Climate change by the numbers:

Leveraging mathematical skills for science learning online

Ian Thacker

University of Texas, San Antonio

ian.thacker@utsa.edu

Author Note

Ian Thacker https://orcid.org/0000-0002-2492-2929

Correspondence concerning this article should be addressed to Ian Thacker, Department of Educational Psychology, University of Texas at San Antonio, San Antonio, TX 78207. Email: ian.thacker@utsa.edu.

I have registered this study with the University of Texas at San Antonio Institutional Review Board (FY20-21-70).

I have no conflicts of interests to disclose.

Color is preferred to be used online only, not in print.

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI: https://doi.org/10.1016/j.learninstruc.2023.101782

Abstract.

The purpose of this preregistered study was to test an online intervention that presents participants with novel numbers about climate change after they estimate those numbers. An experimental study design was used to investigate the impact of the intervention on undergraduate students' climate change understanding and perceptions that human caused climate change is plausible. Findings revealed that posttest climate change knowledge and plausibility perceptions were higher among those randomly assigned to use the intervention compared with those assigned to a control condition, and that supplementing this experience with numeracy instruction was linked with the use of more explicit estimation strategies and greater learning gains for people with adaptive epistemic dispositions. Findings from this study replicate and extend prior research, support the idea that novel data can support knowledge revision, identify estimation strategies used in this context, and offer an open-source online intervention for sharing surprising data with students and teachers.

Keywords: climate change, conceptual change, epistemic dispositions, numerical estimation, plausibility judgments, learning technology

Climate change by the numbers:

Leveraging mathematical skills for science learning online

There is a gap between scientific knowledge and the public understanding of science. Over 98% of climate scientists agree that humans are causing climate change, while only 57% of adults among the general public in the United States concur (Leiserowitz, 2022). Further, 43% of adults in the USA incorrectly believe that there is no scientific consensus around climate change (Leiserowitz, 2022). Misconceptions about socio-scientific topics of this sort can lead the public to distrust scientific viewpoints (Skogstad, 2003; Sinatra & Hofer, 2022) and can influence their policy preferences (Leiserowitz, 2006; Ranney & Clark, 2016).

Correcting such misconceptions is not a simple matter of presenting people with better knowledge of the facts. Shifting misconceptions about climate change can be challenging because of several psychological barriers that make them resistant to change. For example, people sometimes resist engaging with climate change information to avert facing responsibility for their climate inaction or to avoid ostracization from social or political groups (Stoknes, 2015). As such, efforts to support people in correcting their misconceptions need to consider multiple characteristics of the learner (Muis et al., 2015) and support learners with skills required to critically evaluate evidence and claims.

Fortunately, there are approaches that have the potential to bolster critical evaluation skills and shift misconceptions, such as micro-interventions that present people with surprising numbers about climate change after they estimate those numbers (Ranney & Clark, 2016; Thacker & Sinatra, 2022). For example, I encourage the reader to take a moment to estimate the following quantity: What is the percent change in atmospheric levels of methane (a greenhouse gas) since 1750 until now? Savvy readers might take some time to reflect on this question, draw from their prior knowledge about climate change, and roughly estimate a percentage value that they find mathematically and practically reasonable before checking the footnote[1] to compare their estimate with the scientifically accepted value. Theory predicts that explicitly considering evidence in this way (e.g., by formulating a numerical estimate before assessing the scientific value) can activate prior knowledge (Zhang & Fiorella, 2023), elicit more explicit reflection on novel evidence, deeper integration of supported claims (Richter & Maier, 2017), and can subsequently shift people's misconceptions to be more aligned with the scientific consensus (Lombardi et al., 2016). Furthermore, theory and evidence suggest that interventions that provide strategies for explicit reflection on discrepant information can incite greater knowledge revision, and that such effects are moderated by people's dispositions toward reasoning (i.e., epistemic dispositions; Lombardi et al., 2016; Thacker & Sinatra, 2022). For example, Thacker & Sinatra (2022) found that undergraduate students who were presented with surprising numbers about climate change after estimating those numbers had significantly fewer climate change misconceptions compared to a control group, and that supplementing this intervention with numerical estimation instruction and reflective prompts led to larger learning gains among students who reported a willingness to reason with new evidence. Yet, despite these promising findings, the interventions created for prior studies are not easily accessible to the general public and the strategies that individuals employ when estimating quantities are not well understood.

The purpose of this study was to replicate and extend upon a recent experimental study conducted by Thacker & Sinatra (2022). To do so, I made use of a revised version of their intervention: an online, game-based intervention that presents people with novel numbers about climate change after they estimate those numbers. Key revisions to the intervention included (a)

[1] The change in the atmospheric levels of methane since 1750 until now is a 151% increase (NASA, n.d.).

sharing game-like accuracy feedback to enhance learner engagement and encourage reflection on belief-discrepancies, (b) presenting sources of the information to improve its credibility, (c) adding text to explain, contextualize, and improve plausibility of the novel climate change numbers, and (d) providing a "summary" page to encourage further reflection on and integration of the information presented to participants. I also included a variant of the intervention that supplemented the experience with instruction on a targeted collection of numerical estimation strategies: use of given benchmarks, tolerance for error, and digit rounding. Using these interventions, I tested whether the same patterns of learning would emerge with a national sample of undergraduate students in an experimental study. Such replications that purposefully vary elements of context, sample, or operationalization of variables are also called "conceptual replications" (Plucker & Makel., 2021). Furthermore, I preregistered planned analyses and hypotheses in advance of data collection (see https://bit.ly/3rP2m9c for all preregistration information), a practice encouraged to promote transparency in research (Reich, 2021). In addition to replication, I also identified estimation strategies that participants reported using during the study, and assessed whether supplemental instruction on numerical estimation strategies influenced those strategies.

Theoretical Framework

To frame how individuals learn from novel quantitative information about climate change, this study synthesizes theory on Conceptual Change, Plausibility Judgments, and the Integrated Theory of Numerical Development (Dole & Sinatra, 1998; Lombardi et al., 2016; Siegler, 2016).

Conceptual Change

When people encounter novel information that conflicts with their existing conceptions, there is potential for conceptual change. Conceptual change is when conceptual knowledge is restructured after an individual encounters incongruent information that conflicts with their existing perspectives (Dole & Sinatra, 1998; Murphy & Mason, 2006). There are many definitions and operationalizations of conceptual change (for a review, see Vosniadou, 2013). In this study I consider conceptual change to be the shifting of conceptual knowledge to be more aligned with scientific conceptions and less aligned with misconceptions (Dole & Sinatra, 1998). This process is contingent on aspects of the information and characteristics of the learner, such as their prior knowledge, motivation, emotion, and attitudes (Dole & Sinatra, 1998; Pintrich et al., 1993; Sinatra, 2005; Sinatra & Seyranian, 2016).

The Cognitive Reconstruction of Knowledge Model (CRKM) of conceptual change discerns between characteristics of the learner and characteristics of the information that learners engage with (i.e., the "message," Dole & Sinatra, 1998), both of which predict engagement and potential for conceptual change. Learner characteristics include several factors, including the strength and coherence of a learner's existing conception, their commitment to it, and their motivation to learn. Characteristics of the learning materials are also important; an individual will only learn from information if it is comprehensible, coherent, compelling, and if the claims they support seem plausible. Thus, when a learner is presented with a novel climate change number, there is potential for conceptual change only if that information is *comprehensible* to the learner in that they can process the quantitative information and find its meaning relevant and *coherent* with their prior knowledge. To elaborate on processes involved in considering and integrating novel information, I turn to theory on plausibility judgments.

Plausibility Judgments for Conceptual Change

According to the plausibility Judgments for Conceptual Change model (PJCC; Lombardi et al., 2016), when people are presented with novel information (such as surprising climate change data), they initially process the information for validity, judge the plausibility of claims supported by the information, and then potentially restructure their knowledge and change their misconceptions as a result. The PJCC posits that when people initially *pre-process* information, their perceptions of source validity depend on several aspects of the information such as its complexity, corroboration with prior knowledge, perceived conjecture, and perceptions of source credibility (also see Connell & Keane, 2006; Rescher, 1976).

After people have pre-processed novel information for validity, they then judge the *plausibility* of the claims associated with this information. Plausibility is the tentative perception of the potential truth of a claim (Lombardi et al., 2016). For example, an individual who pre-processes and accepts that 99% of scientists believe that climate change is happening may then reflect on their feelings of plausibility and personal endorsements regarding the implied claim that climate change is real. Plausibility judgments can be either implicit or explicit, with more explicit processing depending on the individuals' level of motivation and engagement, emotion, and personal dispositions around reasoning with new information (Lombardi et al., 2016; Richter & Maier, 2017). More explicit plausibility judgments increase the likelihood that individuals will reappraise the plausibility of scientific claims and increases the potential for conceptual change. For instance, when a person estimates a number, they may draw from their prior knowledge and existing assumptions about climate change and apply explicit numerical estimation strategies to arrive at their estimate, such as by using and mathematically manipulating known values to help estimate unknown values. Such an explicitly crafted estimate may better prepare the learner to interpret a scientifically accepted value when later presented with it. If there are notable discrepancies between their estimate and the true value, such a confrontation may prompt the learner to re-appraise the plausibility of the claims supported by that information (c.f., Lombardi et al., 2016; Richter & Maier, 2017).

