Motivated Reasoning in a Causal Explore-Exploit Task

by

Zachary A. Caddick

B.A. in Psychology, California State University, San Bernardino, 2013

M.A. in Experimental and Research Psychology, San José State University, 2016

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Zachary A. Caddick

It was defended on

March 17, 2020

and approved by

Timothy J. Nokes-Malach, Ph.D., Associate Professor, Department of Psychology

Kevin R. Binning, Ph.D., Assistant Professor, Department of Psychology

Thesis Advisor: Benjamin M. Rottman, Ph.D., Associate Professor, Department of Psychology

Motivated Reasoning in a Causal Explore-Exploit Task

Zachary A. Caddick, M.S.

University of Pittsburgh, 2020

The current research investigates how prior preferences affect causal learning. Participants were tasked with repeatedly choosing policies (e.g., increase vs. decrease border security funding) in order to maximize the economic output of an imaginary country, and inferred the influence of the policies on the economy. The task was challenging and ambiguous, allowing participants to interpret the relations between the policies and the economy in multiple ways. In three studies, we found evidence of motivated reasoning despite financial incentives for accuracy. For example, participants who believed that border security funding should be increased were more likely to conclude that increasing border security funding actually caused a better economy in the task. In Study 2, we hypothesized that having neutral preferences (e.g., preferring neither increased nor decreased spending on border security) would lead to more accurate assessments overall compared to having a strong initial preference, however, we did not find evidence for such an effect. In Study 3, we tested whether providing participants with possible functional forms of the policies (e.g., the policy takes some time to work, or initially has a negative influence but eventually a positive influence) would lead to a smaller influence of motivated reasoning, but found little evidence for this effect. This research advances the field of causal learning by studying the role of prior preferences, and in doing so, integrates the fields of causal learning and motivated reasoning using a novel explore-exploit task.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

"*I'm not saying there won't be a little pain … we might lose a little bit … but we're gonna have a much stronger country when we are finished…. So, we may take a hit, and you know what, ultimately we're going to be much stronger for it.*"

- President Donald Trump (Factbase, 2018)

"*The tariffs are beginning to have some impact in a negative way so I hope that we make some progress quickly on some of these other fronts, in particular with China…. If the end result of this is better trading relationships with all of these countries, particularly if it happens sooner rather than later, I think it would be great*."

-Senate Republican Leader Mitch McConnell (Shepardson, 2018)

"*Trump Tariffs Are Short-Term Pain Without Long-Term Gain, Economists Say: Nearly three-fourths of economists in WSJ [Wall Street Journal] survey said they expect short-term trade costs to outweigh any long-term benefits*."

-Wall Street Journal Article (Torry, 2019)

Humans are often faced with the task of evaluating the causal efficacy of an action or policy in dynamic settings, which can be very challenging. For example, when a politician decides to implement a new economic policy (e.g., tariffs), assessing the true impact of the policy is likely to be very difficult because other factors in the economy also change over time, and because one's expectations about how fast the policy will work and the short versus long-term impacts of the policy could lead different people to focus on different evidence. For another example, when a patient is assessing whether a medication is working, it is also very complicated because medications have complicated profiles of how quickly and long they work for, and whether they produce short-term or long-term side-effects.

Even when learning simple cause-effect relations, prior beliefs and expectations have strong impacts on the assessment of the evidence (Alloy & Tabachnik, 1984; Fugelsang & Thompson, 2000, 2003; Goedert, Ellefson, & Rehder, 2014). However, in dynamic causal learning situations like those mentioned above, the task is considerably harder and likely to lead to different interpretations. Furthermore, in many situations, an individual might have strong preferences or engage in wishful thinking, which could bias their interpretations of the evidence. Yet, there is surprisingly little research at the intersection of causal learning and motivated reasoning, particularly in dynamic situations such as assessing economic policies, which was our goal. In the rest of the introduction we first discuss motivated reasoning, then causal learning in dynamic tasks, and finally propose a set of hypotheses that we tested in three studies.

1.1 Motivated Reasoning

Often when we reason about information we already have prior preferences pertaining to the subject matter. Individuals tend to more easily confirm information that is congruent with prior preferences (Nickerson, 1998) and reject information that is incongruent with prior preferences (Kunda, 1990). For example, Taber and Lodge (2006) found that individuals who had strong preferences about gun control or affirmative action were more likely to devalue arguments that were incongruent to their preference[1], regardless of their quality. This two-pronged process is known as motivated reasoning.

The current research specifically examines motivated reasoning about how people learn cause-effect relations. Although much of the recent work does not explicitly involve causality, some of the formative work on motivated reasoning involves how people assess causal claims. For a paradigmatic example, Kunda (1987) found that people tend to believe that their own attributes will lead to positive outcomes and reject the possibility that their attributes might lead to negative outcomes. In the first study, Kunda provided a description of a hypothetical person who had one of two attributes. Participants rated how likely the person was to get divorced based

[1] The terms 'preference', 'belief', and 'attitude' are often used interchangeably in the literature. For simplicity, we use the word 'preference,' and discuss potential differences in the general discussion.

upon this attribute. When the attribute (the cause) matched an attribute of the participant, they were less likely to view this attribute as leading to divorce (the effect). Study 2 was similar, but examined attributes predictive of success in graduate school, and found that individuals who didn't want to go to graduate school (lack of motivation) were less likely to engage in preferential reasoning. Study 3 examined how participants evaluate scientific evidence. Participants read a scientific article stating that caffeine consumption leads to poor health outcomes for women. Women who drank a lot of coffee found the evidence less convincing than those who drank only a little or none; however, for men there was no difference in the ratings of convincingness, presumably since the evidence was only relevant to women and hence men had no motivation to engage in biased reasoning.

Despite the fact that some of the foundational work on motivated reasoning involved causal reasoning (see Kunda, 1987; Kunda, 1990), much of the research that followed has not focused on causality (for examples, see: Campbell & Kay, 2014; Kaplan, Gimbel, & Harris, 2016; Hart & Nisbet, 2011; Klaczynski, 1997; Nyhan & Reifler, 2010; Paharia, Vohs, & Deshpandé, 2013). Additionally, often there is no presentation of statistical evidence between a potential cause and outcome. Instead, much of the research has focused on how people confirm or reject evidence for reasons aside from the data itself such as the news outlet it was reported through or the qualifications of the author of a scientific study, which is known as the credibility heuristic (Kahan, Braman, Cohen, Gastil, & Slovic, 2010). People also prefer information that comes from sources with similar ideological preferences over those with competing preferences–even when these individuals have demonstrated poorer knowledge in the relevant domain (Marks, Copland, Loh, Sunstein, & Sharot, 2018).

Inspired by political discourse around predicting and assessing policies, in the current study we sought to integrate research on motivated reasoning with a causal learning paradigm in which participants assess a cause-effect relation from evidence presented sequentially over time. We focused on whether people are able to learn to distinguish short versus long-term effects from experience and whether their prior preferences distort their assessments of the actual efficacy of the policies.

The studies we conducted differ in a number of ways from prior research on motivated reasoning about causal relations. First, instead of presenting participants with a verbal description of an anecdote or a verbal description of a scientific study (e.g., Kunda's 1987 study on caffeine), we examined how people interpret quantitative data. The closest motivated reasoning study to ours (Kahan, Peters, Dawson, & Slovic, 2017) had participants make causal assessments from data presented in a 2×2 contingency table of cross-sectional data. The contingency table presented evidence about cities that either banned handguns in public or not, and whether there was an increase or decrease in crime. Despite being presented with objective numbers, participants were more likely to make correct inferences about the influence of handgun policies on crime when the data supported their previously held preferences about handguns.

The studies we conducted presented participants with quantitative data that was either congruent or incongruent with their prior preferences. However, we used a much more complex task than the one used by Kahan et al. (2017). First, the data involved a sequence of events over time so it was time-series data rather than cross-sectional data. Second, participants had control over the policies rather than merely observing them, allowing us to measure biased information sampling. Third, by comparing simpler causal relations in which the cause fairly quickly produced the effect, versus more ambiguous relations in which the short-term and long-term influence of the cause are different allowed us to study whether more ambiguous cause-effect relations were more biased by motivated reasoning.

1.2 Causal Learning

One of the core questions addressed by the field of causal learning is how people learn the strength or influence of one or multiple causes on an effect (e.g., Spellman, 1996; Derringer & Rottman, 2018). Most of the research on causal learning has focused on situations that involve 'stable' causes; no matter when the cause is used, or how frequently or infrequently it has been used, it produces the same average outcome on average. In the current study, inspired by the examples of economic policies, we investigated how people learn about more complex causes; causes that take repeated usage to exhibit their full influence, or causes that exhibit different short-term and long-term influences.

1.2.1 Ambiguity. One fundamental features of the causal relations we investigated are that they are ambiguous. Causes that exhibit negative short-term outcomes but positive long-term outcomes (or vice versa) after repeatedly using the cause are especially ambiguous. Suppose a learner tries such a cause, and notices that quickly after trying it, it seems to produce a strong negative outcome. In this case, a learner may decide to stop using it, and may never even experience the long-term benefit. Or, suppose a learner tries a cause that produces a short-term benefit, and continues to use it and later experiences a long-term negative outcome. The learner may be able to detect this long-term negative outcome, or they might instead attribute the negative outcome to something else changing over time. The other causes we studied, in which a cause exhibits a positive or negative effect, but it takes some repeated usage to produce the maximal influence, are less ambiguous. Still, they are ambiguous in the sense that if they are only tested for a short amount of time, the learner will not realize how beneficial or harmful they actually are.

Ambiguity is a core feature of causal inference and has been studied in relation to other causal situations. For example, if a cause initially has a positive outcome (immediate), but later on when that cause is used again it has a negative outcome (also immediate), people often anchor their preferences about a cause-effect relationship based on the initially experienced events (Marsh & Ahn, 2006). This can be explained in that the initial experience influences how people interpret future experiences such as believing that the cause still has a positive influence but that something else has changed and is producing the negative outcomes (Luhmann & Ahn, 2007). As another example, Marsh and Ahn (2009) presented participants with data about a cause and effect, each of which were present or absent. Overall, there was a positive correlation; however, there were some instances in which the effect was present but the cause was ambiguous–it was unclear if it was present or absent. Participants tended to interpret the ambiguous cause as present merely because the effect was present, and because their prior experience led them to believe that the cause and effect were correlated. In sum, the previous research has shown that in ambiguous situations, people tend to interpret the evidence in ways that fit with prior beliefs developed from earlier in learning.

In the current study which involves motivated reasoning, the question is not whether participants' experiences early in learning impact their subsequent inferences, but rather whether there beliefs and preferences about economic policies, prior to any learning, affect the ways that they go about testing the policies and the inferences they make about the policies. We hypothesized that the prior preferences would have an influence on learning about all the causes since the task was hard and there was considerable noise, and that this effect would be even stronger for the more ambiguous policies that have different short versus long-term impacts.

1.2.2 Active Learning and Explore/Exploit Paradigms. Because our focus was on how people come to learn the efficacy of policies and to choose the policies to produce the maximal economic output, we created a dynamic causal explore-exploit task. Importantly, we wanted some of the policies to have fairly fast and clear impacts on the economy, and others to have more ambiguous influences in which the short-term impact contradicts the long-term impact.

To accomplish this goal, we created a task in which participants learned about 6 policies; on each trial they could choose between two different versions of the policy (e.g., increasing vs. decreasing border security funding). Two of the policies worked fairly quickly; after a couple of trials of using the policy it had their maximum impact. Two of the policies exhibited the temporal tradeoff of the short-term versus long-term costs versus benefits. To implement this, we adapted a paradigm sometimes called the ‘Harvard Game’ (see Sims et al., 2013, for a review). This paradigm is known to be difficult; participants often exhibit ‘melioration’ in that they primarily implement the version of the policy that produces the better short-term outcome but is sub-optimal in the long-run. In addition, there were also two more policies for which the different versions of the policies made no difference, which we call the ‘non-causal’ policies. Our goal for these was to test whether people could accurately learn that they were in fact non-causal.

Importantly, the task involved active learning, so participants’ goal was to try to simultaneously learn about which policies were best and to use these policies to produce the best outcome. Thus, similar to prior studies of melioration, we analyzed the percent of trials in which participants chose the optimal policy. But in addition, we also studied whether participants eventually learned explicitly which policies were better and the functional form of the policies.

2.0 Summary of Studies and Hypotheses

We conducted three studies to address the role of ambiguity and prior preferences in causal learning. Study 1 investigated the roles of prior preferences and ambiguity in causal learning. First, we hypothesized that participants would be more accurate in their policy assessments for policies that are less ambiguous (matching short-term effects and long-term effects) compared to policies that are more ambiguous (mismatching short-term and long-term effects). This was a precondition for a number of the subsequent questions.

Second, we tested whether participants would exhibit motivated reasoning; whether their choices when actively testing the policies, and their final assessments of the policies, would be biased by their prior preferences. Specifically, we tested whether this would occur even when those preferences are technically irrelevant to this hypothetical task and participants were incentivized for accuracy. Furthermore, we investigated not just whether motivated reasoning occurred, but specifically how it changed the ways in which they tested the policies.

Third, we tested whether the motivated reasoning effect would be exacerbated for the policies that where more ambiguous. This hypothesis assumes that when a cause-effect relation is more ambiguous, it could reasonably be interpreted in multiple ways, allowing more room for a bias to seep in.

Study 2 compared causal judgments when participants did versus did not have strong prior preferences. The main question was whether having strong preferences on average (across preferences that happen to be right and preferences that happen to be wrong) leads to more biased testing and less accurate judgments, compared to when participants are more open-minded (have neutral preferences).

Study 3 tested whether causal learning and judgments are affected by having more versus less knowledge of the potential functional relations between the causes and effect. We hypothesized that having more knowledge about the potential ways that the causes could influence the effect would lead to better strategies for testing the policies overall, and particularly benefit learning about the policies that have different short and long-term effects.

3.0 Study 1

3.1 Method

3.1.1 Participants. Fifty people participated via MTurk. Participants were paid $6.50 for participation (which amounted to approximately $8-10/hr). In addition, participants could earn up to $3.00 in bonuses contingent upon performance and were informed of their bonus total after the completion of the study.

3.1.2 Design. Each participant learned about six economic policies. Each policy had two options that participants chose between. For example, for the policy of border security funding, the two options were 'increasing border security funding' and 'decreasing border security funding'.

As explained below, with pretesting we selected policies for which each individual participant had very strong preferences that one option was better for the economy and the other was worse. For example, one participant might believe that increasing border security funding is better for the economy, and another participant might believe the opposite.

Independently from participants' preferences, we randomly assigned the six policies to one of six 'payoff functions' (*Figure 1*). The functions determined how each policy choice affected the economic output, which we called the 'Economic Vitality Index' or EVI for short. These two options were randomly assigned to either the better or worse states of the function. Thus, participants' preferences about the influence of a policy could either be preference-congruent (e.g., believing that more funding for border security is better for the economy, and indeed it was better), or preference-incongruent (e.g., believing that more funding for border security is better for the economy, but in fact it was worse). In addition, for two of the policies the options made no difference.

3.1.3 Economic Functions. Functions 1 and 2 were "clear" in that the policies made a change relatively quickly, and the change lasted as long as the policy was used (*Figure 1*). For Function 1, after the cause was turned from 'off' to 'on', it quickly produced an increase in the EVI. Function 2 was simply the opposite of Function 1 (negative coefficient signs); after the cause was turned from 'off' to 'on', it quickly produced a decrease in the EVI. The math behind these functions is based upon the idea of a decaying causal influence, similar to radioactive decay or a medication half-life. For example, imagine that the cause is a drug, which decays in half after each trial. At the end of Trial 1 after starting to take the medicine, 50% of the drug remains. At the end of Trial 2 of taking the medicine, 50% remains from Trial 2, and 25% remains from Trial 1, producing a 75% effect. At the end of Trial 3, 12.5% remains from Trial 1, 25% remains from Trial 2, and 50% remains from Trial 3, etc. If the drug is repeatedly taken, the effect approaches 100%. If and when the policy is turned off, the remaining effect from prior trials continues to decay. Equation 1 shows this function where policy *p* can be either on (1) or off (0) for each trial *t*.

$$EVI_t = .5(EVI_{t-1} + 100p_t) \qquad \text{(Equation 1; Function 1)}$$

Functions 3 and 4 are "ambiguous" in that the short-term effects of the policy are opposite to the long-term effects (*Figure 1*). For Function 3, when the policy is turned from 'off' to 'on', it immediately has a negative influence on the EVI, but eventually has a positive influence. Function 4 is the opposite; it initially has a positive influence but eventually has a negative influence. Functions 3 and 4 are similar to the function used in the melioration literature (e.g., Sims et al., 2013). These functions are somewhat analogous to a fixed-income security (e.g., treasury bond, certificate of deposit), and can also be viewed as somewhat analogous to the decision to buy versus rent a home. Importantly, Functions 3 and 4 have two defining features. First, there is a buy-in cost, which reduces the current EVI (analogous to spending money on the bond, CD, or a down-payment for a house, reducing one's current cash level). Second, there is a defined rate of return over time, and the cumulative return is larger than the initial cost, in this case twice as large. This means that if one keeps on buying the investment (using the policy) over and over again, initially the costs are substantial and one's cash deposits will be low. However, over time, as the dividends start to come in, one's cash level will be higher after repeatedly making the investment than if never investing at all. If one stops investing after having repeatedly invested, they will temporarily have an increased cash flow because of the incoming dividends, but over time they will taper away.

