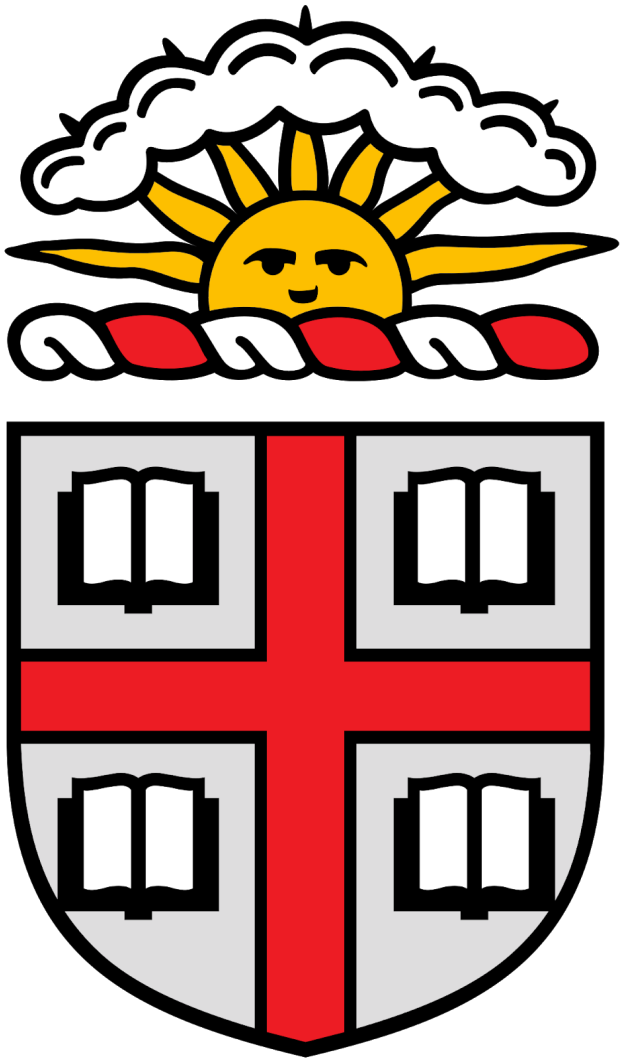
# Motivated Reasoning in a Community of Knowledge



*This thesis is submitted in partial completion of the requirements for the degree of Bachelor of Arts with Honors in the Health and Human Biology concentration of Brown University.*

May 2019

Matthew Orringer _____

Advisor: Steven Sloman, PhD _____

**Table of Contents**

**Abstract**

Though we all have access to a world of knowledge in the palms of our hands, people today are becoming increasingly unlikely to engage with ideas or opinions with which they do not agree. The present study investigates confirmation bias and motivated reasoning and how they may or may not contribute to creating a divide in opinions. Experiment 1 investigates how we update objective and subjective beliefs when given consensus or fact-based evidence. Results from Experiment 1 suggest that we update objective beliefs significantly more willingly than subjective beliefs. Experiment 2 examines motivated reasoning by assessing the credibility judgments of individuals affected and not affected by evidence related to neutral or charged topics. Experiment 2 produced mixed results, with one of the four prompts providing evidence for motivated reasoning in the manner in which we expected and another providing evidence for a community-based motivated reasoning pattern. Experiment 3 investigates how reflectiveness (as measured by the Cognitive Reflection Test) and the order in which topics are presented may affect belief updating patterns. Results from Experiment 3 suggest that high CRT individuals update their beliefs more than low CRT individuals, particularly on neutral topics. We also found evidence of an interaction between CRT scores and the order in which neutral and emotionally charged items are presented for evaluation. Overall, the present study investigated the conditions under which individuals are more or less willing to update their beliefs as well as the mechanisms that are employed to make these decisions.

## Introduction

Aside from the Civil War era, perhaps more today than ever before, the United States has become a country in which two separate realities coexist (Jones, 2019). Many sources have been identified as the cause of this polarity, among them the social media echo chambers and the emergence of "fake news." There are, however, psychological reasons underlying this increasing divide between factions of the American populous. One such factor is confirmation bias, a form of motivated reasoning in which we seek out information that verifies our beliefs and reject information that conflicts with already held views.