Empirical evidence demonstrates that there are strong relationships between plausibility perceptions and conceptual change outcomes regarding climate change. For example, Lombardi and Sinatra (2012) found that students who completed a science course devoted to the topic of climate change had fewer misconceptions and found human induced climate change to be more plausible compared with students in a typical intro-science course. In a related study, Lombardi, Danielson, and Young (2016) found that undergraduate students who were prompted to reflect on their own climate change misconceptions rated scientific explanations of climate change as more plausible and had fewer misconceptions at posttest compared with students who read an expository text.

Potential Moderators of Plausibility Judgments and Conceptual Change

As noted, the PJCC predicts that emotional, motivational, and epistemic dispositional factors are associated with plausibility perceptions and conceptual change outcomes. This study narrows in on two potential moderating factors that may be linked to explicit evaluations of evidence and claims: epistemic dispositions and mathematics anxiety.

Epistemic Dispositions. Epistemic dispositions can be defined as a learners' relatively stable views about knowledge and knowing. For example, a person may have a flexible view of knowledge and may be generally open to reason with new evidence, even if that evidence is inconsistent with their existing beliefs (e.g., Stanovich & Toplak, 2019; Stanovich & West, 1997). Such an epistemic disposition, also called Actively Open-Minded Thinking (AOT), is expected to shape whether people choose to explicitly evaluate novel information and integrate related claims into their existing knowledge structures (Lombardi et al., 2016). Prior research shows that undergraduate students with higher levels of Actively Open-Minded Thinking benefited more from estimation instruction and reflection prompts when learning from novel climate change data (Thacker & Sinatra, 2022). The current study aimed to replicate these findings using a game-based online intervention. Namely, I anticipated that learners' Actively Open-Minded Thinking would moderate the effects of the interventions created for this study.

Mathematics Anxiety. Emotional and personal dispositions, such as trait-level mathematics anxiety, are also expected to shape whether individuals might explicitly engage with novel evidence (Lombardi et al., 2016; Richter & Maier, 2017; Ramirez et al., 2018). Mathematics anxiety can be defined as a relatively enduring disposition that is characterized by feelings of fear and anxiety in response to doing mathematics or considering the prospect of doing mathematics driven partly by a fear of failure (e.g., Ramirez et al., 2018). Despite the many emotional constructs involved in knowledge revision (Muis et al., 2018), mathematics anxiety in particular is linked to university students' negative responses toward learning mathematics (Jackson & Leffingwell, 1999) and may therefore moderate the effects of a mathematical instruction intervention. As such, I expected that undergraduate students with higher levels of mathematics anxiety would be less responsive to an intervention supplemented with explicit instruction on numerical estimation skills. For the current study, I anticipated that individuals with higher math anxiety would benefit less from math instruction targeting numerical estimation strategies.

Magnitude Knowledge

Given that novel information can incite conceptual change, how might things change when that information is quantitative? Are there particular challenges people encounter when evaluating and interpreting quantitative evidence? Theory on magnitude knowledge and numerical estimation suggests that, indeed, some people may require support in making sense of numbers they are presented with, and that emphasizing specific quantitative reasoning skills may facilitate sense-making processes.

Making meaning of number magnitudes is considered to be a core competency in mathematics and science and involves several skills that develop over time (Booth & Siegler, 2006; Cheuk, 2012; Opfer & Siegler, 2007; Sasanguie et al., 2012; Siegler & Booth, 2004). Siegler's (2016) *Integrated Theory of Numerical Development* provides an explanation for how this development occurs, positing that people develop accurate understandings of number magnitudes and their relationships as they connect numbers (e.g., representing rising temperatures) to the things that those numbers refer to (e.g., global climate change). As learners develop, they learn new ways to make meaning of numbers by connecting and comparing them to other numbers, ideas, and representations through processes of *association* and *analogy*, both of which are activities that are considered crucial for both mathematical and scientific learning (Siegler, 2016). One useful context where people tend to apply processes of association and analogy is in situations where they must estimate numbers.

Numerical Estimation

Numerical estimation is an educated guess for a quantity that can draw from a person's prior experiences and understanding of number and operations (Dowker, 2005). Of the common categories of numerical estimation skills (e.g., computational estimation or numerosity; Reys & Reys, 2004), research on measurement estimation is the most relevant for this study. Measurement estimation concerns the explicit estimation of real-world measures (Sowder &

Wheeler, 1989) and is useful for understanding factors that help people judge whether real-world quantities are reasonable and valid.

Siegler and Booth (2005) summarize research supporting that people use a variety of strategies and representations when solving measurement estimation problems. Oftentimes people use multiple strategies when estimating a single number, and the strategies people use are numerous, ranging from: drawing from prior knowledge, rounding digits, using visual representations, making use of given or known information, or flat-out guessing (c.f., Joram, 1998). Siegler & Booth (2005) also posit that people choose strategies adaptively depending on the situation and whether they believe the strategies will lead to more accurate and rapid estimates, as informed by accuracy feedback. People also change estimation strategies when presented with new and advanced strategies for estimating measures, potentially via formal instruction.

Prior studies have established that a particularly important measurement estimation strategy is the *benchmark strategy*—the use of a given or known number to better estimate an unknown number. Prior lab-based research conducted with undergraduate students shows that exposure to benchmark values can improve their estimation accuracy in various contexts. For example, providing students with the distance between two cities before having them estimate the distance between new pairs of cities led students to make more accurate estimates compared with a control (Brown & Siegler, 2001). Similarly, providing benchmark quantities enhanced undergraduates' estimates of a wider variety of every-day quantities (Brown & Siegler, 1993; 1996; Friedman & Brown, 2000; Joram et al., 2005). Further, instruction on when and how to use given benchmarks can improve the frequency that students use benchmark values as well as their estimation accuracy (Joram et al., 2005) and can benefit climate science learning among students with adaptive epistemic dispositions (Thacker & Sinatra, 2022).

Further, empirical findings suggest that people's estimation accuracy and judgments of reasonableness of quantities is associated with a *tolerance for error*—a tolerance for imprecision when estimating numbers, and valuing "reasonable" answers over "correct" answers. Individual interviews with university students, adults, and K-12 students suggest that a tolerance for error is a common characteristic among good estimators (Thacker et al., 2021; Reys et al., 1982; Reys & Reys, 2004; Shimizu & Ishida, 1994). Numerical estimation, by nature, involves some amount of imprecision and uncertainty, and maintaining tolerance for that imprecision is important for learners to assess and learn from their own performance. Moreover, research on self-assessment suggests that people learn more from their errors when their *internal feedback* and self-talk is focused on mastery and correcting deeper understandings rather than when ruminating on failures and worrying about surface level inaccuracies (Zhang & Fiorella, 2023).

Another common and useful strategy used by good estimators is the use of flexible techniques for *rounding* digits to ease computations, reduce estimation time, and improve accuracy (Joram et al., 1998; Reyes & Reyes, 2004; Siegler & Booth, 2005). However, despite the correlational research showing that good estimators tend to use rounding techniques and have a tolerance for error, there are little to no experimental studies testing whether instruction around these techniques can increase frequency of their use nor whether they support learning outcomes. In the current study, I addressed this gap in the literature by experimentally testing an intervention that combined three key estimation strategies—the use of benchmark values, flexible rounding, and tolerance for error—and assessed whether the intervention increased participants' reported use of these strategies and conceptual change outcomes.

Existing Interventions

There is some existing empirical research that studies how people learn from novel climate change numbers after estimating the magnitude of those numbers (Thacker et al., 2021; Ranney et al., 2019; Ranney & Clark, 2016; Thacker & Sinatra, 2022). Ranney & Clark (2016) created an intervention that presents people with surprising climate change data after they estimate the magnitude of that data and found that undergraduate students who engaged with this intervention changed their policy preferences to be more aligned with scientific findings as compared with a control group.

Thacker & Sinatra (2022) extended this intervention by testing differences between multiple variations of the original. Namely, they tested differences between five conditions, (1) an estimation intervention adapted from Ranney & Clark (2016), (2) the estimation intervention supplemented with numerical estimation instruction focused on supporting the benchmark strategy, (3) the estimation intervention supplemented with prompts to activate reflection on inaccuracies of their estimate, (4) the estimation intervention supplemented with both estimation instruction and prompts to activate reflection, and (5) a control group that read an expository text. The authors found that undergraduate students assigned to any of the four estimation conditions had fewer misconceptions about the scientific consensus by .33 standard deviations compared with a control group. Furthermore, such gains were stronger when the intervention was supplemented with instruction on benchmark estimation strategies and reflective prompts among students with higher levels of Active Open-Minded Thinking (Thacker & Sinatra, 2022). The current study aimed to investigate whether these findings would replicate in a somewhat simpler study design that included only three experimental conditions (estimation intervention, estimation intervention + numeracy instruction, control) rather than five, as to focus more specifically on supporting productive estimation strategies and studying their impact on student learning in this context. The numeracy instruction, in particular, was expanded in this study to support a broader range of strategies, including the use of benchmark values (as previously done in Thacker & Sinatra, 2022), as well as two additional estimation strategies: tolerance for error and digit rounding techniques.

Namely, our team designed an open-source, online web-application over the course of 22 interviews with graduate and undergraduate students at a university in the southern United States (Thacker et al., 2021). Through three iterations of interviews, we identified specific mathematical skills that students used as they estimated and interpreted novel climate change data and redesigned the intervention to make these skills salient for participants. The study revealed that participants with above-median learning gains tended to make use of three specific strategies in tandem: applying arithmetic operations to given benchmark values, flexible decimal rounding, and expressing a tolerance for error. The design iterations resulted in an openly accessible, game-based learning intervention with the option to supplement the experience with estimation instruction focused on these three skills. The design was also modified from Thacker & Sinatra's intervention to present participants with: feedback on the accuracy of their estimates that call attention to belief-discrepancies, short explanations to contextualize and improve coherency of the "true values" that they were estimating, links to sources that back up the evidence to improve its credibility, and a summary page to encourage further reflection. That is, we added design features that were not previously included in the original design that might facilitate students' explicit consideration of potentially belief-discrepant numbers about climate change. However, despite the potential utility of this game-based intervention for climate change learning,[2] little is known about its effect on student climate change learning outside of the small, descriptive sample. Furthermore, little is known about what skills or knowledge participants specifically gain from the mathematical instruction provided in this setting as compared with the baseline condition.