For Function 3, the investment function works such that when 100 EVI is invested, 200 EVI is returned over the following 10 trials. The rate of return on investment rises until it peaks 5 trials later, and then decreases; this is why in Function 3 p_{t-5} has the largest coefficient (50). The rate of return follows roughly a normal distribution from t-9 to t-1, which means that the cumulative payoff if left on is sigmoidal.[2] If this function is kept 'on', then eventually every trial will return 200 EVI (a net gain of 100 EVI). Upon being set to 'off', investments will no longer be made and only past investments will be returned (if any investments were made in the last 10 trials). Function 4 is the inverse to Function 3 (with negative instead of positive coefficients), and represents a policy that has short-term benefits but long-term consequences.

[2] Studies in the melioration literature typically use a flat payoff distribution over the prior 10 trials which results in straight lines instead of curved lines in Figure 1. We chose a slightly different payoff function in order to make the returns curved, similar to Functions 1 and 2. However, the general shape of the function is quite similar.

$$EVI_t = -100p_t + 5p_{t-1} + 10p_{t-2} + 20p_{t-3} + 40p_{t-4} + 50p_{t-5} + 40p_{t-6} + 20p_{t-7} + 10p_{t-8} + 5p_{t-9}$$

(Equation 2; Function 3)

For Functions 5 and 6, neither of the two options have any impact on the EVI, so they are called “non-causal”.

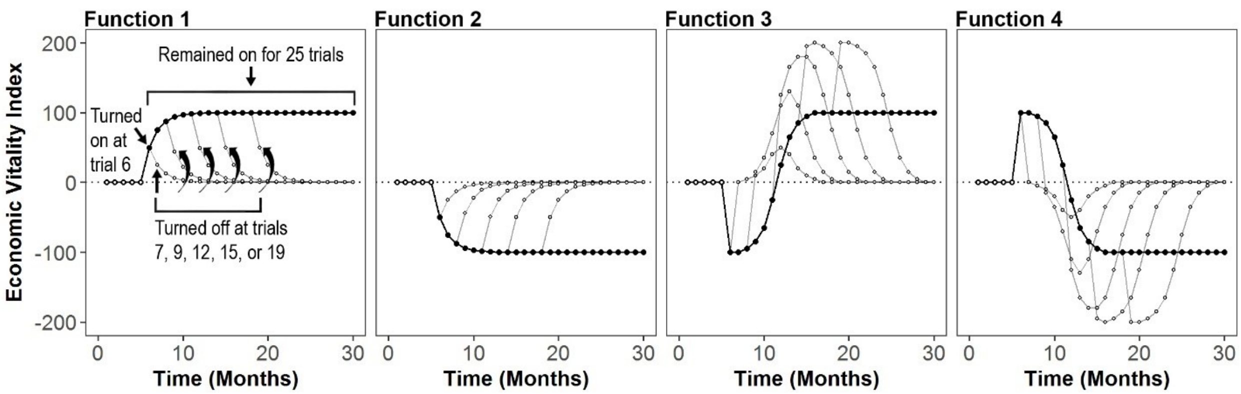

Figure 1. Illustrations of the payoff functions. Note: The first 5 trials in every graph represent an input of 'off' (white dots). The solid black line shows the function if it is turn on at Trial 6 and left on until Trial 30. The gray lines show the pattern economic output if the function is turned 'off' on Trial 7, 9, 12, 15, or 19, instead of being left on. Functions 1 and 2 are the "low ambiguity" (short-term and long-term effects match). Functions 3 and 4 are the "high ambiguity" (short-term and long-term effects are mismatched).

3.1.4 Procedures and Measures.

3.1.4.1 Initial Instructions. Participants were told to imagine that they had just been elected the leader of a large industrialized country. As the leader, they have the responsibility to make important decisions about economic policies with the goal of maximizing economic output. Before taking office, they must first evaluate a set of economic policies which will shape their economic platform.

3.1.4.2 Initial Policy Assessment. In order to choose 6 economic policies for each participant for which they had strong preferences that one option was better than the other, participants rated all 33 policies on two questions. One question was about their subjective preference for a particular policy option (see Appendix B) and the other question their objective belief about whether the policy would have a positive or negative impact on the economy (see Appendix C). For example, for the policy about border security, participants were asked "Would you prefer the government decrease or increase border security spending?" on a scale of 1 = strongly prefer decreasing to 7 = strongly prefer increasing, and they were also asked "Do you believe decreasing or increasing border security spending is better for the economy?" on a scale of 1 = strongly believe decreasing border security spending is better for the economy to 7 = strongly believe increasing border security spending is better for the economy. We asked about both beliefs and preferences because we assumed that they would be strongly correlated, and since we felt that it would be difficult to disentangle the two, we wanted to choose policies for which participants felt strongly for both preferences and beliefs.

After participants answered all 66 questions, the computer selected the six policies for which participants had the most extreme ratings measured as the extremity of the average of the two questions. Most participants had at least six policies that they rated maximally extreme (either a 1 or a 7 on both questions). These six policies formed the participants' policy platform and were used in the subsequent tasks.

3.1.4.3 Party Color Selection. Next, participants selected a color (purple, pink, orange, yellow, green, or brown) to represent their political party. Red and blue were omitted from the choices due to the strong association these colors have with the two main political parties in the United States. After selecting a color, the participant was presented with a color that represented the opposition party.

3.1.4.4 Economic Learning Task. The economic learning task was the primary task for the study. Participants' goal was to select economic policies that produced the highest economic output and to correctly assess which policies were best for the economy. Participants were told that they will receive a payment bonus based upon their average economic output for their time in office, relative to other participants' performance on the task, with a range of zero to two dollars. The six payoff functions were randomly assigned to the six policies.

Participants were presented with the six policies, randomly ordered on the screen (*Figure 2*). The screen presented the participants' preferred option with a square of the color of their party, and the non-preferred option with a square of the color of the opposing party. Initially each of the six policies were randomly set in either the 'on' or 'off' setting, which was framed as the policy selection of the prior administration. This random selection means that some of the prior policy decisions agreed with the participant's preference, and some disagreed.

The screen also displayed the current "Economic Vitality Index," which is intended to be a made-up economic indicator similar to the Gross Domestic Product or the stock market. The Economic Vitality Index (EVI) was a sum of the six payoff function outputs (*Figure 1*), plus a constant of 700, and a noise function. The noise function is a randomly generated Gaussian distribution with a mean of 0 and a standard deviation of 27. This degree of noise was selected to make the task hard, but not impossible. All payoff functions were initially set such that they have already reached their asymptote (see *Figure 1*), as if they have either been "on" or "off" for at least 20 trials.

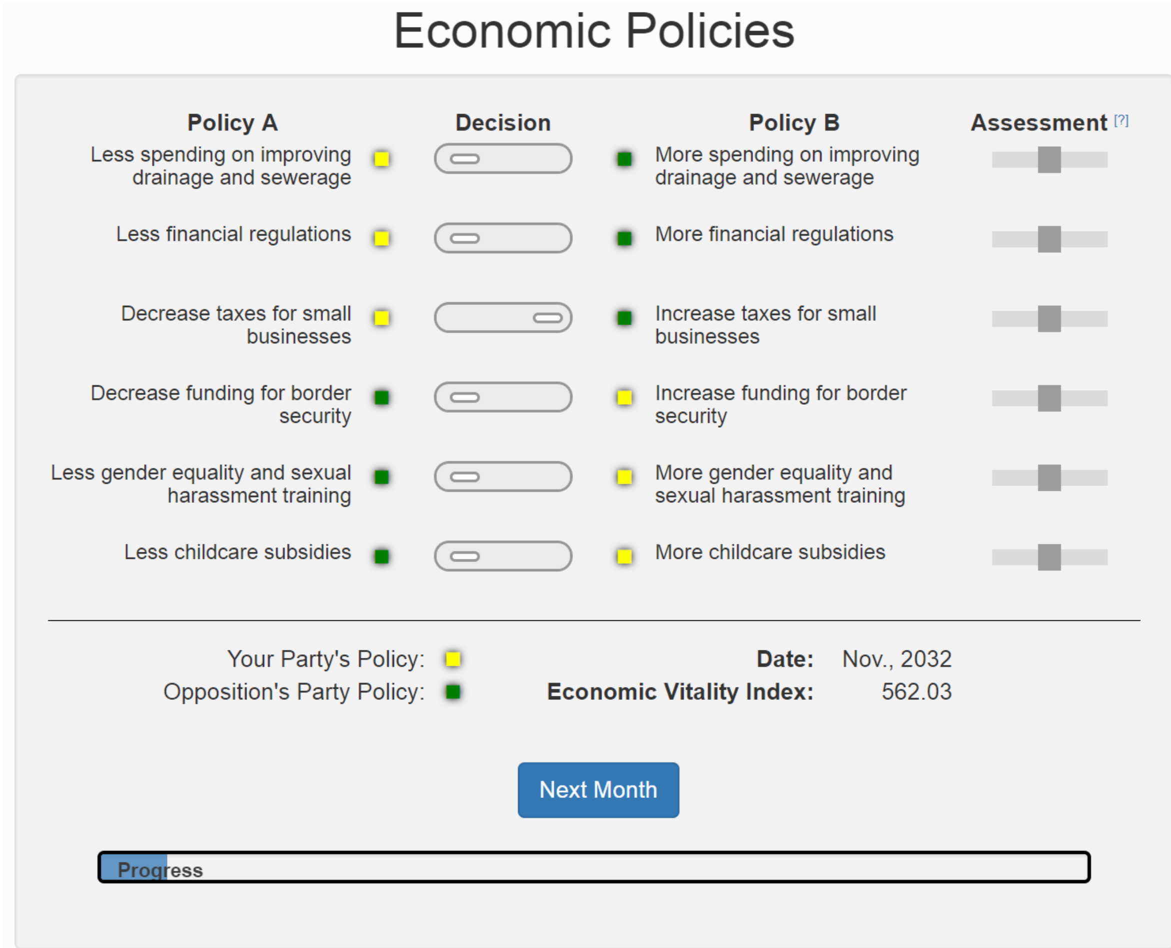

Figure 2. Economic learning task.

Participants experienced 150 trials and each trial represented one month in time. During the first 10 trials, the participants were told that they had not yet assumed power, so they just observed the six policies and observed the EVI of the prior administration. During these 10 trials, the six policies were held constant, and because the policies were already at asymptote, the change in the EVI across the 10 trials was only due to the noise function.

After the 10th trial, participants were told that they had been elected into office and could set the policies for the next 140 trials however they choose using the toggle switches in the ‘decision’ column. After they set the policies as they wish, they pressed the ‘next month’ button to go to the next trial, which revealed the EVI produced that month. At that point they could again make changes to the policies. Additionally, throughout the task participants were encouraged to use the slider in the "Assessment" column to track which policy option, A or B, they thought was better. The slider scale was from -5 to +5 and was initially set to 0. Left means that Policy A was better and right means that Policy B was better. After the last trial, participants were given one last opportunity to update their policy assessments.

3.1.4.5 Function Identification. After the 150 trials were over, participants' understanding of how each policy works was tested by matching each of the six policies to a figure that presents 8 possible functions. These 8 functions present the four unique functions from *Figure 1* plus the "non-causal" function (i.e., Function 5 & 6), as well as three additional functions as lures. We also included textual descriptions of the influence of each policy. Instructions were provided stating that the graphs show different possibilities of what might happen if you switch from Option A (e.g., "decreasing border security funding") to Option B (e.g., "increasing border security funding") and to select the graph that they think would result from this policy change.

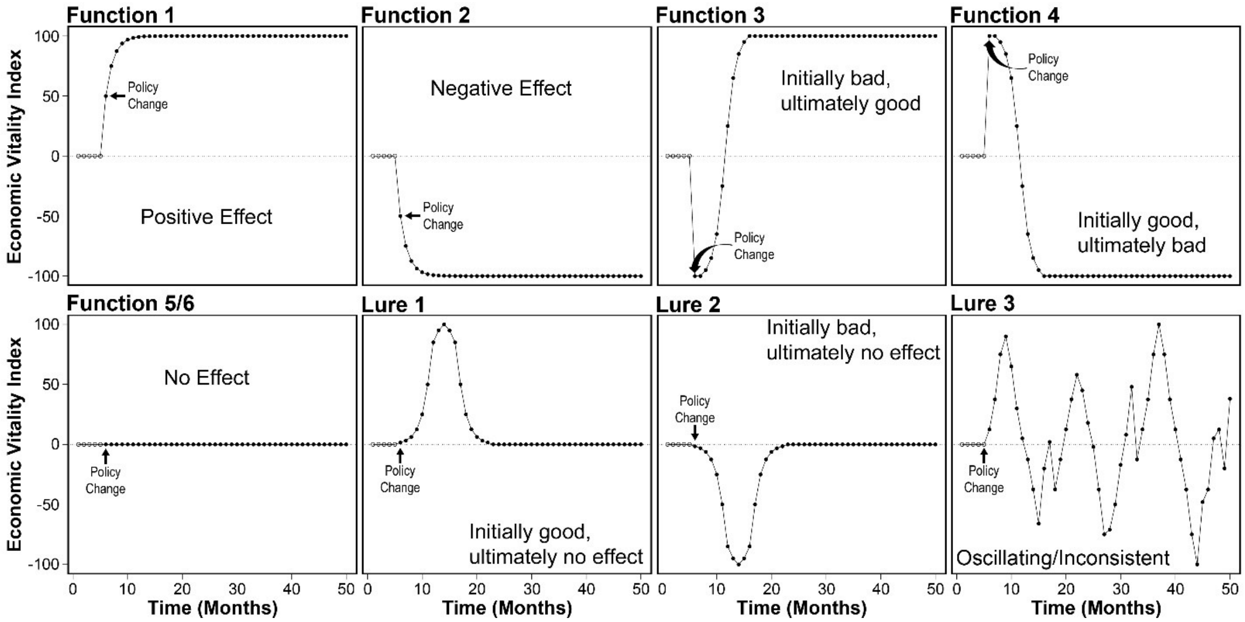

Figure 3. Function identification task payoff functions. Note: The above graphs were included as choices in the Function identification task (Lure 3 was only included in Study 1 and removed for subsequent studies). The first 5 trials in every graph represent an input of 'off' (white dots) before 45 inputs of 'on' (black dots).

3.1.5 Individual Differences. We initially had hypotheses about possible relations between individual difference measures (dogmatism, need for cognition, need for cognitive closure) and performance on the task, particularly around motivated reasoning. Though we measured these for Study 1 and Study 2A, we found few reliable relations, so we stopped measuring them in future studies and do not report the results for concision.

3.2 Results

For some analyses we separated our analyses into two categories for causal and non-causal functions. The causal functions were Functions 1-4 that actually produce an effect and where one policy was better than another (e.g., policy A > policy B). The non-causal functions (Functions 5 and 6) had no impact regardless of which policy was chosen (e.g., policy A = policy B). For causal functions, a policy was called "preference-congruent" if the participant's preferred policy happened to be the optimal policy, and was called "preference-incongruent" if the participant's preferred policy happened to be the suboptimal policy. For the non-causal functions, there is no such thing as preference congruence or incongruence because neither version of the policy is better than the other.

3.2.1 Participants. We removed 9 participants from our sample for making two or fewer policy changes throughout the entire learning task, which we viewed as a lack of task engagement. In all, 41 participants submitted valid data for analysis. Randomization outcomes for the number of participants who experienced each function as preference-congruent versus incongruent are in Appendix D for all studies.

3.2.2 Choices in the Learning Task. In this section, we examined four different ways that participants might test the policies in biased ways.

