbind(imputed.df,
                    dfs$s.post[,c('pre.gw', 'total.gw')])

# Note - this gives the same result as a simple t-test on the difference
# scores, so we're not cheating on our degrees of freedom, or obtaining
# artificially lower variance on the pre-test scores.

with(imputed.df,
     t.test(pre.gw, total.gw, alternative='less', paired=TRUE) )
```