Research Questions and Hypotheses

The current study aimed to test whether the findings of Thacker & Sinatra (2022) would replicate when using a game-based estimation intervention that was modified to (a) emphasize *three* estimation strategies (the benchmark strategy, flexible rounding, and tolerance for error) rather than the one included in the original study (the benchmark strategy), (b) be presented in an accessible online format, (c) include estimation accuracy feedback to help learners detect and correct their misunderstandings, and (d) include an explanation of and sources for each number estimated. In addition to assessing the impacts of the revised intervention, this study extends prior research by also testing whether the supplemental instruction on numerical estimation influenced the estimation strategies that participants reported using during the study. This study also aimed to increase the credibility of the research findings by documenting the research questions, hypotheses, and planned analyses in advance to conducting the study (for the full anonymous preregistration, see https://bit.ly/3rP2m9c).[3] Namely, I sought to answer the following preregistered research questions:

- **Research Question 1 (RQ1).** To what extent would estimation of and exposure to novel climate change data using an online learning intervention improve learners' climate change knowledge and plausibility perceptions compared with reading an expository text?

[2] Hereafter also referred to as the "estimation game," "game-based intervention," or "the intervention."
[3] Note that the preregistration describes an additional study that is beyond the scope of this study.

Regarding RQ1, I hypothesized (H1) that people assigned to the intervention conditions would have greater knowledge at posttest when compared with people assigned to read an expository text. I also anticipated there would be no significant difference in climate change plausibility perceptions between the intervention groups and the comparison group at posttest.

- **Research Question 2 (RQ2).** To what extent do mathematics anxiety and epistemic dispositions moderate the effects of the interventions on knowledge and plausibility?

Regarding RQ2, I hypothesized that (H2.1) individuals with high math anxiety would benefit less from the mathematics skills instruction. That is, I expected that math anxiety would negatively moderate the effects of the intervention modified with estimation instruction when the outcome was climate change knowledge or plausibility perceptions. I also hypothesized that (H2.2) individuals with higher levels of Actively Open-Minded Thinking would benefit more from the intervention and modified intervention compared with the comparison condition when the outcome was climate change knowledge. That is, I expected that Actively Open-Minded Thinking would positively moderate the effects of the intervention. I also anticipated that there would be a significant main effect of Actively Open-Minded Thinking on plausibility perceptions, but no significant interactions with the experimental conditions.

- **Research Question 3 (RQ3).** To what extent does enhancing the intervention with instruction on estimation strategies change learners' climate change knowledge and plausibility?

Regarding RQ3, I hypothesized that (H3) supplementing the intervention with instruction on estimation skills would lead learners to report more scientific knowledge about climate change compared with those who are assigned to the intervention but without estimation instruction. I expected no significant differences in climate change plausibility between these conditions.

- **Research Question 4 (RQ4).** Will individuals who are assigned to interact with the intervention supplemented with instruction on estimation strategies report using the emphasized estimation strategies (i.e., the benchmark strategy, flexible rounding, and tolerance for error)?

Regarding RQ4, I hypothesized that (H4) individuals who were assigned to the intervention supplemented with estimation instruction would more frequently report (a) tolerance for error, (b) rounding techniques, and (c) benchmark strategies compared with those assigned to the intervention without this modification. In addition to testing this preregistered research question, I also explored additional strategies that learners reported and whether there were differences in groups assigned to receive estimation instruction. In addition to testing RQ4, I also explored additional strategies that learners reported and whether there were differences in groups assigned to receive estimation instruction.

Methods

Participants and Procedure

To test these hypotheses, I formed a national online Qualtrics panel of $N = 605$ undergraduate students to participate in an experimental online survey. The sample size obtained was above the target of 500, which was a rounded estimate based on a power analysis using G*Power (Faul et al., 2009) for a regression with six predictors, power = .8, significance = .05, and an effect size of $f^2 = 0.03$ based on previous research (Thacker & Sinatra, 2022).

To obtain this sample, Qualtrics representatives initially used multiple platforms to widely share a survey link online; 2,651 people initially clicked on the link to participate, but

2,046 were dropped from the analysis because they either did not meet the eligibility criteria (over 18 and a full-time undergraduate student), did not pass an attention check, or because they were flagged as a "speeder" by the algorithm created by Qualtrics. There was no missing data at pretest or posttest.

Participants in the main analytic sample were 20.3 years old on average, 79.3% Female, 16.7% Male, 3.5% nonbinary/other, 47.6% White/Caucasian, 16.0% Black/African American, 14.5% Asian American, 4.0% Two or More Races, and 1.5% American Indian or Alaskan Native. All participants (a) completed a pretest to measure their misconceptions about climate change, plausibility judgments about climate change, mathematics anxiety, and prior epistemic dispositions, (b) were automatically directed to a web app that randomly assigned them to one of three conditions (a control group that read an expository text about the greenhouse effect, the estimation game intervention, or the modified estimation game supplemented with estimation instruction), and then (c) were automatically directed back to the original survey where they completed an estimation strategy report (intervention conditions only), identical posttest of knowledge, plausibility perceptions, and a demographics questionnaire. For a summary of the procedures, see Figure 1.

[INSERT FIGURE 1 AROUND HERE]

Materials

All survey materials, intervention texts, data, and analysis scripts are available on the Open Science Framework (https://osf.io/nqbyr/?view_only=bc36d59f82424c4f8e67cfb1538afaaf).

Conditions

There were three experimental conditions: the estimation game intervention group, modified intervention group, and control group (see Appendix A for a summary with screenshots, excerpts, and a link to the intervention). Students assigned to the estimation game intervention estimated 12 climate change related numbers before being presented with the scientifically accepted answer. As with the original study, six of these items were from Ranney & Clark (2016) and mostly prompted participants to estimate unitless percentages; the remaining six items prompted participants to estimate raw quantities (e.g., in units of meters, degrees F, or Billions of tons of CO_2) and presented participants with a benchmark value (Thacker & Sinatra, 2022). When participants were presented with the scientifically accepted answer, the intervention also displayed a short explanations justifying each accepted value, references with links to credible sources of information, and accuracy feedback displayed as images of stars representing different accuracy levels.[4] After estimating all 12 climate change numbers in this way, the game presented participants with a "final score" showing a summary of each of their estimates with the original questions, the true values, the stars they earned for each, and links to information sources. Students in the modified intervention condition also completed the same estimation game, but prior to the game, they engaged with instructional text that emphasized three numerical estimation strategies—tolerance for error, the benchmark strategy, and flexible rounding—with worked examples and two checks for understanding. These three estimation strategies were emphasized because, in prior qualitative research (Thacker et al., 2021), they were used among university students with above-median learning gains for this specific task. Students in the control group were presented with an 812 word expository text about the greenhouse effect (from Lombardi et al., 2013) to take approximately the same amount of time as the intervention conditions. In addition to describing how the greenhouse effect works, the

[4] Scoring was as follows: Five stars were displayed for estimates within 10% of the true value, four if within 20%, three if within 30%, two if more than 30%, and one star if the estimate was in the wrong direction (e.g., they estimated that atmospheric CO_2 levels have *decreased* since 1959).

control text discussed several topics that were also covered in the estimation game conditions. Namely, both noted that greenhouse gasses (such as CO_2 and methane) are responsible for trapping heat in the atmosphere, that humans are responsible for adding greenhouse gasses to the atmosphere, that this causes temperatures to rise and global ice cover to melt, and that 97% of climate scientists agree that climate change is happening. All experimental conditions were presented in an openly accessible, open-source online web app (ianthacker.com/design.html; also see Thacker et al., 2021].

Dependent Variables

Climate Change Knowledge. Knowledge of climate change was measured using seven items adapted from the Human Induced Climate Change Knowledge measure (HICCK; Lombardi et al., 2013), as done in prior research (Thacker & Sinatra, 2022). Participants responded as to whether they believed that climate scientists would believe that certain statements are true (e.g., “Most of the world’s glaciers are decreasing in size. This is evidence of climate change”). Participants reported the extent they thought climate scientists would agree with seven statements on a seven-point agreement scale ranging from 1 (*completely disagree*) to 7 (*completely agree*). As such, agreement on this scale represents accurate conceptions about the scientific consensus around climate change and disagreement represents misconceptions. Participants completed this scale pre-intervention and post-intervention ($\alpha_{pre} = .65$, $\alpha_{post} = .74$).

Plausibility Perceptions. Perceptions of plausibility that humans are responsible for climate change were measured using four items adapted from the Plausibility Perceptions Measure (PPM; Lombardi et al., 2013). These four items from the original eight-item scale were used to shorten the length of the survey and were selected because they had the highest factor loadings from Thacker & Sinatra (2022; all factor loadings > .84). These items were intended to capture participants' personal positions on whether humans are responsible for climate change as they responded to statements ("Evidence from around the world shows that the climate is changing in many regions") on a six-point agreement scale from 1 = *Highly Implausible (or even impossible)* to 6 = *Highly Plausible*. This scale was completed at pretest and posttest ($\alpha_{pre} = .81$, $\alpha_{post} = .85$).