3.2.2.1 Amount of Testing by Preference. For the four causal policies for each participant, we calculated the average percent of trials in which the policy was set to the subjects' preferred option (out of 4 causal policies × 140 trials = 560 observations). If participants were not biased and simply tried to figure out which policy option was better, then they would try their preferred and non-preferred policy options equally. However, we hypothesized that they would try their preferred policy options more frequently than their non-preferred policy options.

Using a one-sample *t*-test we found that participants were more likely to test their preferred policies (M = 67%; SD = 18%) compared to chance (50%), $t(40) = 6.00$, $p < .001$, $d = .94$. We did the same test for the two non-causal functions, and also found that participants were more likely to test the preferred policy (M = 73%; SD = 22%), $t(40) = 6.82$, $p < .001$, $d = 1.06$.

3.2.2.2 Number of Trials Until Testing by Preference (Figure 4). We hypothesized that another way that bias in testing could be measured is that participants would try to test preferred policy options *earlier* than non-preferred policy options. This would manifest in the following way. Suppose that a policy was randomly set to the non-preferred option at start. We hypothesized that after only a few trials participants would tend to switch it to the preferred option. In contrast, we hypothesized that if a policy was randomly set to the preferred option at start, that it would take longer for participants to switch it to the non-preferred option. (Note that a participant cannot learn anything about a given policy until a switch happens.) If a participant never tested a policy at any point during the learning task, that particular policy for that participant was omitted from analysis. See *Figure 4* for density plots for all studies.

Because time until testing is necessarily positive and was skewed, a generalized linear model with a gamma distribution and an inverse link function was used to predict when a policy was first tested by policy preference at start. A random intercept for subject and a random slope for policy preference at start was included in the model. Participants switched non-preferred policies to preferred earlier than they switched preferred to non-preferred ($\beta = -.30$, $SE = .05$, $p < .001$).

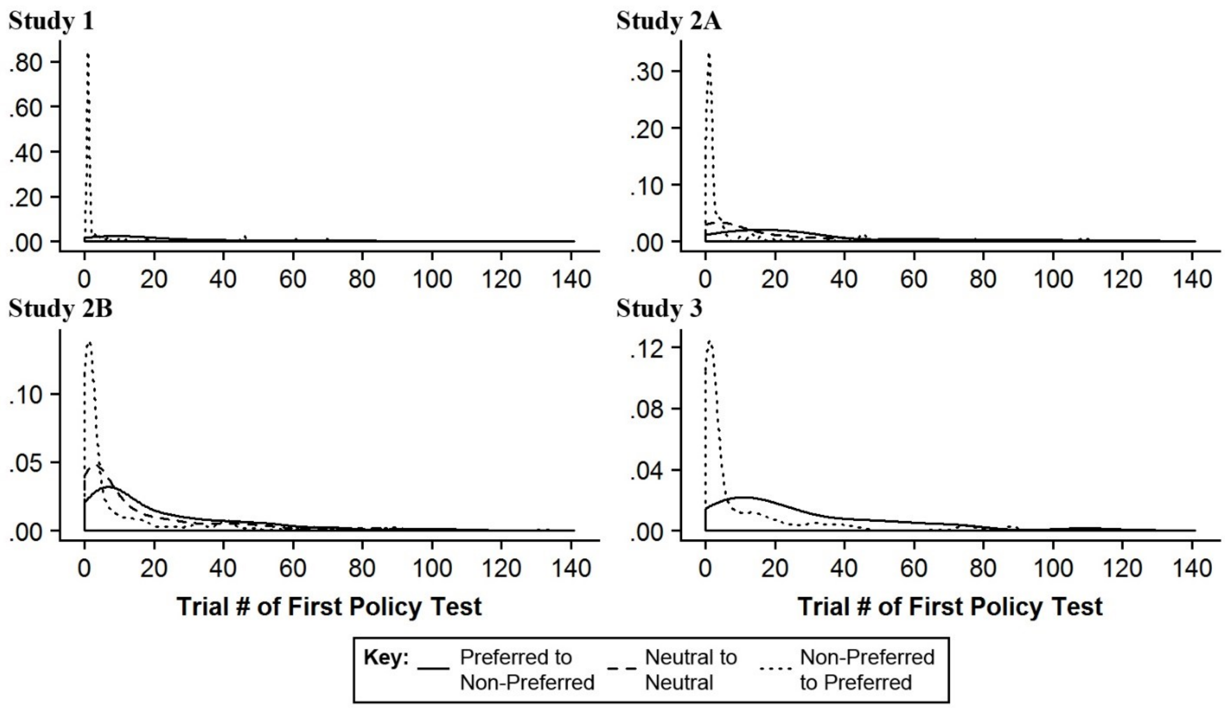

Figure 4. Density plots for number of trials until testing by preference. Note: The Y-axis is the density probability estimation for the trial number of first testing a policy. The X-axis is the trial range (1-140) for the learning task. Lines differentiate initial policy setting at the start of the learning task. All four studies show initial spike from policies initially set to a non-preferred policy being switched to a preferred policy early in the first few trials.

3.2.2.3 Never Testing Bias by Preference (Figure 5). Though most participants indeed tested both versions of each policy, on average across all participants and all policies 7.72% of policies, including both causal and non-causal, were never changed to test the version that was not selected at the start of the learning task. We hypothesized that participants might decide to leave policies that were initially set in their preferred state alone, never testing them, even though this would mean that they would not have an opportunity to determine which version was actually more effective, which would presumably lower their bonus for the task. See *Figure 5* for descriptive results of the percent of policies not tested from all studies.

To analyze this[3], we coded each participant as whether or not they failed to test at least one policy that was initially set to the preferred option, and whether or not they failed to test at least one policy that was initially set to the non-preferred option. We compared these using a McNemar's test of paired proportions. Participants were more likely to have not tested a policy at all, if the initial testing required switching a preferred policy to a non-preferred policy (29.27%)[4], versus if the initial testing required switching a non-preferred policy to a preferred policy (4.88%%), $\chi^2(1) = 6.75$, $p = .009$.

[3] We initially conducted a mixed effects logistic regression predicting whether a policy was ever tested by preference at start (yes/no) with a random intercept of subject. However, we ran into model convergence errors with this approach. This is likely due that fact that we are attempting to detect differences in rare events where large individual differences are present. In response, we simplified the approach.

[4] Note, these means are higher than the overall average (7.72%) and the averages in Figure 5 because the inferential statistics analyze whether a participant failed to test any of the policies initially set to preferred or non-preferred, whereas in Figure 5 we report the likelihood that an individual policy was not tested.

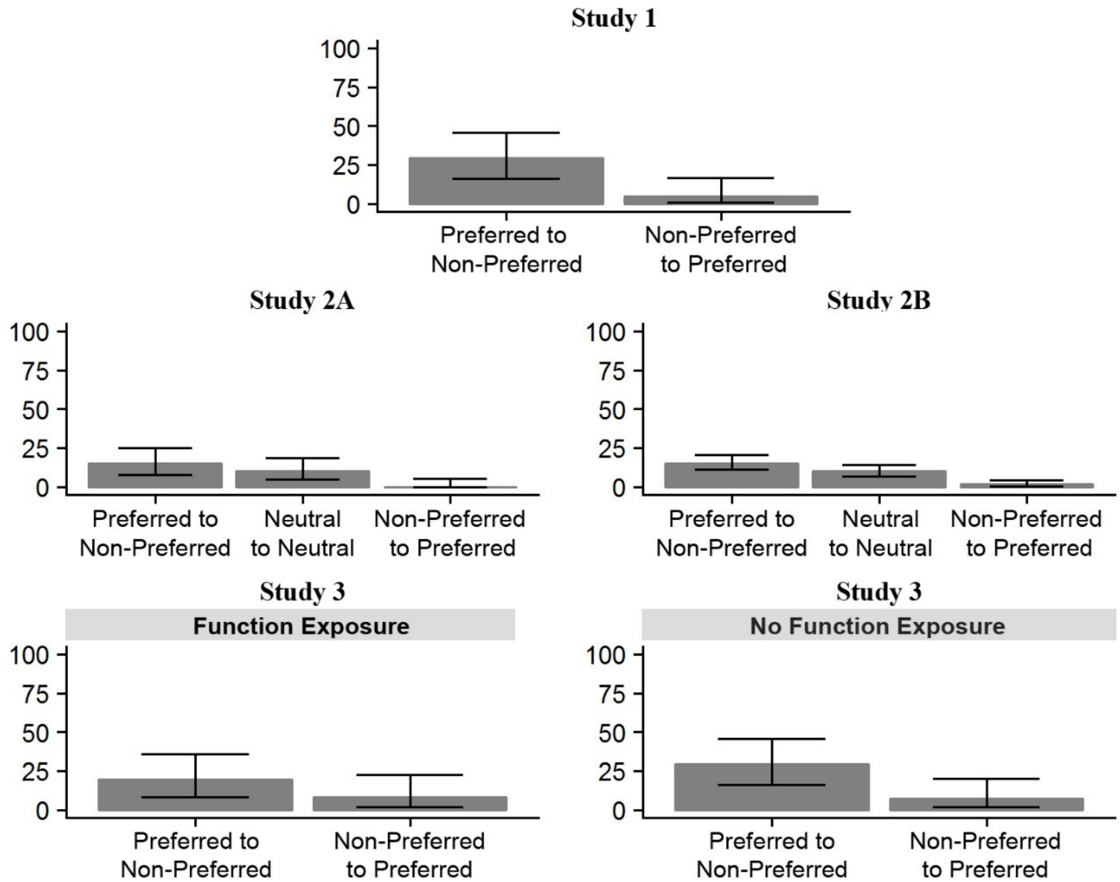

Figure 5. Percent of policies never tested by initial policy state. Note: Groups refer to the initial policy state and the alternative policy state upon entering task. Errors bars are 95% confidence interval.

3.2.2.4 Testing the Optimal Policy by Preference (Figure 6). Because this task is an explore-exploit task, not a pure explore task, it is rational for participants to test the versions of the policies that they actually believe to be better. Knowing from the finding that the participants tended to try their preferred policy options more than their non-preferred policy options, we hypothesized that another way that this bias would appear is in the frequency of testing the optimal policy. Specifically, we hypothesized that participants would be more likely to test the optimal version of the policies when the optimal version was also their preferred version (preference-congruent) compared to when the optimal version was their non-preferred version (preference-incongruent).

We tested preference-congruence and ambiguity as predictors of the likelihood that a policy was set to the optimal version on a given trial with a logistic mixed effects model. The model included predictors for preference-congruence, ambiguity, and their interaction, and had a by-subject random intercept with random slopes for all three predictors.[5]

Both predictors (and all other similar regressions in this manuscript) used effects coding with +.5 preference-congruent and -.5 for preference-incongruent, and +.5 for less ambiguous and -.5 for more ambiguous. This analysis only includes the causal functions, because non-causal functions cannot be categorized as preference-congruent or incongruent. See *Figure 6* for descriptive results from all studies.

Participants were more likely to select the optimal policy when it was also their preferred policy (preference-congruent, as opposed to preference-incongruent; $\beta = 3.47$, $SE = .52$, $p < .001$). Participants were also more likely to select the optimal policy if it was less ambiguous ($\beta = 2.15$, $SE = .54$, $p < .001$). There was no significant interaction between preference-congruence and ambiguity ($\beta = -.32$, $SE = .91$, $p = .726$).

[5] The model did not include a correlation parameter between the random slopes due to convergence difficulties, and dropping the correlation is a recommended way to improve convergence (Barr et al., 2013). A number of other models in this manuscript also drop the correlation parameter and are not specifically identified for concision.

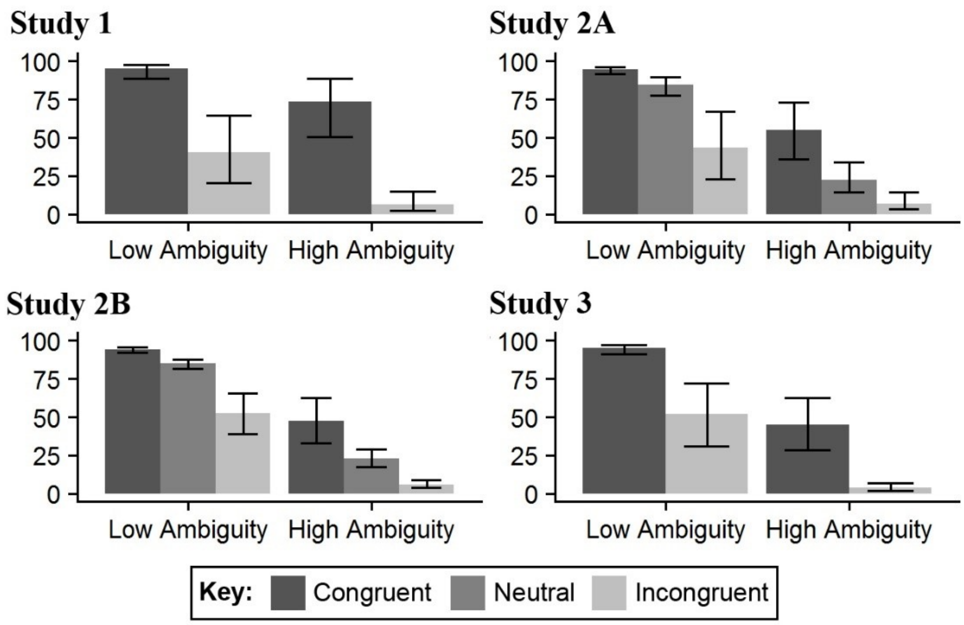

Figure 6. Percentage of trials set to optimal policy by preference-congruence and ambiguity. Note: Error bars represent 95% confidence interval. The wider confidence intervals for instances of low-ambiguity-congruent and high-ambiguity-incongruent are due to bimodal distributions.

3.2.2.5 Summary of Choices During the Learning Task. The previous analyses demonstrated that participants tested their preferred policy options more frequently, earlier, and were less likely to never test their preferred policy options, compared to their non-preferred policy options. These biases also meant that they were less successful at using the optimal policy when the optimal policy was preference-incongruent. The next section focuses on participants' judgments about the policies.

3.2.3 Judgments of Policy Efficacy After the Learning Task.

3.2.3.1 Causal Functions (Figure 7). We hypothesized that when participants judged the efficacy of each policy after the 140 learning trials, that they would tend to believe that the option that they preferred was more effective. We tested this by comparing the accuracy of their judgments about preference-congruent policies (policies for which their preferred option happened to be optimal) versus preference-incongruent policies (policies for which their preferred option happened to be sub-optimal). The dependent variable was the error in the policy assessments. This is measured by taking the absolute value of the difference between the slider position from the ideal slider position. For example, if Policy B is in fact better (which corresponds to +5), and a participant sets the slider to +2, they are 3 points away from the correct answer. See *Figure 7* for descriptive results from all studies; note that future studies use a different dependent measure of accuracy.

Given that we randomized whether a preferred policy was optimal or not, there wasn't necessarily one congruent and one incongruent policy for every function type (see Appendix D). Participants could have up to four preference-congruent policies or as few as zero. This means there were repeated measures for some users (e.g., two preference-congruent with the high ambiguity functions), only between group measures for some, and an absence of measurement for other participants (e.g., no preference-congruence with the high ambiguity functions). The below analysis was conducted at the user-level; when multiple measurements were present, these judgments were averaged. We also used non-parametric tests due to violations of normality.[6]

As expected, a Wilcox Rank Sum found that participants' judgment error was lower in the preference-congruent condition, when their preferred policies happened to be optimal (median = 3), than in the preference-incongruent condition (median = 6), $U = 354.50$, $p < .001$, $r = .440$, 95% CI = .249 – . 620). Additionally, participants were more accurate for the low ambiguity functions (median = 2) than for the high ambiguity functions (median = 5), $W = 94.50$, $p < .001$, $r = .49$, 95% CI = .32 – .65.

[6] Because of the non-normality of the rating scale, we used nonparametric tests and did not test for an interaction between preference-congruence and ambiguity. In subsequent studies we changed the rating scale and tested for an interaction.