Gilovich (1991) defines motivated reasoning with regard to how we evaluate novel evidence. He notes that with evidence which is confirmatory in nature we ask, "Can I believe this?" However, we greet disconfirming evidence with the more skeptical response of "Must I believe this?" Motivated reasoning may present itself in a number of forms. For example, Wyer and Frey (1983) found that individuals who were told that they performed poorly on an intelligence test were more likely to question its validity than were subjects who were told that they performed well on the test.

Kunda (1987) extended these findings and was able to control for the potentially confounding effects that prior beliefs may play in dictating judgments. The study provided evidence that caffeine consumption was harmful for women. As a result, women who were heavy consumers of caffeine found the evidence to be less credible than women who consumed low levels of caffeine. The effect did not hold for male subjects, who, one can reasonably assume, held similar initial beliefs about caffeine to the female subjects. This paradigm serves as the basis for Experiment 2 in the present study.

Dawson et al. (2002) investigated motivated reasoning by assessing subjects' performance on the Wason Selection Task, a well-known logic puzzle that examines one's deductive reasoning skills. The experimenters found that when the stakes of a similar Wason task were raised (by implying early death in one study and by affirming a negative stereotype in the other), subjects who felt threatened greatly outperformed those were not at risk from the results. This suggests a heightened reasoning ability when subjects are motivated by the results at hand.

Kahan et al. (2017) found that when evaluating a neutral topic (skin rash treatments), more numerate individuals were more effective in logically evaluating the data. However, when the topic being discussed was politically and emotionally charged (gun control), more numerate individuals were actually more polarized in their interpretation of the data, suggesting that these individuals use their advanced reasoning skills in order to maintain their views. This work serves as the basis for Experiment 3 of the present study.

Another of Kahan's (2011) studies found that participants were quick to dismiss the credentials of scientists who espoused views that were in conflict with their opinions. A 2015 study found that when responding to politically charged scientific topics (climate change, fracking, etc.), as subjects' scientific literacy increased, so too did the polarization of opinions of liberals and conservatives. However, on neutral scientific topics (artificial food coloring, nanotechnology), scientific literacy, and not political affiliation, was the only predictor of attitudes. These studies suggest that we use our reasoning skills selectively and apply different reasoning mechanisms when we are motivated to arrive at a given result.

One particularly noteworthy motivated reasoning study, conducted by Nyhan and Reifler (2010), sparked interest in the "backfire effect." The researchers found that in specific instances,

presenting participants with information that contradicts prior beliefs actually made subjects more sure of their initial beliefs. This effect has proven to be difficult to recreate as there have been many failures to replicate the original study (see Haglin, 2017, Weeks et al., 2014, Garrett et al., 2013).

Recently, neuroscientists have sought to illustrate a biological basis for confirmation bias and motivated reasoning. One study, conducted by Westen et al. (2006), used fMRI technology to monitor subjects' brain activity while they listened to statements made by both their favored Presidential candidate and his or her opponent. Not surprisingly, subjects accepted exculpatory statements made by their preferred candidate significantly more often than those made by the opposing candidate. The study identified medial prefrontal cortex activation in subjects "processing emotionally threatening information about (their) preferred candidate." Additionally, the posterior cingulate cortex, which has been associated with forgiveness and moral judgements (see Cunningham, Raye, & Johnson, 2004), was specifically activated in these instances of processing threatening information. The study also found that participants arrived at conclusions that maintained their prior beliefs when confronted with negative evidence about their candidate. The researchers found that the dorsolateral prefrontal cortex, which is associated with emotionally-independent reasoning, was not "differentially activated." Thus, this evidence suggests that there is a neurologically distinct process that takes place during motivated reasoning.

Additionally, a 2011 study conducted by Doll et al. suggests there is an underlying genetic mechanism that may explain why certain individuals may be particularly susceptible to confirmation bias. The study identified genetic differences in the Catechol-O-methyltransferase

(COMT) gene, which is involved in the prefrontal dopaminergic pathway, as being implicated in the differential persistence of beliefs in the face of contradictory information.

While there is certainly potential in investigating the neurobiological underpinnings of confirmation bias and motivated reasoning, there are inherent limitations in these studies, namely the fact that we as humans live in a "community of knowledge" and not as distinct, independent cognitive entities. Sloman and Fernbach (2017) discuss how we often think we know more than we do because we know that other people understand information about specific topics and we conflate their knowledge with our own. Yousif and Keil (2018) extended this understanding of human cognition in a recent experiment that investigated how people make decisions. The study found that individuals use the metrics of consensus of the masses and confidence of the masses in order to make decisions. However, participants used these two metrics differently in different contexts. For objective subjects, participants preferred consensus evidence whereas they preferred confidence-based evidence on subjective matters. This study provides the foundation for Experiment 1 of the present study.