Estimation Strategy Reports. Participants assigned to the two intervention conditions also provided open-ended descriptions of strategies that they used to estimate numbers. Participants provided immediate retrospective strategy reports at the conclusion of the estimation game by explaining their responses in a text box in response to the prompt, "Please describe in as much detail as possible how you made your estimates of climate change numbers." (also see Sidney et al., 2019; Siegler & Thompson, 2014). Strategy reports can reflect the impact of experimental conditions on the strategies they support by allowing insight into the specific methods people used to approach problems and whether they applied the techniques emphasized in experimental conditions. In sum, I used strategy reports as evidence that participants' answers were indicative of the effectiveness of the direct instruction on estimation strategies.

The coding scheme initially included several codes that emerged from a prior qualitative investigation exploring estimation strategies individuals employed during think aloud interviews as they engaged earlier versions of the estimation game (Thacker et al., 2021). These codes were adapted for use in the current study, with each code reflecting a different strategy or approach that students reported using to estimate climate change numbers (see Appendix B for the full list of codes and descriptions). Given that numerical estimation often involves the use of multiple strategies in-tandem (Siegler & Booth, 2005), we designed the coding scheme to potentially capture multiple estimation strategies by allowing for codes that were not mutually exclusive.

Three particular strategies from the codebook were assessed across conditions in this study: tolerance for error, rounding techniques, and benchmark strategies. Tolerance for error was reflected when participants reacted positively to accuracy feedback despite estimates that were not perfect, (e.g., "I knew it wasn't going to be 100% accurate," also see, Thacker et al., 2021; Shimizu & Ishida, 1994). Flexible rounding techniques were reflected when the individual reported rounding numbers to make mental computation easier (e.g., by rounding the hints to whole numbers before making estimates; e.g., Joram et al., 1998). Benchmark strategies were reflected when participants reported using given numbers to help them estimate unknown quantities (Brown & Siegler, 2001; Joram et al., 1998, Siegler, 2016). In this sample, students reported using benchmark strategies in multiple ways, ranging from a vague use of benchmarks (e.g., "I used the hints"), to the use of benchmarks to project trends over time ("I used the bit of evidence they gave and extrapolated from there"), or the use of arithmetic operations to rescale or iterate benchmarks ("I went off the amount of years and multipled [*sic*]"). For additional codes identified in the data with examples, see Appendix B.

A graduate student coder and I first collectively coded 50 responses as to calibrate around the codebook. Then we independently coded all remaining responses; interrater reliability was high (> 95% interrater agreement across all codes and responses). We then resolved all disagreements through a conversation. Both coders were blind to condition and all other student data as they coded strategy reports. The graduate student coder was not privy to the full experimental design or hypotheses.

Covariates

Epistemic Dispositions. Baseline epistemic dispositions were measured using the Actively Open-Minded Thinking scale (AOT; Stanovich & West, 1997) that captures participants' willingness to reason with novel evidence using seven items (e.g., "People should take into consideration evidence that goes against their beliefs") using a seven-point agreement scale ($\alpha = .71$).

Math Anxiety. Participants also completed a Mathematics Anxiety Questionnaire (Ganley et al., 2019) consisting of nine items (e.g., "I get a sinking feeling when I think of trying to solve math problems") with five response options ranging from 1 (Not true of me at all) to 5 (Very true of me; $\alpha = .93$).

Analytic Strategy

To assess the effects of the interventions on the knowledge and plausibility outcome variables (RQ1 & RQ3), I used ordinary least squares regression with heteroskedasticity-robust standard errors using a separate model for knowledge and plausibility perceptions. Predictors were the experimental conditions and pre-test scores. To assess moderating effects (RQ2), I repeated these analyses after adding math anxiety and Actively Open-Minded Thinking as moderators of the treatment condition, with separate models for each moderator. I examined differences between all experimental groups, and to parallel analyses from Thacker & Sinatra (2022), I also ran these models using complex contrasts testing whether there were knowledge or plausibility differences between (a) the control and a combination of the intervention two conditions, and (b) the two intervention groups, both after adjusting for pretest scores. All continuous variables were standardized around the mean prior to analyses. Analyses were performed using R Version 3.6.1.

To test whether estimation instruction led to more frequent reports of tolerance for error, rounding, and benchmark strategy techniques (RQ4), I tallied frequency counts for each estimation strategy that students used within each condition and compared frequencies between those assigned to the estimation game intervention and those assigned to the estimation game modified with estimation instruction using two-sample difference in proportions tests.

Results

Preliminary Analyses

Preliminary analyses revealed no significant preexisting differences between conditions for all main outcomes and predictors. Namely, there were no significant between-condition differences in pretest knowledge ($F = 0.14$, $p = .874$), pretest plausibility perceptions ($F = 0.12$, $p = .890$), epistemic dispositions ($F = 0.53$, $p = .587$), mathematics anxiety ($F = 0.35$, $p = .702$), age ($F = 0.02$, $p = .979$), gender ($\chi 2 = 10.2$, $p = .117$), or race ($\chi 2 = 4.6$, $p = .991$). Skew ranged from –.50 to .40 and kurtosis from –1.31 to 0.10 for all continuous variables, which is considered acceptable (Tabachnick & Fidell, 2013). Descriptive statistics by condition and intercorrelations among all main outcomes and covariates are presented in Table 1. As with the original study, I found significant improvements in knowledge ($d = 0.17$, $p < .001$) from pretest to posttest. Further, I confirmed that there were no significant interactions between condition and pretest knowledge ($p = .349$) or condition and pretest plausibility ($p = .094$), suggesting that the data met assumptions for the planned analyses (Murnane & Willett, 2010).

[INSERT TABLE 1 AROUND HERE]

Main Analyses

RQ1: Main Effects of the Intervention. Table 2 displays the full results for all multiple regression analyses. With regards to posttest knowledge as the main outcome, participants in both intervention conditions outperformed the control group. This difference was significant for both intervention variants before and after including moderating variables and interactions. I also found a significant difference in posttest knowledge between the control and the combined intervention groups ($d = 0.30$, $p < .001$), an effect size that is very close to that found in the original study ($d = 0.33$). As such, H1 was confirmed.

[INSERT TABLE 2 AROUND HERE]

When posttest plausibility was the main outcome, I found a significant and positive effect of the baseline estimation game compared with the control group before and after including moderators and interactions. Using contrasts to combine intervention conditions and compare with the control, I found no significant difference in posttest plausibility ($p = .158$).

RQ2: Moderating Effects of Math Anxiety and Actively Open-Minded Thinking. When math anxiety was the moderator, I found no direct or moderating effects on the relations between experimental conditions and knowledge or plausibility. Despite finding a marginally significant negative moderating effect of math anxiety on posttest knowledge among those assigned to the numerical estimation instruction condition ($\beta = -0.126$; $p =.069$), the effect was not significant at the .05 level. Thus, H2.1 was not supported.

When Actively Open-Minded Thinking was the moderator, findings revealed a significant direct effect on knowledge and a moderating effect on the relation between the modified intervention and knowledge. On average, the effects of the intervention modified with estimation instruction were .20 *SD*s stronger for people with Actively Open-Minded Thinking levels 1 SD above the mean. When posttest plausibility was the outcome, I found no significant main or moderating effects. Thus, I found partial support for H2.2 in that epistemic dispositions moderated the effects of the modified intervention on posttest knowledge, but not on plausibility perceptions.

RQ3: Comparison Between Intervention and Modified Intervention. Comparing the two intervention groups to each other revealed no significant differences after controlling for pretest scores. This was true for both outcomes: posttest knowledge (p = .180) and plausibility perceptions (p = .117). That is, I found no effects of supplementing the estimation game with estimation instruction on the main outcome variables, H3 was not supported.

RQ4: Self-Reported Estimation Strategies. A summary of the estimation strategies used by participants by condition is presented in Table 3. As predicted, qualitative reports of estimation strategies revealed that significantly more participants in the modified intervention group shared tolerant reactions to accuracy feedback (H4a) and reported using the benchmark strategy more often (H4c) compared with the baseline intervention group. However, I found that no participants reported using rounding techniques, thus I did not find the predicted differences in rounding strategies between conditions (H4b).

[INSERT TABLE 3 AROUND HERE]

In Table 3, I also report additional estimation strategies that emerged but were not preregistered. Of these strategies, I found that significantly fewer students who received estimation instruction reported using “wild guesses” to estimate climate change numbers and significantly more students reported using arithmetic operations on given numbers compared with the baseline intervention group. Students who received estimation instruction also less often referred to news or social media as a source for estimates. I also found that participants across both conditions reported drawing from prior knowledge (educational experiences, personal experiences, vague statements of prior knowledge, and use of information from previous items), though there were no significant differences in use of prior knowledge between conditions.

Discussion

The purpose of this preregistered experimental study was to investigate whether people revise their conceptions after encountering novel statistics in an online intervention, and whether that learning is enhanced with additional numeracy instruction. Findings indicated that undergraduate students assigned to a game-based estimation intervention outperformed those in an expository text control group, and that effects were more pronounced among participants with adaptive epistemic dispositions who also received additional estimation instruction, replicating findings from prior work (Thacker & Sinatra, 2022; See Table 4 for a summary of all preregistered research questions, hypotheses, and results).