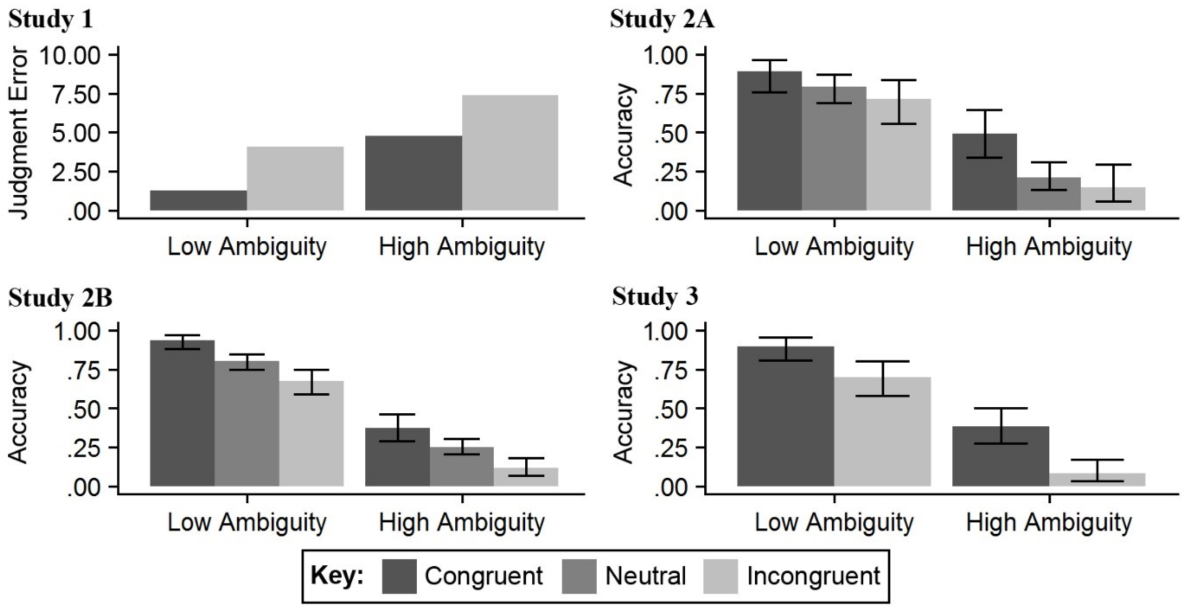

Figure 7. Accuracy of judgments of policy efficacy by congruence and ambiguity for causal functions. Note: Study 1 is a measure of error in judgment, whereas Studies 2A, 2B, & 3 are accuracy percentages. Error bars represent 95% confidence from a binomial test for each subgroup and did not account for repeated measures. There are no error bars for Study 1 because the data presented in this figure were categorized before conducting analysis.

3.2.3.2 Non-Causal Functions (Figure 8). We also examined how accurately participants assessed the non-causal functions and whether participants tended to select their preferred policy option as being better, despite neither policy option being better. To do this, participants' judgments were coded such that 0 represented an assessment that the preferred option produced a much better outcome than the non-preferred option (which was incorrect), 5 represented a correct assessment that there is no difference between the two options, and 10 represented an assessment that the non-preferred option produced a much better outcome (which was also incorrect). In *Figure 8*, instead of using the 11-point scale, we plot the three groups <5, 5, and >5 for consistency with subsequent studies. Participants very rarely concluded correctly that there was no difference, and usually concluded that their preferred policy option was better.

To determine if participants were more likely to assess their preferred policy as being better, we took the average of the scores for the two non-causal functions. A one-sample Wilcoxon signed-rank test against 5 confirmed that the judgments were biased towards the preferred policy (median = 4), $W = 42.50$, $p < .001$, $r = .64$.

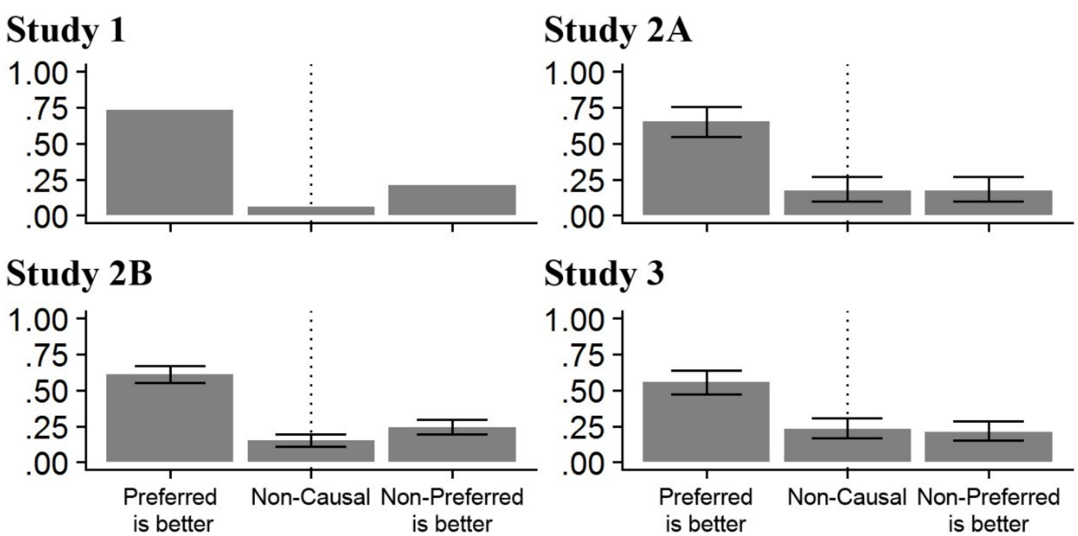

Figure 8. Judgments of policy efficacy after the learning task for non-causal functions. Note: Y-axis is percentage of judgments. The dotted line represents correct judgment. Study 1 data was collapsed into three bins for visual congruence with subsequent studies but not for analysis. The dotted line represents optimal judgment. Policies for which a participant held a neutral preference prior to the study were omitted.

3.2.4 Function Identification. At the end of the study participants were asked to match each of the six policies to a figure that represented different policy functions. We analyzed whether participants were able to accurately identify the mathematical function for each policy.

3.2.4.1 Causal Functions (Table 1). A mixed effects logistic regression analysis was conducted to test for differences in the ability to correctly choose the graph that represented functions by preference-congruence, ambiguity, and their interaction. The model used by-subject random intercept and a random slopes for all three predictors. Participants' accuracy at function identification did not differ based upon congruence (β = .02, SE = .45, p = .973). However, participants were more likely to correctly identify a function if it was less ambiguous (β = 1.48, SE = .46, p = .001). No interaction between congruence and ambiguity was found (β = 1.15, SE = .95, p = .229). In general the accuracy was fairly low, which is expected given that participants were not aware that they would have to do this task during the learning trials, they did not know the set of possible functions in advance, and furthermore, it is an unusual task; people rarely have to interpret graphs such as these in other settings.

Table 3.1. Accuracy of Function Identification for Causal Policies

			Preference			
Study	*Ambiguity*	*Function Exposure*	*Congruent*	*Neutral*	*Incongruent*	*Total*
1	High	No	.08	–	.13	.11
	Low	No	.39	–	.27	.33
2A	High	No	.24	.13	.10	.15
	Low	No	.33	.27	.27	.28
2B	High	No	.21	.15	.14	.16
	Low	No	.34	.21	.22	.25
3	High	No	.20	–	.05	.13
	Low	No	.27	–	.18	.22
	High	Yes	.14	–	.14	.14
	Low	Yes	.44	–	.29	.37

Note: Study 1 chance = 12.50%. Study 2A, 2B, & 3 chance = 14.29%.

3.2.4.2 Non-Causal Functions. Table 2 shows the mean accuracy of correctly identifying that the non-causal policies were non-causal. Participants were very rarely accurate, only 2% of the time; chance performance given the 8 graphs was 12.50%.

Table 3.2. Accuracy of Function Identification for Non-Causal Policies

		Has preference		
Study	*Exposed to Mechanism*	*Yes*	*No*	*Total*
1	No	.02	–	.02
2A	No	.03	.11	.07
2B	No	.06	.18	.12
3	No	.11	–	
3	Yes	.15	–	

Note: Study 1 chance = 12.50%. Study 2A, 2B, & 3 chance = 14.29%. "Exposed to Mechanism" refers to the function exposure manipulation used in Study 3.

3.2.5 Relations Between Choices During the Learning Task and Judgments of Policy Efficacy. We sought to examine relations between choices during learning and judgments afterwards. Though there are some relations, they are not especially reliable, and also are not directly related to questions around motivated reasoning. Thus, we report these findings in Appendix E.

3.3 Study 1 Discussion

In Study 1 we found that participants' testing behavior and policy assessments were greatly influenced by their strong policy preferences. Both participants' choices in the task and their learning outcomes provide evidence for motivated reasoning. Participants tended to test the preferred option of the policy more overall than the non-preferred option, and to test the preferred option earlier by switching the non-preferred option to preferred earlier than the reverse. In instances in which participants did not test a policy at all, the policy tended to already be set to the preferred policy. During the testing phase, participants more frequently tested the optimal policy if it was also the preferred policy. At the end of the learning task, participants were more likely to correctly assess the policies (to correctly determine which version of the policy is better) when they were preference-congruent (when the participant preferred the option that happened to be better).

Given the converging evidence that strong preferences can alter behavior and lead to biased conclusions, the next study investigated whether it is better to be open-minded (have neutral beliefs) than strong beliefs when learning cause-effect relationships.

4.0 Study 2

Study 2 extended the approach of Study 1 by comparing policies for which participants did not have prior preferences versus policies for which participants had prior preferences. We hypothesized that participants may be more accurate at learning about policies when they do not have strong preferences about them as opposed to when they do have strong preferences; their preferences may bias their ability to learn about the policy if they just assume that one version of the policy is better and fail to sufficiently test it.

Consider participants' choices during the learning phase. Study 1 showed that participants tended to choose their preferred policy options more often than their non-preferred policy options. However, this tendency could impede causal learning for both preference-congruent and preference-incongruent policies, because participants tended to mainly select their preferred policy option rather than switch between the two options; switching is necessary to test which option is better. In contrast, if a participant has no preferences, the lack of a bias could lead to more accurate learning.

4.1 Method

Study 2 was very similar to Study 1 except for the following changes. First, instead of only selecting policies for which participants had strong preferences, three policies with strong preferences and three policies with neutral preferences were selected for each participant. Policies with neutral preferences were defined as having policy assessments (the average of the two ratings) between 3 and 5 on the 7-point scale. We first selected policies with ratings of exactly 4 (the middle of the scale), but if a participant did not have enough policies of exactly 4, then policies with ratings of 3.5 and 4.5 were chosen next, followed by policies with ratings of 3 and 5. In cases where there were not enough participant ratings that fell into the "neutral preference" or "strong preference" bins, it was possible for a participant to have more neutral policies than strong preference policies (or vice versa). Most participants in had a perfect balance between strong and neutral policies (MTurk sample: 99%; Intro. Psych sample: 96%).

Second, the policy assessment judgments that participants made during the learning task and right after Trial 140 were changed to a 3-point scale ("Policy A is better"; "No Effect/Uncertain"; "Policy B is better.") instead of the 11-point scale. This was because in Study 1 participants mainly used the extremes of the scale resulting in a non-normal distribution.

Third, during the function identification task, we removed the 'oscillating' lure plot because so many participants chose it and success rates were very low. We worried that participants chose it because it looked like the noise function we were using rather than because it looked like any of the causal functions.

Fourth, we collected two samples for Study 2; MTurk (Study 2A) and undergraduate introduction to psychology students (Study 2B). Because of the similarity of results, we report the results side-by-side.

4.1.1 Participants. In the MTurk sample there were 102 participants. Participants were paid $6.50 for participation (which amounted to approximately $8-10/hr) with an opportunity to be awarded up to $3.00 in bonuses contingent upon performance. We removed 12 participants for making fewer than two policy changes throughout the entire learning task. Additionally, we removed one participant who appeared to have sped through the study after examining completion times and another because they only partially completed the full study. In all, 88 participants were included in our analysis.

In the introduction to psychology (Intro. Psych.) participant pool, there were 385 participants. Participants received course credit for participation. We removed 101 participants for making fewer than two policy changes throughout the entire learning task. The higher rate of disengagement compared to the MTurk sample could be due to the lack of payment and bonus. In all, 283 participants were included in our analysis.

4.2 Results

4.2.1 Choices in the Learning Task.

4.2.1.1 Amount of Testing by Preference.

4.2.1.1.1 Causal Functions. Participants tended to test their preferred policies more often than their non-preferred policies for both the MTurk (M = 66%; SD = 29%, $t(87) = 5.40$, $p < .001$, $d = .58$) and Intro. Psych. samples (M = 64%; SD = 30%, $t(279) = 7.99$, $p < .001$, $d = .48$).

4.2.1.1.2 Non-Causal Functions. Participants were much more likely to test the preferred policy than the non-preferred policy for both the MTurk (M = 74%; SD = 27%, $t(70)$=7.53, $p < .001$, $d = .89$) and Intro. Psych. samples (M = 68%; SD = 30%, $t(225) = 9.25$, $p < .001$, $d = .62$) .

4.2.1.2 Number of Trials Until Testing by Preference (Figure 4). We first replicated our finding from Study 1, excluding neutral policies; participants switched non-preferred policies to preferred earlier than they switched preferred to non-preferred (MTurk: $\beta = -.31$, $SE = .05$, $p < .001$; Intro. Psych.: $\beta = -.21$, $SE = .03$, $p < .001$).

We then compared policies for which participants had neutral preferences versus policies for which they had strong preferences and tested whether they tested neutral policies (switching from one neutral option at start to the other) earlier or later than policies for which they had strong preferences (which includes both switching from preferred at start to non-preferred or non-preferred at start to preferred). We again used a generalized linear model with a gamma distribution and an inverse link function to predict when a policy was first switched. The model included a by-subject random intercept with a random slope of preference strength (strong vs.

neutral). We did not find a difference between strong and neutral preferences (MTurk: $\beta = -.01$, $SE = .04$, $p = .845$; Intro. Psych.: $\beta = .01$, $SE = .02$, $p = .61$).

4.2.1.3 Never Testing Bias by Preference (Figure 5). Though most participants indeed tested both versions of each policy, on average across all participants and all policies some 4.55% of the MTurk sample and 5.77% of the Intro. Psych. sample were never changed. Participants were more likely to have not tested a policy at all, if the initial testing required switching a preferred policy to a non-preferred policy (18% for MTurk, 14% for Intro. Psych) versus if the initial testing required switching a non-preferred policy to a preferred policy (0% for MTurk, 1.40% for Intro. Psych), and these proportions were significantly different (MTurk: McNemar's $\chi^2(1) = 8.10$, $p = .004$; Intro. Psych.: McNemar's $\chi^2(1) = 23.31$, $p < .001$).

In the Intro. Psych. sample, participants were more likely to have not tested a policy at all, if the initial testing required switching a preferred policy to a non-preferred policy (15%), versus if the initial testing required switching a neutral policy to a competing neutral policy (10%), McNemar's $\chi^2(1) = 4.69$, $p = .030$. However in the MTurk sample, we did not find a difference in never testing behavior between policies that were initially set to either preferred (14%) or neutral policies (9%), McNemar's $\chi^2(1) = 1.13$, $p = .289$.

We found that participants were more likely to have not tested a policy at all, if the initial testing required switching a neutral policy to a competing neutral policy (MTurk: 13%; Intro. Psych: 10%) versus if the initial testing required switching a non-preferred policy to a preferred policy (MTurk: 0.00%; Intro. Psych: 2%; MTurk: McNemar's $\chi^2(1) = 7.11$, $p = .008$; Intro. Psych.: McNemar's $\chi^2(1) = 18.38$, $p < .001$).

4.2.1.4 Testing the Optimal Policy by Preference (Figure 6). The same logistic regression from Study 1 produced similar findings. Participants were more likely to select the optimal policy when it was preference-congruent, as opposed to preference-incongruent, (MTurk: β = 2.94, SE = .39, p < .001; Intro Psych.: β = 2.72, SE = .27, p < .001). Participants were also more likely to select the optimal policy if it was less ambiguous (MTurk: β = 2.52, SE = .44, p < .001; Intro Psych.: β = 3.14, SE = .22, p < .001). There was not a significant interaction (MTurk: β = -.57, SE = .53, p = .287; Intro Psych.: β = .19, SE = .37, p = .607).

We then tested whether participants were more likely to test the optimal version of a policy if they had neutral preferences about the policy as opposed to having strong preferences (preference congruent and incongruent). We used logistic mixed effects model predicting the likelihood that a policy was set to the optimal version on a given trial by preference strength (strong vs. neutral) and ambiguity, and their interaction. The model included a by-subject random intercept with random slopes for all three predictors. Similar to the previous finding, participants were more likely to select the optimal policy if it was less ambiguous (MTurk: β = 2.70, SE = .35, p < .001; Intro Psych.: β = 3.03, SE = .17, p < .001). However, there was not a significant difference between strong versus neutral preferences (MTurk: β = .01, SE = .27, p = .967; Intro Psych.: β = .24, SE = .17, p = .153). There was also no significant interaction (MTurk: β = .46, SE = .45, p = .306; Intro Psych.: β = -.19, SE = .28, p = .503).