The present study investigated a number of the issues outlined above. Experiment 1 examined differential changes in subjective and objective beliefs when presented with consensus and fact-based evidence. We hypothesized that subjective beliefs would be more resistant to change because they are grounded more in emotion and less in data as compared to objective beliefs. As such, subjective beliefs may be more intrinsically tied to our self-image and thus we may view an attack on these beliefs as an attack on who we are. We predicted that this would motivate subjects to uphold their subjective beliefs more than their objective beliefs in the face of novel evidence. We also hypothesized that fact-based evidence would be more effective in

changing the beliefs of subjects because empirical evidence is more difficult to argue away, ignore, or justify. Additionally, we investigated whether the credibility of the evidence or one's expertise on a given topic would affect their judgments. We hypothesized that individuals would find evidence that supports their initial beliefs to be more credible due to either normative or motivated reasoning (which we investigated further in Experiment 2). Finally, we predicted that individuals with greater expertise on a given topic would be more hesitant to change their views.

Experiment 2 sought to illuminate the mechanism of reasoning being employed to make credibility judgments. We hypothesized that if subjects were using motivated reasoning, then evidence that positively affects them would be regarded as more credible, while evidence that presents a threat would be regarded as less credible. If one was not affected by the evidence, we hypothesized an intermediate credibility rating from these subjects.

In Experiment 3, we further examined the conditions under which individuals are more or less willing to change their beliefs. We investigated individuals' reflectiveness (as measured by the CRT) and the order in which neutral and charged topics were presented as potential causes of differential belief updating. Experiment 3 examined Kahan's contention that more numerate individuals are better at self-deception. We hypothesized that among more numerate individuals, those who responded to emotionally charged topics first would be less likely to change their beliefs throughout the experiment.

In sum, this collection of experiments sought to evaluate the conditions under which individuals are more or less willing to update their beliefs as well as the mechanisms that are employed to make these decisions. We investigated the effects of the evidence, topic, type of question, and reasoning strategy employed in order to arrive at a holistic interpretation of how

we make judgments in order to maintain maximal consistency with both the evidence presented as well as our initial beliefs.

## Experiment 1

*Introduction:*

In this experiment, we sought to answer a number of questions. First, we investigated the effects of different types of evidence, specifically consensus-based and "fact"-based evidence, on belief updating. We predicted that fact-based evidence would be more effective in changing beliefs as it would be more difficult to explain away or justify. We also studied the differential effects of the type of topic in question, namely subjective and objective topics. We predicted that objective beliefs would be more likely to be changed as they are grounded more in data and less in emotion. In addition, we examined the effects that the credibility of evidence and expertise of the subject on the given topic had on belief updating. We hypothesized that individuals would find evidence that aligns with their initial beliefs to be more credible and that individuals with a higher degree of expertise on a given topic would change their opinions less.

*Methods*

Two hundred Brown University students were recruited to participate in this study (51% female; average age: 19.4) in exchange for either class credit or participation in a lottery for $50. Participants were given a packet and asked to respond to each question with a number between 0 and 100. Each participant gave initial belief responses to 16 statements. Eight unique topics were covered in these statements with half of the statements being objective in nature and half being subjective. The topics included sports, economics, abortion, politics, religion, media, education,

and climate change. See Appendix 1 for the specific questions and evidence of Form 1. Each form had the same subjective and objective statements as well as an equal number of both positive vs. negative and factual vs. consensus-based evidence. Thus, across the four forms, each question had an equal number of each of the four types of evidence (positive vs. negative and consensus vs. fact-based). Participants were randomly assigned to one of the four forms to counterbalance the sample. Once the participants provided their initial belief responses, they were provided a piece of evidence about each topic and asked to update their beliefs on the same 0-100 scale. For each question, a given participant received either positive or negative evidence that was either consensus or "fact"-based. Each consensus-based piece of evidence detailed a description of 72-76% of experts in the field in question having a given belief. Positive evidence was intended to increase the individual's belief while negative evidence was intended to decrease his or her belief. Much of the evidence in this study was created by the researchers which was explained to the subjects at the conclusion of the study. After giving updated beliefs for each statement, participants were asked to rate the degree to which they found the evidence provided to be believable on a scale of 0-100. After providing these credibility ratings for each statement, the participants provided a rating of their individual expertise on the topic discussed in each statement on a scale of 0-100. Finally, demographic information was collected (gender identity, age, education level).