In addition to replication, this study extends prior work in several ways. First, the central intervention used in this study was redesigned to be more openly accessible to students and teachers by presenting novel data using a game-like web app. Second, the online intervention was revised from the original version to better promote climate change plausibility perceptions by providing: (a) explanations for the “true values” to improve credibility and coherence of the information, (b) personalized accuracy feedback to signal reflection on belief-discrepant information as well as enhance engagement, (c) sources of scientific information to improve its credibility, and (d) instruction on a more comprehensive set of estimation strategies. Third, and potentially because of these design revisions, I found previously undetected effects of the intervention on plausibility perceptions when compared with a control group. Fourth, this study assessed estimation strategies that students reported using and showed that students used more explicit estimation strategies when assigned to receive supplemental estimation instruction emphasizing the benchmark strategy, the importance of tolerating imperfect estimates, and flexible rounding strategies.

[INSERT TABLE 4 AROUND HERE]

However, before discussing the findings in more detail, I would like to acknowledge some of the limitations of this study. First, this study relies on the use of self-reported measures, which can have questionable validity in contexts where people share information about sensitive topics such as their beliefs about climate change (Tourangeau et al., 2013). As such, I took necessary precautions to reduce social desirability bias by ensuring participants of their anonymity and allowing them to complete surveys in a private setting. Second, the sample in this study was more female (79.3%) than the national average (58%; Irwin et al., 2022). Future studies might aim to recruit a more nationally representative sample. Third, although the control text and the intervention were parallel in many regards (e.g., both had similar word count, presented information about the mechanisms and impacts of climate change, and emphasized that there is a scientific consensus that climate change is happening), they were not identical in terms of the information presented. Future research might investigate how a wider variety of expository texts perform compared to the estimation game. Fourth, the central intervention in this study tested the effects of an intervention that combined many design features to maximize explicit reflection on belief-discrepant numbers. Though it is expected that people use multiple strategies in combination while estimating real world measures (e.g., Siegler & Booth, 2005), it might be considered a limitation that the research design makes it difficult to disentangle the unique contributions of each strategy. Future research might explore the unique effects of each of the three numerical estimation strategies emphasized in this study and interactions between them.

A Numerical Estimation Game for Climate Change Learning

I found greater learning outcomes among students who engaged with the online estimation game. On average, the two groups assigned to estimate climate change numbers before seeing the true value performed about a third of a standard deviation better than a control group on a climate change knowledge posttest, as predicted. Findings replicate the effects found by Thacker & Sinatra (2022) showing that novel climate change data can support climate change learning when presented in a game-based online intervention.

However, unlike the original study which found no effect of condition on personal plausibility perceptions, I found a significant and positive impact of the baseline estimation game on plausibility perceptions compared with the control. Such shifts in climate change plausibility might be explained by improvements to the intervention design. Unlike the original study, the intervention in the current study leveraged several design principles inferred from conceptual change theory. The design was revised to explain, contextualize, and provide credible sources of novel data as to enhance its comprehensibility, coherency, and credibility—all important preconditions for plausibility appraisals and conceptual change around socio-scientific topics (Dole & Sinatra, 1998; Lombardi et al., 2016). The estimation game also presented learners with game-like accuracy feedback to signal the potential need to explicitly reflect on belief-discrepant information and enhance engagement, both important mechanisms involved in plausibility appraisals and conceptual change (Lombardi et al., 2016). Furthermore, this finding supports research suggesting that people experience deeper learning from their errors when presented with feedback and external guidance to support productive self-explanations, which helps the learner detect and correct their misunderstandings (Zhang & Fiorella, 2023). Future research might test the effects of each of the novel design principles individually rather than in combination.

Moderating Effects of Epistemic Dispositions

Results also revealed that students' willingness to reason with new evidence was a significant predictor of their climate change learning and that it also moderated the effects of the intervention when supplemented with estimation instruction. These results replicated findings from Thacker & Sinatra (2022) and showed that students with higher levels of Actively Open-

Minded Thinking had greater learning outcomes if assigned to the game supplemented with math instruction. The moderating effects of Actively Open-Minded Thinking provides support for the PJCC (Lombardi et al., 2016) in that instruction that emphasizes the explicit evaluation of evidence appears to be most effective among those willing to consider belief-discrepant information. Based on these findings, educators might consider emphasizing to students the importance of keeping an open mind when examining new types of numerical evidence, even if the evidence is contrary to their current beliefs.

I also found a marginally significant moderating effect of mathematics anxiety on the impact of mathematics instruction. People with mathematics anxiety were only slightly negatively affected by the estimation instruction. One reason for this relatively null finding could be that the mathematics instruction presented in this intervention may have been less triggering of anxiety when compared to a typical undergraduate mathematics classroom context. For example, the intervention in this study was completed in a private setting, potentially reducing the performance aspects of anxiety associated with mathematics (Ramirez et al., 2017).

Effects of Supplementing the Game With Numerical Estimation Instruction

Consistent with the original Thacker & Sinatra (2022) study I also found that, on average, people assigned to the intervention modified with estimation instruction had no additional learning or plausibility benefits compared to the baseline intervention. One explanation for this finding is that the baseline intervention may have been equally effective at encouraging explicit evaluation of quantities as the version modified with estimation instruction. Another explanation is that the outcome measures may not have been sensitive to capture the learning that occurred from this very short micro-intervention.

However, despite finding only weak effects of estimation instruction on climate change learning, the current study extends prior research by examining the effect of this instruction on reported estimation strategies, revealing more nuanced learning outcomes. Participants who received additional estimation instruction when playing the estimation game reported using more explicit computation strategies such as making use of given information, mathematically manipulating that information, and fewer wild guesses. They also tended to report more tolerant reactions to imprecise estimates. In other words, students who received light touch estimation instruction reported using more explicit numerical estimation strategies compared with those who did not. These explicit processes might explain the learning benefits experienced by people who were open to reason with new evidence. As posited by the PJCC, conceptual change is more likely to occur when learners process information *and* are willing to explicitly engage with potentially belief-discrepant claims (Lombardi et al., 2016; Richter & Maier, 2017). Findings from this study support this idea, showing that estimation instruction encouraged students to apply more explicit data processing strategies and increased their awareness of the discrepancies between their estimates and the true values—leading to more conceptual change among those willing to change their mind based on such discrepancies. As such, findings show that bolstering numeracy skills can shape problem solving strategies as to enhance explicit, analytical reasoning with data, the very skills that may be crucial for enhancing learning amongst students with adaptive epistemic dispositions discussed earlier.

I should also note that, contrary to what was hypothesized, there was little evidence that participants employed digit rounding techniques that were emphasized in the estimation instruction. One reason for this could be explained by the context. In this study, participants responded to an open-ended item after estimating 12 climate change numbers and retrospectively reported strategies they had used. In previous research (Thacker et al., 2021), participants had reported their strategies in-situ during think-aloud interviews. As such, ephemeral estimation strategies may have been under reported in the current study. Future research might improve upon this study design by integrating strategy reports to appear immediately after participants were presented with accuracy feedback.

Implications and Future Directions

This study provides empirical support for the PJCC (Lombardi et al., 2016) and replicates and extends findings from Thacker & Sinatra (2022). Results demonstrated that a just handful of novel climate change numbers presented in a brief online intervention incited knowledge revision and this change moderated by epistemic dispositions, as predicted by the PJCC. Estimation instruction also enhanced learning outcomes for students who were open to reason with novel evidence and significantly increased reported use of benchmark strategies, a tolerance for error, and reduced the frequency of vague guessing, supporting the idea that explicit estimation strategies can be enhanced with light-touch instruction. Indeed, findings suggest that bolstering estimation heuristics may benefit people with adaptive epistemic dispositions potentially *because* it enhances explicit processing—a combination of factors that are predicted to promote shifts in plausibility perceptions and conceptions (Lombardi et al., 2016). Future experimental research might utilize interventions that enhance Actively Open-minded Thinking (c.f., Chang et al., 2016) to test causal relationships with explicit reasoning and climate change learning. Future research might also investigate additional factors that are predicted to moderate learning processes, such as motivational and emotional factors as well assess additional outcomes such as behaviors and intentions (c.f., Sinatra et al., 2012).

The intervention presented in this study and the design principles that it embodies also present a potential toolkit for the communication of climate change science. This study demonstrates that people can shift their misconceptions when prompted to make predictions before being presented with scientific evidence, and that supporting this evidence with personalized feedback and information that bolsters its credibility and coherence can also enhance plausibility perceptions of climate change. Consistent with prior research, evidence from this study supports the idea that effective climate change communication requires more than merely presenting people with accurate information; it is important to consider people's dispositions, beliefs about knowledge, emotion, and motivation (see e.g., Muis et al., 2015; Sinatra et al., 2014). Indeed, communication with the public about climate change involves overcoming many unique psychological barriers that are connected with their social and political identities (Stoknes, 2015), and supporting people in overcoming those barriers requires careful attention to building their skills and dispositions around evaluating evidence and claims.

Findings from this study also show that explicitly emphasizing mathematical skills may have important implications for educators. Adding numerical estimation instruction to the baseline intervention supported the use of explicit estimation strategies and bolstered climate change learning among people with adaptive epistemic dispositions. As such, educators might consider supplementing number estimation activities with some basic advice around useful estimation strategies, such as emphasizing when and how to use benchmarks, and by reminding students that rough estimates of numbers need not be perfect and that some level of imprecision is inevitable. Instructors might also consider that such explicit estimation strategies seem to be more effective for students who are more willing to engage with belief-discrepant information and tailor the level of mathematics instruction based on the makeup of their classroom.

Instructors might additionally, aim strengthen students' adaptive epistemic dispositions by communicating the importance of keeping an open mind when engaging with belief-discrepant evidence.