In the analysis of the percent of trials set to the optimal policy during learning (*Figure 6*), there does appear to be some asymmetries between strong versus neutral preferences for Study 2. For the low ambiguity functions, participants were almost as good at using the optimal policy versions for neutral policies as they were for congruent policies, suggesting good learning for neutral policies. In contrast, they were considerably worse for incongruent policies. However, for high ambiguity policies for which participants were much worse overall, the neutral policies were more in the middle between congruent and incongruent policies. If anything, the neutral policies were closer to the incongruent than the congruent policies.

Note also, that in the percent of trials set to the optimal policy (*Figure 6*), the error bars for the low-ambiguity preference-incongruent policies, and for the high-ambiguity preference-congruent policies were fairly large. We think what is happening is that when learning about the low-ambiguity policies, participants tend to be good, but when their preferences are incongruent, for about 50% of policies people still primarily tested their preferred (suboptimal version).

In contrast, when learning the high-ambiguity functions, in general learning was fairly poor and people tended to use the suboptimal policy. Stated another way, most people tended to learn the short-term influence of the high-ambiguity functions, and therefore use the suboptimal policy. However, when their preferences were congruent, about 50% of people still primarily tested their preferred option. It seems plausible that these participants did not realize that their preferred option was optimal in the long run, but instead just continued to use this option because they preferred it despite it being suboptimal in the short-term.

In summary, a plausible interpretation of these results is that the wide error bars in both the low-ambiguity preference-incongruent and the high-ambiguity preference-incongruent conditions are that for about 50% of policies participants either failed to learn and instead primarily tested their preferred option, or, they did learn the short-term influence, but still primarily tested their preferred option that they understood was suboptimal in the long-term.

Still, despite some apparent asymmetries, they did not show up as a main effect of preference-strength or an interaction between preference-strength and ambiguity. This is likely due to the high variability in these two (low-ambiguity incongruent, high-ambiguity congruent) conditions.

4.2.1.5 Summary of Choices During the Learning Task. The previous analyses replicated the results from Study 1: participants tested their preferred policy options more frequently, earlier, and were less likely to never test their preferred policy options, compared to their non-preferred policy options, and participants were less successful at using the optimal policy when the optimal policy was preference-incongruent.

However, whereas we had speculated that perhaps participants would be better at testing policies for which they had neutral preferences compared to policies for which they had strong preferences (an average of congruent and incongruent), we found few differences.

4.2.2 Judgments of Policy Efficacy After the Learning Task.

4.2.2.1 Causal Functions (Figure 7). Given that participants rarely selected "no effect" as their final judgment of a policy, we collapsed the responses from three levels into two (correct vs. incorrect) for ease of analysis. We first replicated our finding from Study 1, excluding neutral policies. A mixed effects logistic regression analysis was conducted to test for differences in the ability to correctly identify which policy option was better for economic output. The main effects and interaction between ambiguity and preference-congruence were included. In the MTurk sample, there was a by-subject random intercept and random slopes for all three predictors. For the Intro. Psych. sample the random slope for the interaction was dropped due to non-convergence.

Replicating Study 1, MTurk participants were less likely to correctly assess preference-incongruent policies than preference-congruent, (β = -1.68, *SE* = .55, p = .002). However, no significant effect of preference-congruence was found for the Intro Psych. sample (β = -2.31, *SE* = 1.42, p = .104). Participants were significantly worse at assessing policies with high ambiguity, compared to low ambiguity (MTurk: β = -2.85, *SE* = .74, p < .001; Intro. Psych.: β = -21.28, *SE* = 1.89, p < .001). There was no interaction (MTurk: β = .72, *SE* = .91, p = .430; Intro. Psych.: β

= -.33, *SE* = 2.33, *p* = .888).

Next, we tested whether participants were better at assessing policies for which they had strong preferences (preference congruent or incongruent) versus no preferences. A mixed effects logistic regression analysis was conducted with preference-strength (strong vs. weak) and ambiguity as predictors, and a by-subject random intercept with random slopes for all three predictors.

There was no significant difference in correctly assessing policies when participants did versus did not have a preference (MTurk: β = .44, *SE* = .32, *p* = .168; Intro. Psych.: β = -.09, *SE* = .19, *p* = .621). Participants were significantly worse at assessing policies with high ambiguity compared to low ambiguity (MTurk: β = -3.24, *SE* = .55, *p* < .001; Intro. Psych.: β = -3.80, *SE* = .39, *p* < .001). There was no interaction between preference-strength and ambiguity (MTurk: β = -.70, *SE* = .62, *p* = .259; Intro. Psych.: β = .05, *SE* = .37, *p* = .898).

4.2.2.2 Non-Causal Function (Figure 8). We first replicated our results from Study 1 demonstrating that participants were more likely to assess their preferred policy as being better, despite no actual difference. To do this, we used the subset of policies for which participants had a preference, and for which they failed to correctly assess the policy as non-causal (which was relatively rare). A logistic mixed effects model was run predicting judgment bias (1 = assessing preferred policy as being better; 0 = assessing non-preferred policy as being better) with only a by-subject random intercept to account for repeated measures (each participant had between 0-2 observations). When participants had an initial preference for one policy version over another, after testing it they were still more likely to view the preferred option as the better policy (MTurk: *M* = .79; *CI* = .68 – .82; β = 1.34, *SE* = .29, *p* < .001; Intro. Psych.: *M* = .73; *CI* = .65 – .80; β = .99, *SE* = .20, *p* < .001).

We also tested whether participants would be more likely to make accurate judgments of policy efficacy for the non-causal functions if they held neutral preferences versus strong preferences. We conducted a logistic mixed effects regression with preference strength (strong vs. weak) predicting accuracy (correct vs. incorrect) with a by-subject random intercept and a random slope for preference strength. Though participants were a bit more accurate when they held neutral preferences (MTurk: 25.84%; Intro. Psych: 27.50%) than strong preferences (MTurk: 17.24%; Intro. Psych.: 15.03%), the difference was not significant (MTurk: $\beta = .93$, $SE = 2.61$, $p = .721$; Intro. Psych.: $\beta = -1.00$, $SE = .86$, $p = .244$).

4.2.3 Function Identification.

4.2.3.1 Causal Functions (Table 1). We first replicated our finding from Study 1, excluding policies for which participants had no preferences. For the MTurk sample random slopes were included for all three predictors, but for the Intro. Psych. sample the random slope for the interaction was dropped due to non-convergence. Participants were significantly better at function identification with low ambiguity than high ambiguity in the Intro. Psych sample ($\beta = .66$, $SE = .23$, $p = .004$); this finding was marginal for the MTurk sample ($\beta = .89$, $SE = .46$, $p = .051$). Participants were better at function identification for policies that were preference-congruent than incongruent for Intro. Psych. ($\beta = .61$, $SE = .23$, $p < .001$), though this was marginal for MTurk ($\beta = .80$, $SE = .44$, $p = .071$). No interaction was found (MTurk: $\beta = -.81$, $SE = .86$, $p = .347$; Intro. Psych: $\beta = .14$, $SE = .45$, $p = .760$).

We also tested whether participants were more likely to correctly identify functions if they held neutral preferences versus strong preferences. We used a mixed effects logistic regression with the predictors preference strength (strong vs. neutral) and ambiguity. The model included a by-subject random intercept with random slopes for preference-congruence and ambiguity but not the interaction due to non-convergence. Participants were significantly better at function identification with low ambiguity (MTurk: $\beta = .82$, $SE = .30$, $p = .006$; Intro. Psych: $\beta = .57$, $SE = .17$, $p < .001$). There was not a significant effect of preference strength (MTurk: $\beta = -.26$, $SE = .28$, $p = .351$; Intro. Psych: $\beta = -.28$, $SE = .17$, $p = .109$). There was no interaction (MTurk: $\beta = .18$, $SE = .55$, $p = .742$; Intro. Psych: $\beta = -.25$, $SE = .32$, $p = .436$).

4.2.3.2 Non-Causal Functions (Table 2). We conducted a mixed effects logistic regression to test for differences in non-causal function identification by preference-strength (strong vs. weak). A by-subject random intercept with a random slope was used. Though the accuracy of function identification was a bit higher with neutral preferences than strong preferences, the difference was not significant in the MTurk sample ($\beta = 3.59$, $SE = 6.76$, $p = .595$). However, participants in the Intro. Psych sample were more likely to correctly identify the non-causal functions if they did not have preferences ($\beta = 9.63$, $SE = 1.67$, $p < .001$).

4.3 Study 2A & 2B Discussion

Study 2 largely replicated the findings in Study 1. In addition, Study 2 found that when participants had neutral preferences, their performance was in the middle between preference-congruence and preference-incongruence such that there was no difference in performance between having strong and weak preferences in most all cases. Stated another way, the benefits of preference congruence (when the participant's preference happens to be right) and the costs of preference incongruence (when the participant's preference happens to be wrong) roughly cancel out. There were some hints that the neutral condition might not be right in the middle, or might actually flip for low versus high ambiguity functions, but these were not statistically significant. In sum, this study reconfirms that preferences have a strong influence on causal learning and judgments, however, it does not provide evidence that having preferences, on the whole, leads to better or worse learning and judgments compared to not having preferences.

5.0 Study 3

One of the central challenges participants faced in the previous studies is that they did not know in advance about the possible functions for how the policies worked. For example, if a participant assumed that the policies worked immediately, they might make a change to one policy on one trial, and then make a change to another policy on the subsequent trial, and because it actually takes a number of trials for the policies to work, their causal attributions could be wrong. For another example, a participant might not even consider the possibility that a policy could have short-term costs but long-term benefits, and upon noticing a short-term cost they might switch away from that policy without investigating whether there are long-term benefits.

On the one hand, in many real-world situations decision makers don't know the possible functions, or might only have rough guesses about the length of time it might take for a policy to produce its full impact, or whether it is possible for a policy to have different short versus long-term influences. On the other hand, in some situations more informed decision makers might have hypotheses about possible functional forms (for example, see the quotes at the beginning of the introduction).

The goal of Study 3 was to investigate whether being more informed about the potential types of influences ('function exposure') would improve learning, which would appear as a main effect of function exposure. Furthermore, we hypothesized that if participants are exposed to the possible functional forms in advance, it might reduce the biases seen due to preference, which would appear as an interaction between preference (congruent vs. incongruent) and function exposure.

Previous studies using the 'melioration' paradigm have tested a couple ways to improve performance on the task, with various success. It has been found that giving participants a perceptual cue that corresponds with the underlying state of the payoff function (how many times the optimal choice has been chosen in the past 10 trials) can improve learning (Gureckis & Love, 2009; Herrnstein et al., 1993; Otto, Gureckis, Markman, & Love, 2009; Stillwell & Tunney, 2009). However, this approach would have been very confusing with 6 causes instead of just 1, and furthermore, we wanted to test whether a more explicit form of knowledge of the possible functions could matter. Part of the reason why was that unlike the previous studies on the melioration, not only did we study the percent of optimal choices, but we also studied participants' explicit beliefs about which policy option was better and their beliefs about the functional form of the payoff. Herrnstein et al., (1993) found that giving participants explicit instructions improved performance. However, even these instructions did not clearly state that the different options could have different the short-term versus long-term consequences. In the current study, we explicitly told some participants about the possibility of such temporal tradeoffs.

5.1 Method

5.1.1 Participants. One-hundred participants were recruited via MTurk. They were paid $5.50 for participation (which amounted to approximately $8-10/hr) with an opportunity to be awarded up to $3.00 in bonuses contingent upon performance.

5.1.2 Design. Study 3 was very similar to Study 1 with the following changes. First, half of the participants were exposed to the possible functional forms of the policies before starting the task, and the other half were not (like in Studies 1 and 2). Second, similar to Study 1, Study 3 focused on learning in the context of strong preferences, so only policies with strong prior preferences were selected. However, if a participant did not have six policies with strong prior preferences the computer would choose the "next most-extreme" to be included in the task. In these cases, the policies that did not meet our criteria to be categorized as "strong prior preferences" would be omitted from analysis (but not the participant altogether). Third, as an improvement to Study 1, we counterbalanced the causal functions such one of the low-ambiguity functions was preference-congruent and one was preference-incongruent and same for the two high-ambiguity functions.

5.1.3 Function Exposure Task. In the function exposure task participants read the following.

Training intervention instruction text:

> "In the following task, you will pretend to be the elected the leader of a large industrialized country, and you will be responsible for making important decisions about economic policies. But before doing so, **we want you to reflect on the possible ways that your changes to economic policies might influence the economy**.
>
> A change to a policy might:
> - Have no influence on the economy.
> - Have a positive or negative influence, but it might also take some amount of time for these positive or negative influences to appear.
> - Initially have a positive influence, but eventually have a negative influence, or vice versa.
> - Have a temporary positive or negative influence, but no long-term influence.
>
> Thinking about the possible ways that your policy changes might influence the economy will help you to determine which policies are best in order to maximize the economy's output."

Then participants were shown graphs of the 7 functions (5 that were present in learning task, and 2 lures; see *Figure 3*), and for each graph they had to match the function to text describing the function before moving on (*Figure 9*).

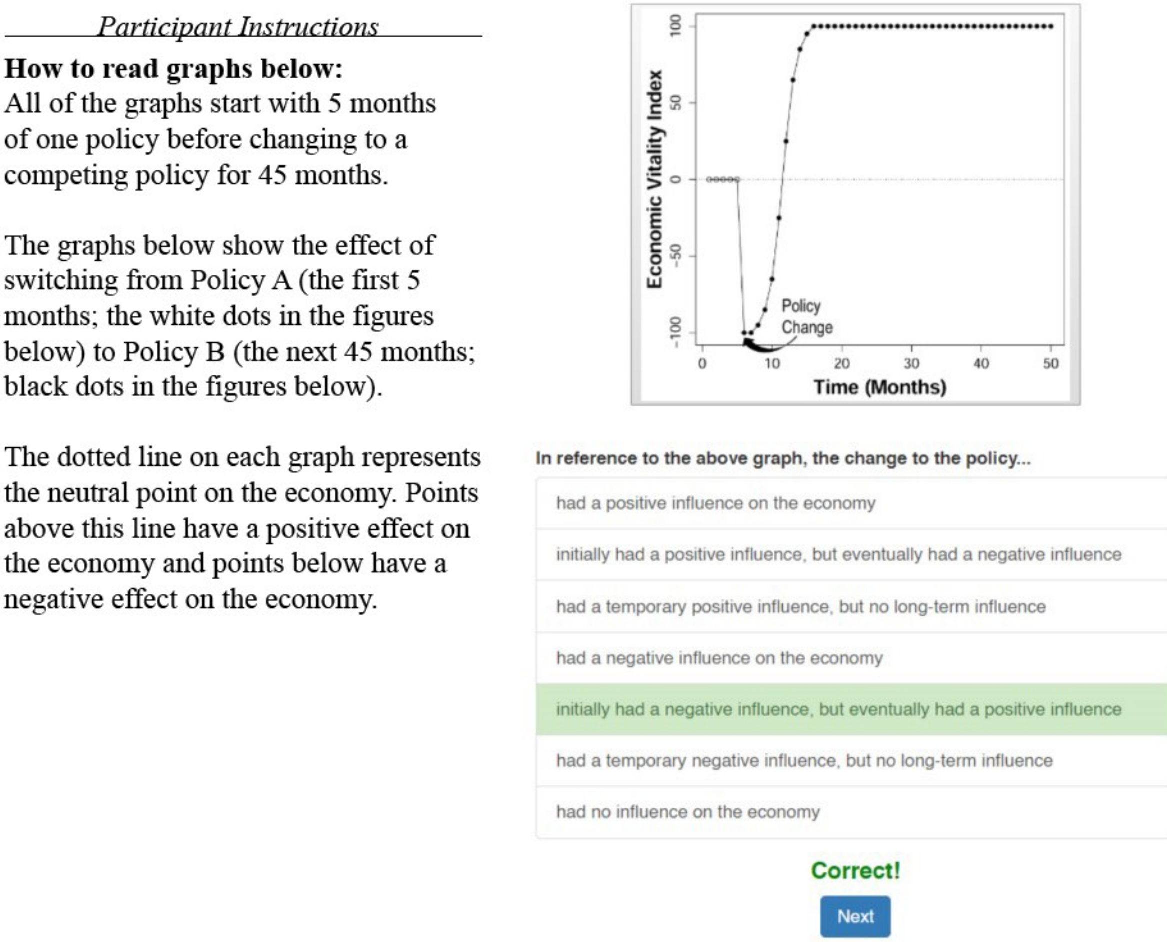

Figure 9. Function exposure test. Participants had to correctly choose the text that best descripted the function.