*Results*:

We reverse coded changes in belief that responded to negative evidence such that all of the changes in belief were positive if no backfire effect occurred.

We used the equation $|FB - IB| = Difference\ in\ belief$

where FB is final belief and IB is initial belief.

For example, when responding to positive evidence, an initial belief of 50 and a final belief of 80 was scored as a difference of 30 in the same way that when responding to negative evidence, an initial belief of 50 and a final belief of 20 was scored as a difference of 30.

We confirmed that there were not significant differences in average initial beliefs for objective vs. subjective topics. The average initial belief for objective topics was 54.3 (Standard error = 5.3) while the average initial belief for subjective statements was 52.4 (Standard error = 5.7).

We found the following average changes in belief for each of the four combinations of type of issue (Objective versus Subjective) crossed with type of evidence (Factual versus Consensus):

**Table 1:**

| Mean Change in Belief (Standard Error) | Objective | Subjective |
|---|---|---|
| Factual | 15.4 (1.9) | 10.8 (1.2) |
| Consensus | 13.7 (2.0) | 8.6 (1.2) |

The results indicate objective beliefs were changed significantly more than subjective beliefs, as predicted. This was confirmed by an ANOVA analysis ($F=9.091$, $p=.004$). Additionally, we found that factual evidence was more effective in changing beliefs than was

consensus-based evidence which is also in line with our hypothesis, however this result was not significant (F=1.519, p=.223). Additionally, we found no effect of an interaction between the two variables (F=.029, p=.866).

We found no significant effect of the direction of evidence as the average change in belief for positive evidence was 12.05 (standard error = 1.16) and the average change in belief for negative evidence was 12.15 (standard error = 2.15).

In order to control for the possibility of floor or ceiling effects skewing the results of our experiment, we conducted a related analysis using log ratios.

For positive evidence, we used the equation:

$$Log\ ((FB + 1)/(IB + 1))$$

We added 1 to each initial belief in order to avoid dividing by 0 and we added 1 to each final belief in order to stay as close to the true ratio as possible.

For negative evidence, we used the following equation:

$$Log\ (100\ + (FB + 1)(1 - 100/(IB + 1)))$$

The equation for negative evidence represents the proportion that the final belief differs from the initial belief, added to the distance of the initial belief from 100. We used this specific equation because negative evidence should produce ratios of less than 1 which would produce negative log values.

In our analysis, we excluded the few instances (<3%) in which the subjects' responses went in the opposite direction from the expectation as these responses produced outlier values.

We found the following log ratio changes in belief after using these two equations:

**Table 2:**

| Mean Log Ratio (Standard Error) | Objective | Subjective |
|---|---|---|
| Factual | .133 (.022) | .113 (.016) |
| Consensus | .121 (.028) | .075 (.017) |

Again, these results suggest that individuals are more apt to change objective beliefs as compared to subjective beliefs and that factual evidence was more powerful in changing the minds of our subjects than was consensus evidence. These results follow the same pattern as those above using simple differences. As such, we proceeded with using simple differences for analysis in this experiment.

We calculated a correlation between the change in belief and expertise values for each of the 16 statements for each of the 200 subjects (3200 data points). We found a weak negative correlation of -.157, which suggests that individuals with more expertise on a given subject changed their beliefs less. This is logical and supports our prediction as one who already knows more about a given field may either already be aware of the novel evidence provided or might be aware of evidence that would disprove the novel evidence.

We calculated the correlation between credibility rating and change in belief and found a very weak correlation of .048. This value is likely so low because individuals who find evidence to be very credible may not change their beliefs at all if they are already aware of the novel evidence and have incorporated this knowledge into their initial belief rating. Additionally, if one

has an extreme initial belief (close to 0 or close to 100), then confirmatory evidence would not be able to move this individual's belief by very much, even if they found the evidence to be very credible. For example, if I have an initial belief of 100 in a given statement and then see evidence that confirms this view, I would still only be able to give a final belief rating of 100 even if I found the evidence to be very credible as the scale used was constrained from 0-100.

We calculated the correlation between credibility and initial beliefs as well. We predicted that people would find evidence that confirms their initial beliefs to be more believable. As such, we separated positive evidence questions from negative evidence questions.