Lastly, this study provides both mathematics and science instructors with a tool that enables applications of mathematical skills to learn about climate change. The central intervention used in this study provides educators concerned with public understanding of science with an easily accessible learning application that can be easily shared with students, can be customized to their needs, and adapted to provide opportunities for students to apply numeracy skills towards making meaning of key numbers that describe our changing environment.

Conclusion

This study expands our understanding of conceptual change processes when people learn from climate change data. Supporting numeracy skills that promote data comprehension (such as connecting data to current understandings, the use of benchmarks, and a tolerance for estimation error) can support conceptual change outcomes, particularly among people willing to engage with belief-discrepant evidence. This contributes to the idea that conceptual change about human-induced climate change is not a simple matter of presenting people with scientific evidence, but rather, involves engaging the learner with that information, supporting their comprehension skills, eliciting predictions and explicit reasoning, providing feedback, and attending to their openness to consider new evidence. Such skills are critical to develop scientific habits of mind and make meaning of information in a way that helps people make informed decisions that can benefit society and the environment.

Table 1.
Descriptives by Condition and Intercorrelations Among Variables for the Main Analytic Sample of Undergraduate Students.

	Total			**Control**			**Estimation Game**			**Modified Estimation Game**			**Correlations**				
	n	M	SD	n	M	SD	n	M	SD	n	M	SD	pre.kn	post.kn	pre.pl	post.pl	anx
Pretest Knowledge	605	4.83	0.55	203	4.85	0.53	204	4.82	0.53	198	4.83	0.58	-				
Posttest Knowledge	605	4.94	0.63	203	4.81	0.62	204	5.02	0.59	198	4.97	0.67	.66***	-			
Pretest Plausibility	605	4.86	1.19	203	4.89	1.18	204	4.86	1.21	198	4.83	1.18	.58***	.55***	-		
Posttest Plausibility	605	4.9	1.28	203	4.86	1.26	204	4.99	1.23	198	4.84	1.36	.51***	.63***	.79***	-	
Mathematics Anxiety	605	2.9	1.03	203	2.93	1.02	204	2.93	1.02	198	2.85	1.05	0.04	-0.05	0.02	0.03	-
AOT	605	4.93	0.90	203	4.91	0.93	204	4.98	0.85	198	4.9	0.92	.47***	.54***	.46***	.44***	-.13**

Note. AOT = Actively Open-Minded Thinking. ***p* < .01; ****p* < .001

Table 2.
*Effects of Experimental Conditions on Posttest Knowledge and Plausibility and Moderating Effects of Actively Open-Minded Thinking and Math Anxiety (*N *= 605).*

	Posttest Knowledge			**Posttest Plausibility Perceptions**		
	No Moderator	Math Anxiety as Moderator	AOT as Moderator	No Moderator	Math Anxiety as Moderator	AOT as Moderator
	β (*SE*) *p*	β (*SE*) *p*	β (*SE*) *p*	β (*SE*) *p*	β (*SE*) *p*	β (*SE*) *p*
Estimation Game	**0.366***** **(0.070)** **p < .001**	**0.366***** **(0.070)** **p < .001**	**0.339***** **(0.068)** **p < .001**	**0.124*** **(0.058)** **p = .034**	**0.123*** **(0.058)** **p = .035**	**0.113*** **(0.056)** **p = .045**
Estimation Game + Estimation Instruction	**0.265***** **(0.074)** **p < .001**	**0.257***** **(0.073)** **p < .001**	**0.272***** **(0.068)** **p < .001**	0.025 (0.060) p = .684	0.025 (0.060) p = .677	0.024 (0.060) p = .687
Moderator		-0.022 (0.045) p = .620	**0.207***** **(0.044)** **p < .001**		-0.013 (0.042) p = .752	**0.072~** **(0.038)** **p = .062**
Intervention * Moderator		-0.021 (0.070) p = .762	0.057 (0.063) p = .365		0.047 (0.060) p = .430	0.060 (0.054) p = .268
Modified Intervention * Moderator		**-0.126~** **(0.069)** **p = .067**	**0.201**** **(0.064)** **p = .002**		0.027 (0.062) p = .657	0.017 (0.060) p = .773

Note: AOT = Actively Open-Minded Thinking. The comparison condition is the control condition in which participants read an expository text about the greenhouse effect. All models include pretest scores as a covariate, namely, posttest knowledge is adjusted for pretest knowledge and posttest plausibility is adjusted for pretest plausibility perceptions. All continuous variables were standardized around the mean. Standard errors are heteroskedasticity-robust. Boldfaced values indicate significant results for predictors.
~$p < .1$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 3.

Estimation Strategies Employed by Undergraduate Students While Estimating Climate Change Numbers

Strategy	**Total** Number of participants who reported using strategy	**Estimation Game Condition** % in this group who reported using strategy	**Modified Estimation Game** % in this group who reported using strategy	*z*	*p*
Preregistered Strategy Comparisons					
Tolerance for Error / as Reaction to Feedback	34	5.9%	11.1%	**3.72**	**<.001**
Flexible Rounding	0	0.0%	0.0%	na	na
Benchmark (unspecified usage)	71	12.7%	22.1%	**2.62**	**.004**
Exploratory Strategy Comparisons					
Benchmark + Extrapolation	56	15.7%	11.8%	-1.03	.151
Benchmark + Mental Iteration or Proportional Reasoning	10	1.0%	3.9%	1.97	.024
Prior Educational Experiences	48	14.2%	9.3%	-1.43	.077
News or Social Media	27	8.8%	4.4%	-1.71	.043
Personal experiences	11	2.5%	2.9%	0.36	.361
Prior Item	6	1.0%	2.0%	0.86	.195
External Resources	11	2.0%	3.4%	0.97	.167
Vague Guess	121	34.3%	25.0%	**-1.87**	**.031**
Nonsense	18	4.9%	4.0%	0.41	.338
n	402	204	198	-	-

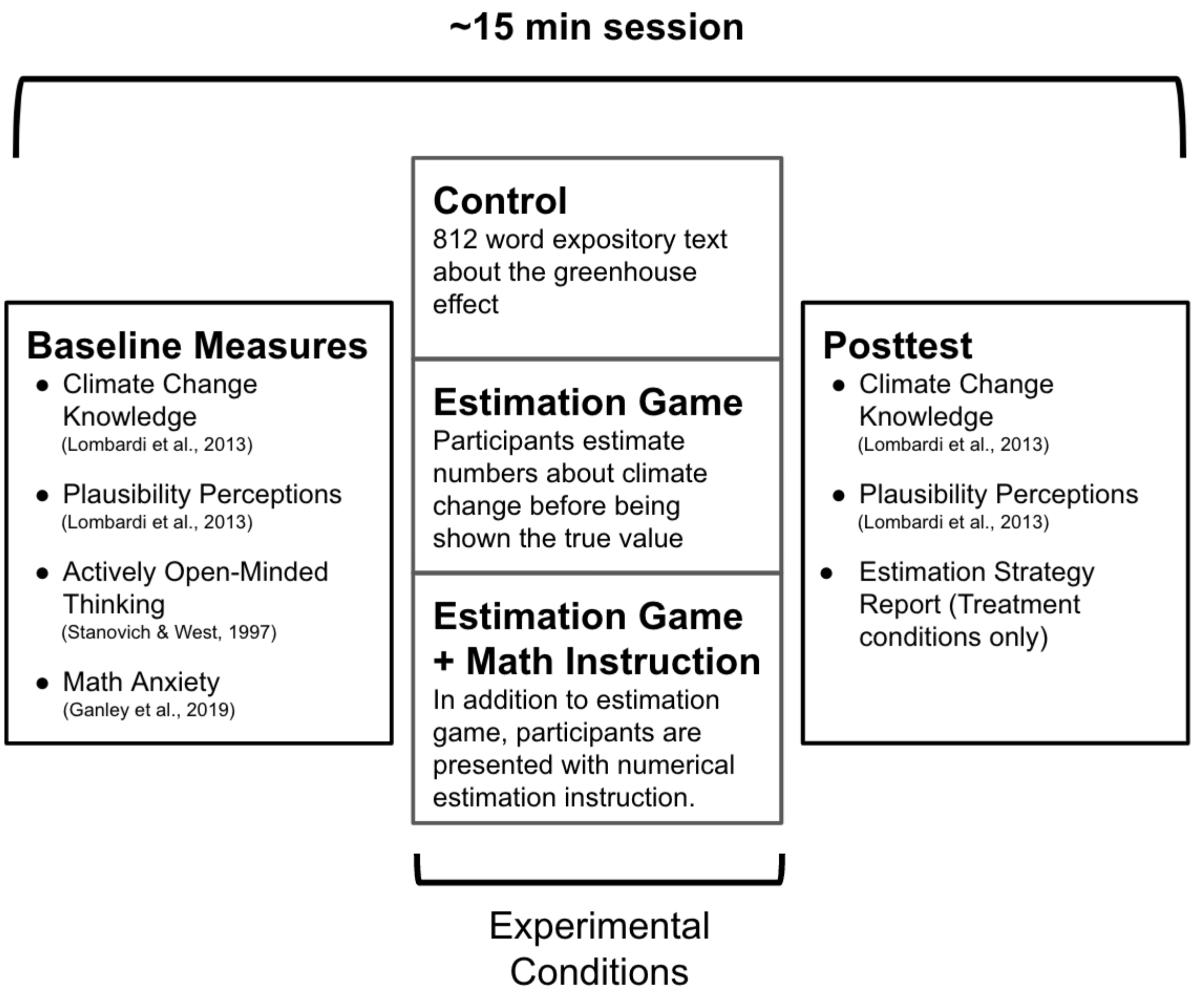
Note. Participants in the control group did not estimate numbers or complete a strategy reports, thus are not included here. The z and p-values were calculated based on a two-sample difference in proportions. Bold values are significant after Bonferroni adjusting p-values to account for the family-wide error rate among strategy types.