5.2 Results

In Study 3, we only investigated learning in the presence of strong prior preferences. However, as explained in the methods it was possible that for some of the policies that participants would hold moderate views. In the few cases in which participants did not have strong prior preferences for certain functions, these were omitted from the analysis. Sixty-eight participants had strong preferences for all six policies. Three participants had 5 preferred policies and 1 neutral policy. Seven participants had 4 preferred policies and 2 neutral policies. Seven participants had three or fewer policies with strong preferences, and these participants were dropped entirely from the study.

In addition, 14 participants were removed from analyses for making fewer than two policy changes throughout the entire learning task, and one was removed for not following directions. In all, 78 participants were included in analyses.

5.2.1 Choices in the Learning Task.

5.2.1.1 Amount of Testing by Preference.

5.2.1.1.1 Causal Functions. Participants tested their preferred version of the policies more than their non-preferred version, $M = 65\%$; $SD = 18\%$, $t(77) = 7.36$, $p < .001$, $d = .83$. There was no difference between those who received the function exposure ($M = 64\%$; $SD = 16\%$) and those who did not ($M = 65\%$; $SD = 19\%$), $t(75.84) = .18$, $p = .86$, $d = .04$.

5.2.1.1.2 Non-Causal Functions. For non-causal functions, participants also tested their preferred version more frequently than their non-preferred version (M = 68%; SD = 24%), $t(76) = 6.63$, $p < .001$, $d = .76$. And there were no differences between the participants who received the function exposure ($M = 66\%$; $SD = 24\%$) or not ($M = 71\%$; $SD = 24\%$), $t(72.25) = .90$, $p = .37$, $d = .20$.

5.2.1.2 Number of Trials Until Testing by Preference (Figure 4). A gamma mixed effects regression was conducted predicting time until testing by the interaction of function exposure condition and policy preference at start, with a by-subject random intercept and a random slope of preference at start. Replicating the prior studies, participants switched non-preferred policies to preferred earlier than they switched preferred policies to non-preferred ($\beta = -.30$, $SE = .05$, $p < .001$). There was no effect of function exposure ($\beta = .03$, $SE = .05$, $p = .584$), nor an interaction ($\beta = .08$, $SE = .09$, $p = .339$).

5.2.1.3 Never Testing Bias by Preference (Figure 5). On average across all participants and all policies 10% were never changed (function exposure condition: 8%; no exposure condition: 12%). Within the function exposure group, participants did not differ in never testing behavior contingent on if the initial testing required switching a preferred policy to a non-preferred policy (19%) versus the reverse (8%), $\chi^2(1) = 2.25$, $p = .134$. However within the no function exposure group, participants were more likely to have not tested a policy at all, if the initial testing required switching a preferred policy to a non-preferred policy (30%) versus the reverse (8%), $\chi^2(1) = 5.82$, $p = .016$.

5.2.1.4 Testing the Optimal Policy by Preference (Figure 6). We used a logistic mixed effects model to predict the likelihood that a policy was set to the optimal version on a given trial with the predictors preference-congruence, ambiguity, function exposure, and their interactions. The model had a by-subject random intercept with random slopes for preference-congruence and ambiguity. (It did not have a random slope for the interaction between these two as the model could not converge given that there was only one observation per cell.)

Confirming findings from Study 1 and Study 2, participants were more likely to select the optimal policy for preference-congruent as opposed to preference-incongruent policies ($\beta = 3.16$, $SE = .44$, $p < .001$) and for less ambiguous policies ($\beta = 3.45$, $SE = .39$, $p < .001$). We did not find a significant effect of function exposure ($\beta = .05$, $SE = .37$, $p = .892$).

There were also significant interactions (*Figure 10*). Though we report these below, it is obvious from the regression weights that the main effects are considerably stronger. First, there was a significant interaction between congruence and ambiguity ($\beta = .29$, $SE = .08$, $p < .001$); this means that the effect of preference congruence was somewhat smaller in the low ambiguity than in the high ambiguity condition. There was not a significant interaction between congruence and function exposure (β = -.59, *SE* = .83, *p* = .477) or ambiguity and function exposure (β = -.29, *SE* = .73, *p* = .691). However, a significant three-way interaction between congruence, ambiguity, and function exposure was found (β = 1.14, *SE* = .17, *p* < .001). In the low ambiguity conditions, the no-treatment group actually performed slightly better than the treatment group. However, in the high ambiguity preference-congruent condition, there was no effect of function exposure, and in the high ambiguity preference-incongruent condition, there was a slight benefit of function exposure. To summarize, there may be a slight benefit of function exposure in the hardest case (high ambiguity, preference-incongruent), but there is no overall benefit of function exposure.

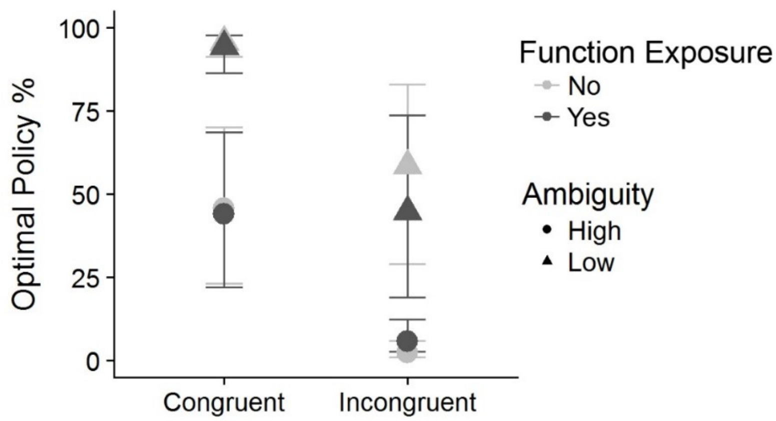

Figure 10. Three-way interaction effect for optimal policy percentage.

5.2.2 Judgments of Policy Efficacy After the Learning Task.

5.2.2.1 Causal Functions (Figure 7). A near-identical approach was taken here as that of Study 2A and 2B, the only exception being that function exposure condition (between-subjects) and its interactions with the other predictors were included as predictors. The model included a by-subject random intercept, but no random slopes.[7]

First, participants were less likely to correctly assess policies if they were preference-incongruent than congruent (β = -1.66, SE = .34, p < .001). Second, participants were significantly worse at assessing policies with high ambiguity compared to low ambiguity (β = -3.01, SE = .34, p < .001). Third, and most relevant to Study 3, there was no effect of function exposure (β = -.30, SE = .34, p = .383); participants were about equally accurate in the function exposure condition (M = 45.77%) as in the no-exposure condition (M = 50.63%). There were also no significant two or three-way interactions.

5.2.2.2 Non-Causal Functions (Figure 8). We first replicated the finding that participants were more likely to assess their preferred policy as being better, despite there being no difference. We used the same approach as in Study 2, and for Study 3 only used the no-function-exposure group for comparability. Replicating prior results, we found that when participants had an *a priori* preference, after testing it they were still more likely to view it as the better policy (M = .73; CI = .61 – .83; β = .99, SE = .28, p < .001).

We next tested whether participants who were in the function exposure condition performed better on this task compared to those who were not. To test for this difference, we conducted a mixed effects logistic regression with function exposure condition predicting accuracy (correct vs. incorrect) with a by-subject random intercept. The mean accuracy in the function exposure condition (22.06%) and the no exposure condition (24.10%) were similar, and the effect of condition was not significant, β = .15, SE = .51, p = .775.

[7] Though preference congruence and ambiguity were within-subjects, there was only one observation per person per cell, and this was the maximal model that would converge here and for other similar models in Study 3.

5.2.3 Function Identification.

5.2.3.1 Causal Functions (Table 1). A mixed effects logistic regression analysis was used to predict the ability to correctly choose the graph that represented the function of each policy from preference-congruence, ambiguity, and function exposure condition. A by-subject random intercept was used with no random slopes. There was a positive effect of preference-congruence ($\beta = .68$, $SE = .20$, $p = .040$) and lower ambiguity ($\beta = 1.09$, $SE = .33$, $p = .001$). However, there was no main effect of function exposure ($\beta = .52$, $SE = .34$, $p = .120$), nor were there any significant two or three-way interactions.

5.2.3.2 Non-Causal Functions (Table 2). A mixed effects logistic regression analysis was used to predict the ability to correctly identify that the two non-causal policies per participant were non-causal based on condition. The model included a by-subject random intercept. Being exposed to the functions prior to learning did not the function identification ($\beta = .40$, $SE = .61$, $p = .513$).

5.3 Study 3 Discussion

Study 3 replicated many of the findings from the prior studies. The added intervention of being exposed to the possible functional forms for the policies was largely ineffective at improving performance.

6.0 General Discussion

In three studies we found evidence that people's testing behavior and learning outcomes were greatly influenced by their *a priori* preferences for some policy options over others (e.g., increasing border security funding vs. decreasing it), which could be viewed as a type of motivated reasoning. We identified four specific biased habits during testing, all of which could be viewed as different manifestations of positive testing (e.g., Klayman & Ha, 1987). First, participants tended to test the preferred option of the policy more overall than the non-preferred option. Second, they tested the preferred option earlier; they switched the non-preferred option to preferred earlier than the reverse. Third, in instances in which participants did not test a policy at all by switching it from one option to the other, the policy tended to already set to the preferred policy. Fourth, all of these habits lead participants to use the optimal policy more when the optimal policy was congruent with their preferences and less when it was incongruent.

After the initial testing phase, when assessing the policies, participants were more accurate when their preferences aligned with the actual efficacy of the policies. In the introduction we raised a number of additional hypotheses, which we address in the following below.

6.1 Ambiguity

As expected, participants were much worse at learning the high-ambiguity policies than the low-ambiguity ones. We had hypothesized that, in addition, the motivated reasoning effect would be magnified for the high-ambiguity policies because the ambiguity could license interpreting these policies in the ways that participants' preferred; however, we did not find evidence for this hypothesis. We have a couple of speculations why.

One possibility is that both the low-ambiguity functions and the high-ambiguity functions were viewed as fairly ambiguous. Indeed, the low-ambiguity functions themselves took multiple trials to reach their full influence, and there was also noise that made all the functions harder to detect. Ambiguity in both the low and high-ambiguity policies could have led participants to allow their preferences to bias their learning and judgments, which is consistent with considerable bias in both conditions. Still, the difference in the ability to learn the low versus high ambiguity policies speaks against this possibility.

Another possibility is that when learning about the high-ambiguity policies, that participants did not notice the ambiguity (the opposing short-term and long-term influences) at all, and instead, only noticed the short-term influence. In Figure 1, the 'high-ambiguity' Functions 3 and 4 produce strong influences on the very first trial that they are implemented. In fact, the immediate influence is stronger than the immediate influence for the 'low-ambiguity' Functions 1 and 2, which take a couple trials to reach their full strength. It is possible that most participants therefore viewed Function 3 as fairly strong unambiguous evidence for a negative effect, and Function 4 as fairly strong unambiguous evidence for a positive effect, when in reality their long-term influences are the opposite. In fact, Sims et al. (2013) argued that when learning about policies with different short versus long-term influences, that the data that participants experience is not sufficient for them to learn the true functional form, and that learning the short-term relation is rational. Stated another way, even if these policies are ambiguous from the perspective of the experimenter, perhaps they were not ambiguous from the perspective of the participant.

Under this possibility, the participants' subjective experiences and interpretations would have been quite similar for the low and high ambiguity policies. This fits with the very poor performance for the high ambiguity policies (*Figures 6* & *7*), because poor performance of learning the long-term influence can be reinterpreted as very *good* performance for learning the short-term influence, just like the good performance of learning the low ambiguity functions. We believe that this explanation is a better fit to the results.

Regardless which of these possibilities is correct, both of them posit that people have considerable difficulty learning about the high ambiguity policies for which the short-term and long-term influences contradict each other, which is consistent with the prior findings using similar payoff functions (Gureckis & Love, 2009, and citations therein). This is especially problematic given that many economic policies (e.g., President Trump's justification for a trade war with China, providing universal early education, free college tuition) are believed to involve a trade-off between the short versus long-term.

Even though in this paper we did not find support for biased reasoning increasing in response to greater ambiguity, we suspect that such a pattern might be found in other situations. For example, it might be found when comparing the current policies to a policy that is truly unambiguous (it has an immediate and strong influence). Alternatively, it might be found when comparing a learning situation that involves very little noise (low-ambiguity) to a learning situation with considerable noise. Or, if the long-term benefit of the ambiguous policies came earlier, perhaps participants would become more aware of the temporal tradeoff and the ambiguity therein. In sum, there are many different ways in which ambiguity can arise, and other sorts of ambiguity could potentially moderate the motivated reasoning effect.

6.2 Open-Mindedness and Neutral Preferences

In Study 2, we had speculated that perhaps having neutral preferences would lead to better learning and more accurate causal judgments, compared to having strong preferences. The hypothesis was that when a learner has neutral preferences, they might be less biased, which could lead to more accurate learning. In contrast, when a learner has strong preferences, sometimes those preferences would be ‘congruent’ (their preferred policy option happens to be better) but sometimes they would be ‘incongruent (their preferred policy option happens to be worse). We speculated that perhaps the costs of preference-incongruence compared to neutral preferences would be larger than the benefits of preference-congruence compared to neutral preferences. The reason was that, if participants avoid testing their non-preferred options, they would learn little about them, potentially leading to very poor learning. In fact, avoiding testing non-preferred options could hurt both preference-incongruent as well as preference-congruent policies, because if a preferred option is repeatedly utilized, a learner does not get to test the comparison between the preferred versus the non-preferred option, which is critical for determining which policy option is better.[8]

Despite some apparent asymmetries in the means between preference-congruent, neutral, and preference-incongruent policies (*Figure 6*, *Figure 7*, and Table 1), no asymmetries were significant. On the one hand, this could be thought of as a lucky finding; even if people are biased, being biased on average in this task did not lead to worse causal learning. On the other hand, in the current study, randomization was used such that on average there were the same number of preference-congruent and preference-incongruent policies. However, in the real world, it is entirely possible that a population might in general have more preference-incongruent views than congruent (i.e., they might tend to prefer policies that are actually worse for the economy). If so, holding more neutral views initially could still be beneficial.

[8] In theory, these factors could play out differently for different measures. For example, if a participant blindly uses a preferred option during learning and rarely, if ever, tests the non-preferred option, then they would do very well at selecting the optimal choice during learning, but when identifying the functional form, they could do very poorly if they barely learned anything about the policy. For neutral policies, the performance on both tasks presumably depends largely on the task difficulty, which could affect the relative performance compared to preference-congruent and incongruent policies.

6.3 Expertise

In Study 3, we tested whether participants would perform better at learning and when making causal assessments if they were given initial instructions about possible types of functional forms of the policies. Most importantly, we wanted them to consider the possibility that a policy might have no influence on the economy at all, or that a policy might have a short-term benefit and a long-term consequence, or vice versa, since participants had so much difficulty learning about all of these policies. In a sense, having some more knowledge about potential functional forms could be viewed as a very light manipulation of expertise; true experts would presumably have more specific views about the timeframes within which a policy could play out.

Despite this hypothesis, there was little evidence that this manipulation made a difference. It did not seem to help them identify when a policy was non-causal (Table 2). It also did not improve the accuracy of assessing high-ambiguity functions (Table 1). Participants were about 15% more accurate in the function identification for the low-ambiguity functions (Table 1); we did not test whether this particular difference was significant only for low-ambiguity functions, but it was not significant for both low and high ambiguity functions.

There are a couple potential explanations for the failure of the intervention. First, perhaps the task is just so hard for the neutral policies and the high-ambiguous policies that the instructions were insufficient to make a difference. Second, it is possible that upon starting the task participants did not think back to the instructions. Third, though we think that this is fairly unlikely, perhaps even though participants passed the questions requiring some amount of understanding the instructions, they did not really understand all the functions.

Other research has found that though people can use prior knowledge about aspects like delay and carryover effects to adapt their causal testing strategies, they have difficulty using other knowledge such as wave-like changes over time (Rottman, 2016). Thus, it appears that adapting testing strategies can be a very challenging thing. Furthermore, other studies on the melioration paradigm have found that giving explicit instructions can help, but the largest benefit came when essentially telling participants which option is better in the long-run (Herrnstein et al., 1993). The current research suggests that even with some forewarning, people still have considerable difficulty learning about policies that have different short and long-term influences, but perhaps other forms of instruction or expertise could help.