We found that the correlation between positive evidence and initial beliefs was .500 ($t<10^{-18}$) and the correlation between negative evidence and initial beliefs was -.326 ($t<10^{-11}$). This confirms our prediction that people would use either normative or motivated reasoning in order to produce credibility ratings that align with their initial beliefs. Note that these t values are exceedingly small in part because of the large (1600) sample sizes.

Next, we calculated a correlation between the expertise ratings and extremity of initial beliefs (EIB). We used the following equation to find the extremity of initial belief values:

$$EIB = |IB - 50|$$

We found a positive correlation of .349, suggesting that those with more expertise on a given topic were more likely to have extreme initial beliefs. Interestingly, this may conflict with results from a 2019 study by Fernbach et al. which found that those with the least expertise are often the *most* extreme in their beliefs.

We did not find a meaningful correlation between EIB and change in belief (correlation = -.040). This suggests, perhaps unintuitively, that those with more extreme initial beliefs were

willing to update their beliefs to a similar degree as those who were less sure of their beliefs initially.

We did not find evidence of backfire effects in this experiment, which provides further support for the notion that backfire effects are particularly elusive and may only occur under specific conditions.

*Discussion*:

The results of Experiment 1 validated many of our predictions. First, we found that subjects were, in fact, more likely to update objective beliefs as compared to subjective beliefs. This may be due to the fact that objective beliefs are grounded more in data and less in emotion than subjective beliefs. Because subjective beliefs are often tied to emotion or even one's morals and values, evidence that opposes these beliefs may be seen as more threatening as it may challenge one's worldview and identity in certain cases. Thus, we may be more motivated to maintain our subjective views as we feel that they are more intrinsically linked to who we are. This may be an identity-protection mechanism.

Though our experiment indicated that factual evidence may be more likely than consensus evidence to change beliefs, the result was not significant ($F=1.519$, $p=.223$). We additionally did not find an interaction between the type of question and type of evidence ($F=.029$, $p=.866$).

This experiment also suggested that there is a weak negative correlation between expertise and changing one's beliefs. This may be because an individual with a high degree of expertise may already be aware of the evidence that is presented. If this were the case, then the

individual would have already presumably incorporated the evidence into his or her initial judgment and thus the evidence would be less effective in changing his or her mind. If the evidence opposes the expert's initial beliefs then he or she is likely aware of research that supports his or her position and thus the evidence would again be less powerful is shifting that individual's views.

The strong correlation between initial beliefs and credibility ratings (when separated into positive and negative evidence groups) suggests that individuals use either normative or motivated reasoning to make initial belief judgments. If these individuals are using motivated reasoning, then we would expect this relationship because people are motivated to confirm that evidence that agrees with their views is credible evidence. On the other hand, individuals may be using normative reasoning to arrive at these conclusions. It seems logical to be skeptical of evidence that conflicts with one's beliefs. For example, most people would likely be quite skeptical if they were told that aliens had landed a spaceship in New York City. As such, there is clearly more work to be done to clarify this distinction. Experiment 2 sought to address this question.

Finally, this experiment suggests that those with more expertise on a given subject have more extreme initial beliefs. This appears logical insofar as those with a high degree of expertise should be more sure of their views which may lead to extreme opinions. Similarly, it seems reasonable that one with a moderate level of expertise would not hold extreme views as he or she may not know enough about the subject to have a strong opinion. However, this result may conflict with well-known research conducted by Fernbach et al. (2019) which suggests that those who are the most sure of their beliefs often have the lowest degree of expertise. Of course,

extremity of initial beliefs and confidence in one's beliefs are not necessarily equivalent measures, so it is possible that both may be true. However, one would assume that there exists a strong correlation between extremity of beliefs and confidence in those beliefs. For example, I may be extremely confident that there is a 0% chance that aliens will land in New York City tomorrow but only moderately confident that I am a slightly above average soccer player. More research thus needs to be conducted to investigate this distinction. Experiment 3 investigates the related relationship between reflectiveness and belief updating.