Table 4.
Preregistered Research Questions, Hypotheses, and Findings.

Research Question	Hypotheses	Findings
RQ1. To what extent does estimation of and exposure to novel climate change data using an online learning intervention improve learners' climate change knowledge and plausibility judgments compared with reading an expository text?	H1. I hypothesize that people assigned to the intervention conditions will have greater knowledge at posttest when compared with people assigned to read an expository text. I anticipate no significant differences in climate change plausibility perceptions between the intervention groups and the comparison group at posttest.	✓ Intervention conditions had significantly greater posttest climate change knowledge when compared with the control (d = .30, p < .001). ✓ No significant differences between posttest plausibility perceptions when comparing intervention conditions to control conditions (p = .158).
RQ2. To what extent do warm constructs (i.e., mathematics anxiety and epistemic dispositions) moderate the effects of the interventions on knowledge and plausibility?	H2.1. Math anxiety: I anticipate that individuals with high math anxiety will benefit less from the mathematics skills instruction. That is, I expect that math anxiety will negatively moderate the effects of the intervention modified with estimation instruction when the outcome is climate change knowledge or plausibility perceptions.	✕ Only a marginally significant moderating effect of mathematics anxiety on knowledge (p = .084)... ✕ …and no significant effect on plausibility perceptions (p = .877).
	H2.2. Epistemic dispositions: I anticipate that individuals with higher levels of Active Open-minded Thinking (AOT) will benefit more from the intervention and modified intervention compared with the comparison condition when the outcome is climate change knowledge. That is, I expect that active open-minded thinking will positively moderate the effects of the intervention. I also anticipate that there will be a significant main effect of active open minded thinking on plausibility perceptions, but no significant interactions with the experimental conditions.	✓ AOT had significant main effects (β = 0.21, p < .001) and moderating effects (β = 0.20, p = .003) on posttest knowledge, ✕ No significant main or moderating effects on posttest plausibility perceptions (p = .266, p = .808 respectively).
RQ3. To what extent does enhancing the intervention with instruction on estimation strategies change learners' climate change knowledge and plausibility?	H3. I hypothesize that supplementing the intervention with instruction on estimation skills will lead learners to report more scientific knowledge about climate change compared with those who are assigned to the intervention but without estimation instruction. I expect no significant differences in climate change plausibility between these conditions.	✕ No significant differences between the intervention group and modified intervention group with regard to posttest climate change knowledge (p = .180)… ✓ …nor plausibility perceptions (p = .623).
RQ4. Will individuals who are assigned to interact with the intervention supplemented with instruction on estimation strategies report using the emphasized estimation strategies?	H4. I hypothesize that individuals who are assigned to the intervention supplemented with estimation instruction will more frequently report (a) tolerance for error, (b) rounding techniques, and (c) benchmark strategies compared with those assigned to the intervention without this modification.	✓ Estimation instruction predicted significantly more frequently shared tolerant reactions to accuracy feedback (p < .001)… ✓ … more frequent use of benchmark strategies (p = .004)… ✕ … but found that no students reported using rounding techniques.

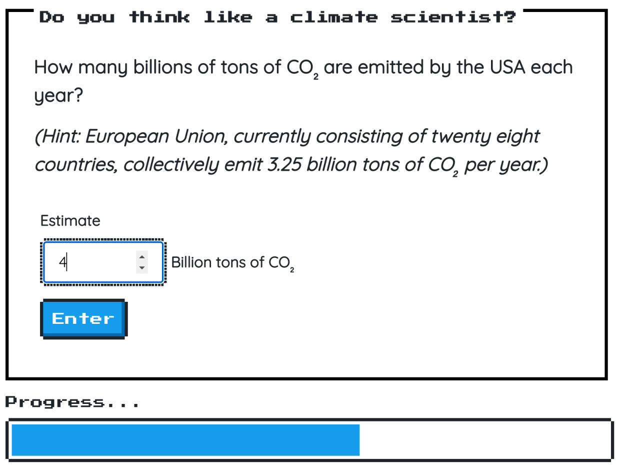
Note. The official preregistration can be accessed using the following link: https://osf.io/9uh6j?view_only=d51fd61e6fa5458c82499b5dff40a6d6.

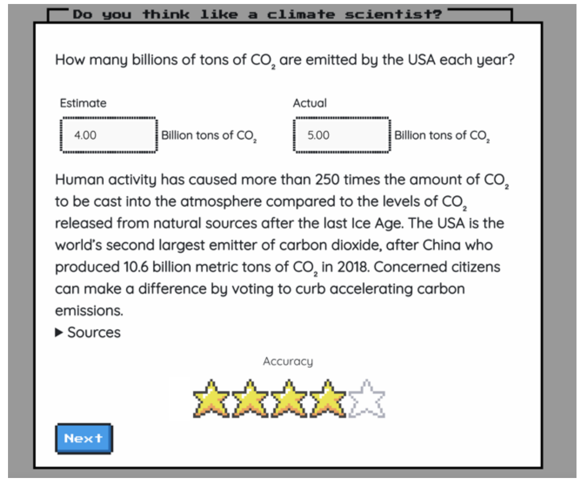
Figure 1
Visual Representation of the Survey Flow, Materials, and Procedures



Appendix A
Examples of The Experimental Conditions

Intervention Group (Screenshots from the "Estimation Game")





Modified Intervention Group (Excerpts from estimation instruction that preceded Estimation Game)

Sometimes people estimate everyday numbers in their head. For example, you might quickly estimate the cost of tax and tip in your head before ordering a meal at a restaurant. These calculations are naturally very rough and imprecise, and it is okay if your guess is not perfect…

REFERENCE NUMBERS

Numbers that you already know (reference numbers) can help you estimate numbers that you do not know. For example, if you know that about 300 pennies fit in a small, 8oz milk carton, you can use this information to estimate the number of pennies that fit in a gallon…

SIMPLIFYING NUMBERS

When using reference numbers, you may want to round values to make mental computation easier. For example, let's estimate the population of California given that the population of Kentucky is 4.47 million. Before making our estimate, we first round the Kentucky population to 4 million to make the math easier, and scale this value according to our beliefs about the size of California compared to Kentucky. If you were to guess that California…

Control Group (Excerpt of Expository Text adapted from Lombardi et al., 2013)

THE ENHANCED GREENHOUSE EFFECT

Many people have heard of the "greenhouse effect", but not everyone knows what the "greenhouse effect" is exactly. The greenhouse effect refers to the way that certain gasses in earth's atmosphere keep the planet warmer than it would otherwise be. The earth's greenhouse effect is a natural occurrence…

Note. All intervention texts appeared within the estimation game environment (ianthacker.com/design.html). For complete intervention materials, as well as survey materials, data and analysis script, see the supplemental materials (https://osf.io/nqbyr/?view_only=bc36d59f82424c4f8e67cfb1538afaaf).

Appendix B
Estimation Strategy Usage

Strategy Report Coding Scheme with Definitions and Examples.

Code	**Definition**	**Example**
Tolerance for Error	Evidence that the individual reacted positively to accuracy feedback despite estimates that were not perfect or precise.	● I knew it wasn't going to be 100% accurate ● I didn't remember all the numbers exactly but I made rough estimates based on what I did remember and just common sense
Flexible Rounding	Evidence that the individual was rounded numbers to make mental computation easier.	● I would round numbers up or down to make things easier*
Benchmark Estimation (unspecified)	Evidence that the individual used the given benchmark values without specifying what they specifically did with the benchmarks.	● I used the hints ● I used the given information ● I read the information above and did my best to estimate based on the questions
Benchmark Estimation + Extrapolation	Evidence that the individual used the given benchmark information to project trends to estimate the unknown quantities.	● I used the data from the past to make educated guesses based on the trends ● Depended on the increase or decrease which I would follow the trend ● I used the bit of evidence they gave and extrapolated from there. ● I took the change provided and tried to make an equivalent increase or decrease based off the information given ● I made my increase or decrease decisions by aligning all my answers with the fact that our environment is experiencing Increased emission and temperatures. Then I tried to choose my numbers based on hints given.
Benchmark estimation + Mental iteration *or* proportional reasoning	Evidence that the individual is using arithmetic operations to estimate an unknown number (e.g., repeated addition or proportional reasoning)	● I went off the amount of years and multipled (*sic*) ● For questions about increase/decrease of greenhouse gasses, water level, etc, I tended to double or triple numbers from previous statistics. ● I calculated the previous amount and subtracted. ● I just guessed that the answer would be from adding it subtracting half of the values

Prior Educational Experiences	The individual references information from their prior learning experiences to estimate or make sense of unknown quantities.	• From my memory of chemistry concepts • i tried to remember what i learned in environmental sciences class • I just guessed based on books I've read a while ago
News or Social Media	Evidence that the individual drew from information they had learned from various news sources.	• I used inferences based on what I've seen and heard on the news • knowledge I've gained through social media and outside of the media.
Personal Experiences	The individual references information from their personal experiences to estimate or make sense of unknown quantities.	• I think the levels are rising. Based on experience • personal info I knew before • Personal knowledge
Prior Item	Evidence that the individual used information from a previous item in the game to estimate the given number.	• I tried to use remember stats from the previous questions • I tried to choose my numbers based on previous questions
External Resource	Evidence that the individual used external resources to estimate quantities while answering questions. This includes use of a calculator, looking up answers or "Googling answers."	• I looked up the questions on google and did my research • Calculator • I use my calculator and thought of the object that they were using
Vague Guess	Vague guessing strategy specified by the participant	• i guessed • I estimated most of them • I just used my intuition
Nonsense	Nonsense or gibberish entered to trick forced response options.	• Tyijcxdhjlp ghkouff

Note. All examples represent real excerpts from student data, unless marked with "*" which represents an illustrative example that was not observed.