6.4 Incentives and Taking the Task Seriously

One important question is the extent to which participants thought that their preferences and beliefs prior to the learning task could actually help them perform well during the learning task. For example, consider a participant who fervently believed that certain policies help the economy and others hurt, and imagine that they believed that the study was programmed to reflect how the actual economy works. In this case, it would be entirely rational to use the prior preferences and beliefs to guide learning. For a participant like this, the current study would be a good simulation of how motivated reasoning could play out in more real-world high-stakes situations.

Alternatively, consider another participant who believed that the study was just a game and that their real-world beliefs and preferences were irrelevant to performing well on the study. If so, then presumably they would be able to hold their preferences at bay, and try to learn in the most rational way as possible in order to maximize their bonus rewards for the task; accuracy was incentivized with bonuses all studies except 2B. In fact, accuracy incentives have been found to reduce and sometimes eliminate the partisan bias effect when assessing the current state of the economy (Bullock et al., 2013; Prior, Sood, & Kahanna, 2015). Yet, in our study, we still observed strong and reliable effects of prior preferences when learning about economic policies, which suggests that in some cases people do not just ignore their preferences even when financially motivated to do so.

It is entirely possible that the participants in these studies included a mixture of both of these sorts of beliefs, or primarily one type more than the other. However, we believe that the results of the current study are important regardless of the composition of the participants. In the first case, the study is a fairly good simulation of more real-world learning. In the second, it shows the power of preferences even when participants believe them to be irrelevant and are incentivized not to use them. Furthermore, this research revealed not just that preferences bias learning and judgment, but specific ways in which they bias learning and judgment.

6.5 Preferences versus Beliefs and Motivated Reasoning

In this paper we have extensively used the term ‘preference’, and at the beginning clarified that we would use ‘preference’ to also include both ‘beliefs’. For example, a person might prefer an increase to border security funding due to a belief that it would be good for the economy, or they might prefer an increase in border security funding for other reasons (e.g., security), even if they do not necessarily think that it would improve the economy – they might even prefer an increase in border security funding despite believing that it would hurt the economy. Distinguishing preferences versus beliefs is an important issue in the literature on rational versus affective accounts of motivated reasoning (Jern, Chang, Kemp, 2014; Nisbett & Ross, 1980; Tappin, Pennycook, Rand, 2019).

In the current study, we did not try to distinguish beliefs from preferences because we felt that they would often be correlated and would likely be hard to distinguish empirically. Thus, it is possible that some of the motivated reasoning could be due to participants importing their actual beliefs about economic policies and thinking that using such beliefs would help them perform better on this task if this task is an accurate simulation of the actual economy. Though this changes the nature of the motivation, we still think that it is important, perhaps even moreso, to understand how prior beliefs affect learning about policies. Future research could try to study how people learn about and test policies for which they prefer one option even if they believe it to be harmful to the economy (e.g., perhaps it has other benefits such as fairness).

6.6 Conclusions

The current research integrates paradigms from motivated reasoning and causal reasoning / reinforcement learning to understand how prior preferences affect how people go about testing the causal impact of policies and how people draw conclusions about policies. We found strong impacts of participants' prior preferences, even in this artificial task and even despite accuracy incentives. Similar processes may occur in real-world situations when one's preferences are even more likely to determine one's willingness to implement certain policies over others.

Appendix A Policy List

Public transportation safety standards	Internet infrastructure	Taxes on imported goods
Maternity/Paternity Leave	Flood risk management	Military spending
Workplace Discriminatory Policies	Drainage and sewerage	Counterterrorism spending
Equal pay for equal work	Carbon tax	Drug treatment
Social Security	Affordable housing	Police spending
Childcare Subsidies	Financial regulations	K-12 Education spending
Road maintenance	Taxes for the rich	University spending
Public transportation	Taxes for the poor	Border security
Large-scale 'green' tech.	Monopolies	Immigration
Subsidize public transit	Reduce drug prices	Marijuana legalization
Air travel infrastructure	Corporate tax rate	Small business tax rate
Gender equality and sexual harassment training		

Appendix B Initial Policy Assessment: Subjective Questions

#	*Policy Question Text: Subjective Version*
1	Would you prefer the government weakening or strengthening guidelines regarding gender equality and sexual harassment training in the workplace? (1=Strongly prefer weakening; 7 = Strongly prefer strengthening)
2	Would you prefer that the government require that all workplaces allow parents to take maternity/paternity leave? (1=Strongly prefer no requirement; 7 = Strongly prefer a requirement)
3	Would you prefer that the government does less or more to prohibit discriminatory policies in the workplace? (1=Strongly prefer less; 7 = Strongly prefer more)
4	Would you prefer that the government to implement less or more policies to ensure all workers are paid equally for equal work? (1=Strongly prefer less; 7 = Strongly prefer more)
5	Would you prefer that the government shrink or expand social security (that is, workers pay a percent of their income as taxes, and then are given money during retirement or if they are unemployed or disabled)? (1=Strongly prefer to shrink ; 7 = Strongly prefer to expand)
6	Would you prefer less or more government subsidies to help pay for child care costs for low-income families? (1=Strongly prefer less; 7 = Strongly prefer more)
7	Would you prefer that the government spends less or more on road maintenance? (1=Strongly prefer less; 7 = Strongly prefer more)
8	Would you prefer that the government spend less or more on public transportation (buses, trains/subways)? (1=Strongly prefer less; 7 = Strongly prefer more)
9	Would you prefer the government spend less or more on large-scale "green" technology public infrastructure projects (e.g., solar farms, hydroelectric dams, wind farms)? (1=Strongly prefer less; 7 = Strongly prefer more)
10	Would you prefer the government spend less or more on air travel infrastructure? (1=Strongly prefer less; 7 = Strongly prefer more)
11	Would you prefer the government spend less or more on safety standards for public transportation? (1=Strongly prefer less; 7 = Strongly prefer more)
12	Would you prefer the government spend less or more on high-speed internet infrastructure? (1=Strongly prefer less; 7 = Strongly prefer more)
13	Would you prefer the government spend less or more on addressing flood risk management? (1=Strongly prefer less; 7 = Strongly prefer more)
14	Would you prefer the government spend less or more on improving drainage and sewerage? (1=Strongly prefer less; 7 = Strongly prefer more)

15	Would you prefer the government pass laws to implement a "carbon tax" in an effort to reduce carbon dioxide emissions and promote "green" technology? (1=Strongly prefer not passing laws; 7 = Strongly prefer passing laws)
16	Would you prefer the government spend less or more subsidizing affordable housing for low-income citizens? (1=Strongly prefer less; 7 = Strongly prefer more)
17	Would you prefer less or more governmental financial regulations? (1=Strongly prefer less; 7 = Strongly prefer more)
18	Would you prefer lower or higher taxes for the wealthiest individuals? (1=Strongly prefer lower taxes for the wealthiest individuals; 7 = Strongly prefer higher taxes for the wealthiest individuals)
19	Would you prefer less or more governmental policies aimed at redistributing wealth (i.e., intending to help the poorest by decreasing the wealth of the richest)? (1=Strongly prefer less; 7 = Strongly prefer more)
20	Would you prefer that the government more aggressively breaks up large companies that could be viewed as monopolies, or just leave them alone? (1=Strongly prefer leaving them alone; 7 = Strongly prefer breaking them up)
21	Would you prefer that the government does less or more to reduce drug prices? (1=Strongly prefer less; 7 = Strongly prefer more)
22	Would you prefer a lower or higher corporate tax rate? (1=Strongly prefer lower; 7 = Strongly prefer higher)
23	Would you prefer lower or higher taxes on imported goods? (1=Strongly prefer lower; 7 = Strongly prefer higher)
24	Would you prefer the government decrease or increase military spending? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
25	Would you prefer the government decrease or increase counterterrorism spending? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
26	Would you prefer the government decrease or increase drug treatment spending (e.g., treating people with opioid addiction)? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
27	Would you prefer the government decrease or increase police spending? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
28	Would you prefer the government decrease or increase Kindergarten-12th Grade public education spending (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
29	Would you prefer the government decrease or increase university spending? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
30	Would you prefer the government decrease or increase border security spending? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
31	Would you prefer the government allow fewer or more immigrants from foreign countries? (1=Strongly prefer fewer; 7 = Strongly prefer more)
32	Would you prefer the government make marijuana illegal or legalize marijuana and tax it? (1=Strongly prefer marijuana being illegal; 7 = Strongly prefer legalizing marijuana and taxing it)

33	Would you prefer the government decrease or increase the tax rate for small businesses? (1=Strongly prefer decreasing; 7 = Strongly prefer increasing)
Note: Likert scale ratings on 7-point scale; rating of 4 = no preference.	

Appendix C Initial Policy Assessment: Objective Questions

#	*Policy Question Text: Objective Version*
1	Do you believe the government weakening or strengthening guidelines regarding gender equality and sexual harassment training in the workplace is better for the economy? (1=Strongly believe weakening guidelines regarding gender equality and sexual harassment training is better for the economy; 7 = Strongly believe strengthening guidelines regarding gender equality and sexual harassment training is better for the economy)
2	Do you believe the government requiring that all workplaces allow parents to take maternity/paternity leave is better for the economy? (1=Strongly prefer no requirement for maternity/paternity leave is better for the economy ; 7 = Strongly prefer a requirement for maternity/paternity leave is better for the economy)
3	Do you believe the government doing less or more to prohibit discriminatory policies in the workplace is better for the economy? (1=Strongly believe that the gov. doing less to prohibit discrimination in the workplace is better for the economy; 7 = Strongly believe that the gov. doing more to prohibit discrimination in the workplace is better for the economy)
4	Do you believe the government to implement less or more policies to ensure all workers are paid equally for equal work? (1=Strongly believe less policies to ensure workers are paid equally for equal work; 7 = Strongly believe more policies to ensure workers are paid equally for equal work)
5	Do you believe the government shrinking or expanding social security is better for the economy– that is, workers pay a percent of their income as taxes, and then are given money during retirement or if they are unemployed or disabled. (1=Strongly believe shrinking social security is better for the economy; 7 = Strongly believe expanding social security is better for the economy)
6	Do you believe less or more government subsidies to help pay for child care costs for low-income families is better for the economy? (1=Strongly believe less gov. subsidies for child care costs for low-income families is better for the economy; 7 = Strongly believe more gov. subsidies for child care costs for low-income families is better for the economy)
7	Do you believe the government spending less or more on road maintenance is better for the economy? (1=Strongly believe that the gov. spending less on roads is better for the economy; 7 = Strongly believe that the gov. spending more on roads is better for the economy)
8	Do you believe the government spending less or more on public transportation (buses, trains, subways/railways) is better for the economy? (1=Strongly believe that the gov. spending less on public transportation is better for the economy; 7 = Strongly believe that the gov. spending more on public transportation is better for the economy)
9	Do you believe the government spending less or more on large-scale "green" technology public infrastructure projects (e.g., solar farms, hydroelectric dams, wind farms) is better for the economy? (1=Strongly believe gov. spending less on large-scale "green" infrastructure projects is better for the economy; 7 = Strongly believe gov. spending more on large-scale "green" infrastructure projects is better for the economy)

10	Do you believe the government spending less or more on air travel infrastructure is better for the economy? (1=Strongly believe the gov. spending less in air travel infrastructure is better for the economy; 7 = Strongly believe the gov. spending more in air travel infrastructure is better for the economy)
11	Do you believe the government spending less or more on safety standards for public transportation is better for the economy? (1=Strongly believe the gov. spending less on public transportation safety standards is better for the economy; 7 = Strongly believe the gov. spending more on public transportation safety standards is better for the economy)
12	Do you believe the government spending less or more on high-speed internet infrastructure is better for the economy? (1=Strongly believe the gov. spending less on high-speed internet infrastructure is better for the economy; 7 = Strongly believe the gov. spending more on high-speed internet infrastructure is better for the economy)
13	Do you believe the government spending less or more on addressing flood risk management is better for the economy? (1=Strongly believe the gov. spending less on addressing flood risk management is better for the economy; 7 = Strongly believe the gov. spending more on addressing flood risk management is better for the economy)
14	Do you believe the government spending less or more on improving drainage and sewerage is better for the economy? (1=Strongly believe the gov. spending less on improving drainage and sewerage is better for the economy; 7 = Strongly believe the gov. spending more on improving drainage and sewerage is better for the economy)
15	Do you believe the government passing laws to implement a "carbon tax" in an effort to reduce carbon dioxide emissions and promote "green" technology is better for the economy? (1=Strongly believe the gov. not passing laws to implement a "carbon tax" is better for the economy; 7 = Strongly believe the gov. passing laws to implement a "carbon tax" is better for the economy)
16	Do you believe the government spending less or more subsidizing affordable housing for low-income citizens is better for the economy? (1=Strongly believe the gov. spending less subsiding affordable housing for low-income citizens is better for the economy; 7 = Strongly believe the gov. spending more subsiding affordable housing for low-income citizens is better for the economy)
17	Do you believe less or more governmental financial regulations is better for the economy? (1=Strongly believe the gov. having less financial regulations is better for the economy; 7 = Strongly believe the gov. having more financial regulations is better for the economy)
18	Do you believe lower or higher taxes for the wealthiest individuals is better for the economy? (1=Strongly believe lower taxes for the wealthiest individuals is better for the economy; 7 = Strongly believe higher taxes for the wealthiest individuals is better for the economy)
19	Do you believe less or more governmental policies aimed at redistributing wealth (i.e., intending to help the poorest by decreasing the wealth of the richest) is better for the economy? (1=Strongly believe less gov. policies aimed at redistributing wealth is better for the economy; 7 = Strongly believe more gov. policies aimed at redistributing wealth is better for the economy)
20	Do you believe the government more aggressively breaking up large companies that could be viewed as monopolies, or just

	leaving them alone is better for the economy? (1=Strongly believe the gov. leaving alone large companies that could be viewed as monopolies is better for the economy; 7 = Strongly believe the gov. breaking up large companies that could be viewed as monopolies is better for the economy)
21	Do you believe the government doing less or more to reduce drug prices is better for the economy? (1=Strongly believe the gov. doing less to reduce drug prices is better for the economy; 7 = Strongly believe the gov. doing more to reduce drug prices is better for the economy)
22	Do you believe a lower or higher corporate tax rate is better for the economy? (1=Strongly believe a lower corporate tax rate is better for the economy; 7 = Strongly believe a higher corporate tax rate is better for the economy)
23	Do you believe lower or higher taxes on imported goods is better for the economy? (1=Strongly believe lower taxes on imported goods is better for the economy; 7 = Strongly believe higher taxes on imported goods is better for the economy)
24	Do you believe decreasing or increasing military spending is better for the economy? (1=Strongly believe decreasing military spending is better for the economy; 7 = Strongly believe increasing military spending is better for the economy)
25	Do you believe decreasing or increasing counterterrorism spending is better for the economy? (1=Strongly believe decreasing counterterrorism spending is better for the economy; 7 = Strongly believe increasing counterterrorism spending is better for the economy)
26	Do you believe decreasing or increasing government spending for drug treatment is better for the economy (e.g., treating people with opioid addiction)? (1=Strongly believe decreasing drug treatment spending is better for the economy; 7 = Strongly believe increasing drug treatment spending is better for the economy)
27	Do you believe decreasing or increasing police spending is better for the economy? (1=Strongly believe decreasing police spending is better for the economy; 7 = Strongly believe increasing police spending is better for the economy)
28	Do you believe decreasing or increasing Kindergarten-12th Grade public education spending is better for the economy? (1=Strongly believe decreasing K-12 public education spending is better for the economy; 7 = Strongly believe increasing K-12 public education spending is better for the economy)
29	Do you believe decreasing or increasing university spending is better for the economy? (1=Strongly believe decreasing university spending is better for the economy; 7 = Strongly believe increasing university spending is better for the economy)
30	Do you believe decreasing or increasing border security spending is better for the economy? (1=Strongly believe decreasing border security spending is better for the economy; 7 = Strongly believe increasing border security spending is better for the economy)
31	Do you believe allowing fewer or more immigrants from foreign countries is better for the economy? (1=Strongly believe allowing fewer immigrants is better for the economy; 7 = Strongly believe allowing more immigrants is better for the economy)
32	Do you believe making marijuana illegal or legalizing marijuana and taxing it is better for the economy? (1=Strongly believe marijuana being illegal is better for the economy; 7 = Strongly believe legalizing marijuana and taxing it is better for the economy)

33	Do you believe decreasing or increasing the tax rate for small businesses is better for the economy? (1=Strongly believe decreasing the tax rate for small businesses is better for the economy; 7 = Strongly believe increasing the tax rate for small businesses is better for the economy)
Note: Likert scale ratings on 7-point scale; rating of 4 = no preference.	

Appendix D Randomization Outcomes

Table D1. Number of Preference-Congruent, Incongruent, and Neutral Policies for Each Type of Causal Function That Arose Through Randomization.

Study	*Causal Functions*	*Preference*		
		Congruent	*Neutral*	*Incongruent*
1	Low-Ambiguity: Positive	19	–	22
N = 41	Low-Ambiguity: Negative	19	–	22
	High-Ambiguity: Positive	23	–	18
	High-Ambiguity: Negative	14	–	27
2A	Low-Ambiguity: Positive	24	45	19
N = 88	Low-Ambiguity: Negative	21	41	26
	High-Ambiguity: Positive	20	51	17
	High-Ambiguity: Negative	25	39	24
2B	Low-Ambiguity: Positive	67	146	70
N = 283	Low-Ambiguity: Negative	72	135	76
	High-Ambiguity: Positive	65	145	73
	High-Ambiguity: Negative	67	151	65
3	Low-Ambiguity: Positive	37 (20, 17)	–	37 (15, 22)
N = 78	Low-Ambiguity: Negative	40 (16, 24)	–	36 (19, 17)
	High-Ambiguity: Positive	36 (15, 21)	–	38 (21, 17)
	High-Ambiguity: Negative	40 (21, 19)	–	36 (15, 21)

Note: Cell values are participant frequency counts. Parenthetical values in Study 3 refer to condition assignment for seeing underlying mechanisms (value 1) or not (value 2).

Table D2. Randomization for Non-Causal Functions.

Study	Preference Present	
	Yes	No
Study 1 (*N* = 41)	82	–
Study 2A (*N* = 88)	87	89
Study 2B (*N* = 283)	280	286
Study 3 (*N* = 77)	151 (68, 83)	–

Note: Cell values are frequency counts. Parenthetical values in Study 3 refer to condition assignment for seeing underlying mechanisms (value 1) or not (value 2).

Appendix E Strategy Assessment Variables and Outcomes

This section reports analyses of the relation between different testing strategies (T) during learning and learning outcomes (O). It did not make sense to include all the measures of choice during testing. For example, even though time until testing could be a marker of bias for or against a particular policy, when considering testing as a whole, the tendency to make changes earlier during the 140 trials, or later, for all the policies in general, is not a clear measure of the quality of testing. In this section, we studied two of the testing strategies from the main paper, Preferred Policy Percent, and Optimal Policy Percent. In addition, we also studied five other measures of testing strategies that we thought might be related to participants' outcomes. The two outcomes were Policy Judgment Accuracy and Function Identification Accuracy, both of which were studied in the main paper. Some of the testing strategies, especially the percent of choices during the learning task that were the optimal policy (Optimal Policy Percent) could be viewed as either testing strategies or outcomes. Since this occurred during the testing phase, it is labeled as a testing strategy.

Overall the testing strategy variables we identified shared a tenuous relationship with learning outcomes. Below is a quick summary of the relationship between the variables. Strategies tended to correlate with one another which makes sense as all strategy variables measure the act of testing in one form or another. For instance, participants with greater controlled changes also tended to have high amounts of confounded changes (Study 1, 2A, 2B). Additionally, participants with higher standard deviation scores for their testing behavior had fewer controlled changes (Study 1, 2A, 2B, 3) and cofounded changes (Study 1, 2A). Notably, skewness of testing largely did not share a relationship with the other strategy variables.

We found some support that participants with higher standard deviation scores were associated with a greater number of trials having chosen preferred policies (Study 2A, 2B) and a greater numbers of trials with optimal policies (Study 1, 2A, 2B). We did not find evidence that a testing distribution's skew impacted learning outcomes. We found supportive evidence that having a greater number of controlled changes during testing was associated with fewer trials having chosen preferred policies (Study 2A, 2B, 3). Also, having more confounded changes during testing was associated with fewer trials having chosen preferred policies (Study 2B, 3), fewer trials having chosen optimal policies (Study 2A), and poorer accuracy in judgments of policy efficacy (Study 2A). However, given the mixed nature of these results caution is warranted in drawing any firm conclusions.

Appendix E.1 Five Testing Strategies

E.1.1 Testing Strategy #1: Standard Deviation. This variable measures the average distance (in trials) between policy changes. To compute this, we examine whether a policy was changed and, if so, what trial number did this change occur on. For each participant, a list is created counting the distance between trials for every policy change in the learning task. If a participant makes several policy changes on the same trial they will all be included. This list is than averaged to create a single number representing, on average, how spaced out testing was for each participant. We predict that higher scores would be associated with positive learning outcomes because with greater distances between policy changes there is more time to witness the effect of causes (policy changes).

E.1.2 Testing Strategy #2: Skew. This variable measures how skewed the distribution of policy changes across trials is (i.e., to what extent were policy changes equally spread out, or not). This variable used the same list as the 'Standard Deviation' variable, except instead of averaging the distance between trial changes, this list was inputted into *R*'s skewness function (type 3). We predict less skew in policy testing to be associated with positive outcomes. The reasoning is that equal testing intervals (skewness = 0) allow for more controlled comparisons across policy changes.

E.1.3 Testing Strategy #3: Controlled Changes. This variable measures the total number trials out of 140 during which a participant made a change to exactly one of the six policies. We expected there to be positive relations between the number of controlled changes with causal learning because controlled changes permit an opportunity to attribute a change in the EVI to a particular policy. (However, if controlled changes are made in close proximity it participants could attribute changes in EVI to the wrong policy.)

E.1.4 Testing Strategy #4: Confounded Changes. This variable measures when the total number of instances in which a participant made a change to 2 or more out of the six policies simultaneously. We predict that confounded changes being associated with poorer learning outcomes because participants would not know which policy to attribute a change in the EVI to. To calculate this, each trial was given a score of x when x policy changes are made, excluding trials of single policy changes (range of x: 2-6). These scores were summed across all trails. For example, if a participant made 5 policy changes on Trial 15 and 2 policy changes on Trial 20, but otherwise only made a single policy change on all other trials (if a change was made at), the confounded changes score would be 7.

E.1.5 Testing Strategy #5: Total Policy Changes. This variable is a count of all policy changes in the learning task. The sum of controlled changes and confounded changes is equal to total policy changes. We did not have predictions about how total policy changes may relate to outcomes because too few would not provide enough opportunities for learning and too many would lead to difficulty attributing changes in EVI to a particular policy. This variable is included to show how it relates to other variables.

E.1.6 Testing Strategy #6: Preferred Policy %. This is the percent of instances that a participant chose their preferred policy option out of all 140 trials and all the levers for which they had a preference (up to 840 instances depending on the study).

E.1.7 Testing Strategy #7: Optimal Policy %. This is the percent of instances that a participant chose the optimal policy out of all 140 trials and for the 4 causal policies (560 instances).

Appendix E.2 Two Outcomes

E.2.1 Outcome #1: Policy Judgment Accuracy. This is the percent of policies correctly assessed at the end of the learning task for the six policies.

E.2.2 Outcome #2: Function Identification Accuracy. This is the percent of policies that were successfully matched in the function identification task.

Appendix E.3 Strategy and Outcomes Matrices

Table E1. Strategy-Outcome Correlation Matrix for Study 1

	T1	T2	T3	T4	T5	T6	T7	O1	O2
T1. Standard Deviation	–								
T2. Skew	-.32*	–							
T3. Controlled Changes	-.29†	.19	–						
T4. Confounded Changes	-.30†	.20	.35*	–					
T5. Total Policy Changes	-.36*	.24	.89***	.74***	–				
T6. Preferred Policy %	-.05	-.27†	-.05	-.09	-.08	–			
T7. Optimal Policy %	.30†	.10	-.22	-.23	-.27†	-.09	–		
O1. Policy Judgment Accuracy	-.04	-.02	-.11	-.06	-.11	-.12	.21	–	
O2. Function ID Accuracy	.09	.08	-.09	.09	-.02	-.11	.24	.34*	–

Note: $N = 41$. Cell values are Pearson r coefficients. Significance levels: †$p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. T denotes testing strategy and O denotes outcome.

Table E2. Strategy-Outcome Correlation Matrix for Study 2A

	T1	T2	T3	T4	T5	T6	T7	O1	O2
T1. Standard Deviation	–								
T2. Skew	-.02	–							
T3. Controlled Changes	-.40***	-.15	–						
T4. Confounded Changes	-.41***	-.06	.25*	–					
T5. Total Policy Changes	-.50***	-.15	.92***	.61***	–				
T6. Preferred Policy %	.24*	-.01	-.30**	-.10	-.29**	–			
T7. Optimal Policy %	.34**	-.02	-.16	-.18†	-.20†	.16	–		
O1. Policy Judgment Accuracy	.36***	.14	.01	-.33**	-.13	-.05	.53***	–	
O2. Function ID Accuracy	.01	.02	.09	-.08	.04	-.01	.00	-.08	–

Note: $N = 88$. Cell values are Pearson r coefficients. Significance levels: †$p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$. T denotes testing denotes strategy and O denotes outcome.

Table E3. Strategy-Outcome Correlation Matrix for Study 2B

	T1	T2	T3	T4	T5	T6	T7	O1	O2
T1. Standard Deviation	–								
T2. Skew	-.16**	–							
T3. Controlled Changes	-.21***	-.09	–						
T4. Confounded Changes	-.09	-.07	.39***	–					
T5. Total Policy Changes	-.19**	-.10	.88***	.78***	–				
T6. Preferred Policy %	.16**	.00	-.17**	-.10†	-.17**	–			
T7. Optimal Policy %	.11†	.11†	-.15**	-.07	-.14*	-.03	–		
O1. Policy Judgment Accuracy	-.09	.09	-.06	-.02	-.05	.05	.47***	–	
O2. Function ID Accuracy	-.15**	.01	-.05	-.02	-.04	.00	.07	.30***	–

Note: $N = 283$. Cell values are Pearson r coefficients. Significance levels: †$p < .10$; *$p < .05$; **$p < .01$; ***$p <.001$. T denotes testing denotes strategy and O denotes outcome.

Table E4. Strategy-Outcome Correlation Matrix for Study 3

	T1	T2	T3	T4	T5	T6	T7	O1	O2	X1.
T1. Standard Deviation	–									
T2. Skew	-.05	–								
T3. Controlled Changes	-.21†	.09	–							
T4. Confounded Changes	-.17	-.16	.11	–						
T5. Total Policy Changes	-.26*	.00	.89***	.55***	–					
T6. Preferred Policy %	.12	-.31**	-.29**	-.19†	-.33**	–				
T7. Optimal Policy %	-.01	-.17	-.08	.10	-.02	-.09	–			
O1. Policy Judgment Accuracy	-.12	-.02	-.02	-.05	-.04	-.05	.48***	–		
O2. Function ID Accuracy	-.04	.09	.05	-.05	.02	.16	-.08	.03	–	
X1. Function Exposure	-.01	.16	.00	.12	.06	-.06	-.10	-.11	.20†	–

Note: $N = 78$. Cell values are Pearson r coefficients. Significance levels: †$p < .10$; *$p < .05$; **$p < .01$; ***$p <.001$. T denotes testing denotes strategy; O denotes outcome; X denotes treatment/intervention.

References

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: the joint influence of prior expectations and current situational information. *Psychological Review*, *91*(1), 112–49.

Bullock, J. G., Gerber, A. S., Hill, S. J., & Huber, G. A. (2015). Partisan Bias in Factual Beliefs about Politics. *Quarterly Journal of Political Science*, *10*(4), 519-578.

Campbell, T. H., & Kay, A. C. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology*, *107*(5), 809–824.

Derringer, C., & Rottman, B. M. (2018). How people learn about causal influence when there are many possible causes: A model based on informative transitions. *Cognitive Psychology*, *102*, 41–71.

Factbase (2018). Interview: WABC Radio's Bernie McGuirk and Sid Rosenberg Interview Donald Trump on Bernie & Sid - April 6, 2018. Retrieved January 14, 2020, from https://factba.se/transcript/donald-trump-interview-bernie-sid-show-wabc-april-6-2018.

Fugelsang, J. A., & Thompson, V. (2000). Strategy selection in causal reasoning: when beliefs and covariation collide. *Canadian Journal of Experimental Psychology*, *54*(c), 15–32.

Fugelsang, J. A., & Thompson, V. A. (2003). A dual-process model of belief and evidence interactions in causal reasoning. *Memory & Cognition*, *31*(5), 800–815.

Goedert, K. M., Ellefson, M. R., & Rehder, B. (2014). Differences in the weighting and choice of evidence for plausible versus implausible causes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(3), 683–702.

Gureckis, T. M., & Love, B. C. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*, *113*(3), 293–313.

Hart, P. S., & Nisbet, E. C. (2011). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research*, *39*(6), 701–723.

Herrnstein, R. J., Loewenstein, G. F., Prelec, D., & Vaughan, W., Jr. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making, 6,* 149–185.

Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political preferences in the face of counterevidence. *Scientific Reports*, *6*, 39589.

Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental*

Psychology, *33*(2), 273–283.

Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (2010). Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. *Law and human behavior*, *34*(6), 501–516.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86.

Jern, A., Chang, K. K., & Kemp, C. (2014). Belief polarization is not always irrational. *Psychological Review*, *121*(2), 206–224.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*(4), 636–647.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480.

Luhmann, C. C., & Ahn, W. (2007). BUCKLE: A Model of Unobserved Cause Learning. *Psychological Review*, *114*(3), 657–677.

Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharot, T. (2018). Epistemic Spillovers: Learning Others' Political Views Reduces the Ability to Assess and Use Their Expertise in Nonpolitical Domains. *Harvard Public Law Working Paper*, *No. 18–22*.

Marsh, J. K., & Ahn, W. (2006). Order effects in contingency learning: The role of task complexity. *Memory & Cognition*, *34*(3), 568–576.

Marsh, J. K., & Ahn, W. (2009). Spontaneous Assimilation of Continuous Values and Temporal Information in Causal Induction. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(2), 334–352.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*(2), 303–330.

Otto, A. R., Gureckis, T. M., Markman, A. B., & Love, B. C. (2009). Navigating through abstract decision spaces: Evaluating the role of state generalization in a dynamic decision-making task. *Psychonomic Bulletin & Review, 16,* 957–963.

Paharia, N., Vohs, K. D., & Deshpandé, R. (2013). Sweatshop labor is wrong unless the shoes are cute: Cognition can both help and hurt moral motivated reasoning. *Organizational Behavior and Human Decision Processes*, *121*(1), 81–88.

Prior, M., Sood, G., & Khanna, K. (2015). You cannot be serious: The impact of accuracy incentives on partisan bias in reports of economic perceptions. *Quarterly Journal of Political Science*, *10*(4), 489-518.

Rottman, B. M. (2016). Searching for the best cause: Roles of mechanism preferences, autocorrelation, and exploitation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *42*(8), 1233–1256.

Shepardson, D. (2018). Senator McConnell says tariffs hurt, urges trade progress. Retrieved January 28, 2020, from https://www.reuters.com/article/us-usa-senate-mcconnell-trade/senator-mcconnell-says-tariffs-hurt-urges-trade-progress-idUSKCN1MR2UZ

Sims, C. R., Neth, H., Jacobs, R. A., & Gray, W. D. (2013). Melioration as rational choice: Sequential decision making in uncertain environments. *Psychological Review*, *120*(1), 139.

Spellman, B. A. (1996). Conditionalizing causality. *Psychology of learning and motivation*, *34*, 167–206.

Stillwell, D. J., & Tunney, R. J. (2009). Melioration behavior in the Harvard game is reduced by simplifying decision outcomes. *Quarterly Journal of Experimental Psychology, 62,* 2252–2261.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political preferences. *American Journal of Political Science*, *50*(3), 755–769.

Tappin, B. M., Pennycook, G., & Rand, D. G. (2019, November 3). Thinking clearly about causal inferences of politically motivated reasoning: Why paradigmatic study designs often undermine causal inference. https://doi.org/10.31234/osf.io/k3jhy

Torry, H. (2019). Trupt Tariffs Are Short-Term Pain Without Long-Term Gain, Economists Say. Retrieved January 28, 2020, from https://www.wsj.com/articles/trump-tariffs-are-short-term-pain-without-long-term-gain-economists-say-11560440436