References

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189–201. https://doi.org/10.1037/0012-1649.41.6.189

Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: a framework for understanding real-world quantitative estimation. *Psychological Review, 100*(3), 511.

Brown, N. R., & Siegler, R. S. (1996). Long-term benefits of seeding the knowledge base. *Psychonomic Bulletin & Review, 3*(3), 385-388.

Brown, N. R., & Siegler, R. (2001). Seeds aren't anchors. *Memory & Cognition*, *29*(3), 405–412.

Cartiff, B. M., Duke, R. F., & Greene, J. A. (2021). The effect of epistemic cognition interventions on academic achievement: A meta-analysis. *Journal of Educational Psychology, 113*(3), 477.

Chang, W., Chen, E., Mellers, B., & Tetlock, P. (2016). Developing expert political judgment: The impact of training and practice on judgmental accuracy in geopolitical forecasting tournaments. *Judgment and Decision making, 11*(5), 509-526.

Cheuk, T. (2012). *Relationships and convergences found in the Common Core State Standards in mathematics (practices), Common Core State Standards in ELA/literacy (student portraits), and a framework for K-12 science education (science & engineering practices)*. http://nstahosted.org/pdfs/ngss/explanationofvenndiagram.pdf

Connell, L., & Keane, M. T. (2006). A model of plausibility. *Cognitive Science, 30*(1), 95–120. https://doi.org/10.1207/s15516709cog0000_53

Dole, J. A., & Sinatra, G. M. (1998). Reconceptalizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2-3), 109-128.

Dowker, A. (Ed.). (2005). A good guess: Estimation and individual differences: Implications for psychology, neuroscience and education. *Individual differences in arithmetic* (pp. 141–151). Psychology Press.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Friedman, A., & Brown, N. R. (2000). Updating geographical knowledge: Principles of coherence and inertia. J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 900–914. https://doi.org/10.1037/0278-7393.26.4.900

Ganley, C. M., Schoen, R. C., LaVenia, M., & Tazaz, A. M. (2019). The construct validation of the math anxiety scale for teachers. *AERA Open, 5*(1), 2332858419839702.

Irwin, V., De La Rosa, J., Wang, K., Hein, S., Zhang, J., Burr, R., Roberts, A., Barmer, A., Bullock Mann, F., Dilig, R., and Parker, S. (2022). Report on the Condition of Education 2022 (NCES 2022-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [date] from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022144.

Jackson, C. D., & Leffingwell, R. J. (1999). The role of instructors in creating math anxiety in students from kindergarten through college*. The Mathematics Teacher, 92*(7), 583-586.

Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research, 68*(4), 413–449.

Joram, E., Gabriele, A. J., Bertheau, M., Gelman, R., & Subrahmanyam, K. (2005). Children's use of the reference point strategy for measurement estimation. *Journal for Research in*

Mathematics Education, 36(1), 4-23.

Leiserowitz, A., Maibach, E., Rosenthal, S., Kotcher, J., Carman, J., Neyens, L., Myers, T., Goldberg, M., Campbell, E., Lacroix, K., & Marlon, J. (2022). *Climate Change in the American Mind, April 2022.* Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.

Leiserowitz, A. (2006). Climate change risk perception and policy preferences: The role of affect, imagery, and values. *Climatic Change, 77*(1), 45-72.

Lombardi, D., Danielson, R. W., & Young, N. (2016). A plausible connection: Models examining the relations between evaluation, plausibility, and the refutation text effect. *Learning and Instruction, 44*, 74–86.https://doi.org/10.1016/j.learninstruc.2016.03.003

Lombardi, D., & Sinatra, G. M. (2012). College students' perceptions about the plausibility of human-induced climate change. *Research in Science Education, 42*(2), 201–217. https://doi.org/10.1007/s11165-010-9196-z

Lombardi, D., Nussbaum, E. M., & Sinatra, G. M. (2016). Plausibility judgments in conceptual change and epistemic cognition. *Educational Psychologist*, *51*(1), 35–56.

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students' climate change conceptions. *Learning and Instruction*, *27*, 50–62.

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist, 53*(3), 165–184. https://doi.org/10.1080/00461520.2017.1421465

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. Learning and Instruction, 39, 168-183.

Murphy, P. K., & Mason, L. (2006). Changing knowledge and changing beliefs. In P. A. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 305–324). Mahwah, NJ: Erlbaum.

Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research.* Oxford University Press.

NASA Global Climate Change. (n.d.a). *Climate change: How do we know?* Retrieved March 10, 2021, from https://climate.nasa.gov/evidence

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive psychology, 55*(3), 169-195.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational research, 63*(2), 167-199.

Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist, 56*(2), 90-100.

Ranney, M., & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Science*, *8*(1), 49–75.

Ranney, M. A., Shonman, M., Fricke, K., Lamprey, L. N., & Kumar, P. (2019). Information that boosts normative global warming acceptance without polarization: Toward JS Mill's political ethology of national character. In D. Wilkenfeld & R. Samuels (Eds.) *Anthology on the experimental philosophy of science* (pp. 61-96). (In Bloomsbury's Advances in Experimental Philosophy series). New York: Bloomsbury

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist, 53*(3), 145–164. https://doi.org/10.1080/00461520.2018.1447384

Reich, J. (2021). Preregistration and registered reports. *Educational Psychologist, 56*(2), 101-109.

Rescher, N. (1976). *Plausible reasoning: An introduction to the theory and practice of plausible inference.* Van Gorcum.

Reys, R. E., & Reys, B. J. (2004). Estimation in the mathematics curriculum: A progress report. *Beyond Written Computation*, *29*, 101–112.

Reys, R. E., Rybolt, J. F., Bestgen, B. J., & Wyatt, J. W. (1982). Processes used by good computational estimators. *Journal for Research in Mathematics Education, 13*(3), 183-201.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52*(3), 148-166.

Rinne, L. F., Ranney, M. a, & Lurie, N. H. (2006). *Estimation as a catalyst for numeracy: Micro-interventions that increase the use of numerical information in decision-making.* Proceedings of the Seventh International Conference of the Learning Sciences, 571–577.

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*(2), 344-357.

Shimizu, K., & Ishida, J. (1994). The cognitive processes and use of strategies of good Japanese estimators. In R. E. Reys & N. Nohda (Eds.), *Computational Alternatives for the twenty-first century: Cross-cultural perspectives from Japan and the United States* (pp. 161-178). Reston, VA: National Council of Teachers of Mathematics.

Sidney, P. G., Thompson, C. A., Fitzsimmons, C., & Taber, J. M. (2019). Children's and Adults' Math Attitudes Are Differentiated by Number Type. *The Journal of Experimental Education,* 1-32.

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science, 19*(3), 341-361.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation: A review. In J. I. D. Campbell, (Ed.). (2005), *Handbook of mathematical cognition*. New York: Psychology Press.

Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful—except when they're not. *Journal of Experimental Child Psychology, 120*, 39-58.

Sinatra, G. M. (2005). The" warming trend" in conceptual change research: The legacy of Paul R. Pintrich. *Educational Psychologist, 40*(2), 107-115.

Sinatra, G. M., & Hofer, B. K. (2021). *Science Denial: Why it happens and what to do about it.* Oxford University Press.

Sinatra, G. M., Kardash, C. M., Taasoobshirazi, G., & Lombardi, D. (2012). Promoting attitude change and expressed willingness to take action toward climate change in college students. *Instructional Science, 40*, 201-217.

Sinatra, G. M., Kienhues, D., & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist, 49*(2), 123-138.

Sinatra, G. M., & Seyranian, V. (2016). Warm change about hot topics: The role of motivation and emotion in attitude and conceptual change. In L. Corno & E. Anderman (Eds.), *APA handbook of educational psychology*. Washington, DC: APA Publications.

Sowder, J. T., & Wheeler, M. M. (1989). The development of concepts and strategies used in computational estimation. *Journal for Research in Mathematics Education*, 130–146.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342–357.

Stanovich, K. E., & Toplak, M. E. (2019). The need for intellectual diversity in psychological science: Our own studies of actively open-minded thinking as a case study. *Cognition, 187*, 156-166.

Stoknes, P. E. (2015). *What we think about when we try not to think about global warming: Toward a new psychology of climate action*. Chelsea Green Publishing.

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). Pearson.

Thacker, I. & Sinatra, G. M. (2022). Supporting climate change understanding with novel data, estimation instruction, and epistemic prompts. *Journal of Educational Psychology, 114*(5), 910–927. https://doi.org/10.1037/edu0000729

Thacker, I., Broadway, R., & Feder, S. (2021). Estimating climate change numbers: How tolerance for error can support science learning. In D. Olanoff, K. Johnson, & S. M. Spitzer (Eds.). *Proceedings of the forty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1738-1743). Philadelphia, PA.

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys.* Oxford University Press.

Vosniadou, S. (2013). *International handbook of research on conceptual change* (2nd ed.). Amsterdam, The Netherlands: Springer.

Zhang, Q., & Fiorella, L. (2023). An integrated model of learning from errors. *Educational Psychologist, 58*(1), 18-34.