## Experiment 2

*Introduction*:

In Experiment 1, we found that individuals are more likely to find evidence that confirms their initial beliefs to be credible. In Experiment 2, we sought to distinguish whether subjects were using motivated or normative reasoning when making credibility judgments. We used the paradigm established by Kunda (1987) in order to investigate this distinction. We predicted that if individuals were using motivated reasoning in order to make decisions about the credibility of novel evidence, then the distribution of responses would follow a predictable pattern. Those positively affected by the evidence would find the statement to be the most credible, while those negatively affected by the evidence would find the statement to be the least credible. Individuals who were unaffected by the evidence, we predicted, would produce intermediate credibility ratings.

*Methods*:

One hundred forty nine participants (44% female) were recruited using Amazon's Mechanical Turk and compensated a small amount for completing the study. Participants were given one of two unique versions of the study. They were given information about a given topic and asked to rate how believable they found the evidence to be on a scale from 0-100. Questions from Experiment 3 served as filler items for Experiment 2 and vice versa. Each of the questions was closely related to one of the statements from Experiment 1 in order to draw more direct conclusions about the results from Experiment 1. The topics covered included education, religion, and politics. See Appendix 2 for the specific questions from Form 1 given to subjects. In each question, the evidence suggested that either males or females were affected by the evidence provided while the opposite gender was unaffected. The evidence was also either positive or negative in nature which created a positively affected group, negatively affected group, and two unaffected groups. In each of the two versions of the study, the statements were equivalent, but the genders were reversed. Demographic information relevant to the experiment (gender, political affiliation, degree of religious identity, and type of high school attended - public or private) was collected at the conclusion of the study so as to avoid the possibility of stereotype threat or demand characteristics. Subjects who chose "other" as their gender identity were excluded from the analysis. As in Experiment 1, the evidence provided was created by the researchers for the purpose of the study. This was explained to the participants in a disclaimer following completion of the study.

*Results*:

The first prompt stated: "Brookings Institute conducted a study which found that male college students are 2.7x more likely to graduate college if they attended a private high school compared to a public high school. Female graduates of private high schools did not show a higher probability of graduating from college compared to female public school graduates. Rate the degree to which you find this evidence to be credible."

This created four groups of subjects:

Males who attended public high schools

Males who attended private high schools

Females who attended public high schools

Females who attended private high schools

However, because on half of the forms the gendered evidence was reversed such that females were affected by the evidence and males were unaffected, the groups are better described in the following manner.

Affected / public high school

Affected / private high school

Non-affected / public high school

Non-affected / private high school

The demographics are described in the chart below:

**Table 3:**

| | | |
|---|---|---|
| Affected by Evidence? | Affected 76 | Non-affected 70 |
| Type of School Attended | Public 125 | Private 21 |

The chart below summarizes the mean credibility ratings as well as the standard errors provided by each of the four groups.

**Table 4:**

| | Public School | Private School | Total |
|---|---|---|---|
| Affected | 49.88 (3.24) | 70.20 (7.33) | 52.55 (3.06) |
| Non-affected | 55.93 (3.08) | 53.73 (7.51) | 55.59 (2.83) |
| Total | 52.74 (2.25) | 61.57 (5.45) | 54.01 (2.09) |

As is evident in the chart, the affected/private school group showed the highest average rating of credibility at 70.20. The affected/public school group showed the lowest average rating of credibility at 49.88. The two non-affected groups showed intermediate ratings of 53.73 for the private school attendees and 55.93 for the public school attendees. This follows our prediction of the pattern that motivated reasoning should create.

We also conducted a two-factor ANOVA analysis of the data described above. The affected vs. non-affected variable was not statistically significant alone ($F=.781$, $p=.378$), nor was the type of school attended ($F=2.362$, $p=.127$). However, the interaction between the two variables was marginally significant ($F=3.652$, $p=.058$). This suggests that individuals were

motivated to believe evidence that positively affected them while subjects who were negatively affected by the evidence were motivated to doubt the credibility of the statement. In addition, we saw that those who were unaffected by the evidence provided intermediate credibility ratings. Thus, this first prompt produced the pattern of credibility ratings we would expect if subjects were using motivated reasoning.

The second prompt stated: "A recent study conducted by the economic think tank, National Bureau of Economics Research, found that when a Democrat is president, the unemployment rate is significantly lower for women ($p=.002$). The unemployment rate for men, however, is generally unaffected by the political party of the president. Rate the degree to which you find this evidence to be credible."

Because we included three possible political affiliations (Democrat, Republican, Independent), this question produced six possible groups:

Affected / Democrat

Affected / Republican

Affected / Independent

Non-affected / Democrat

Non-affected / Republican

Non-affected / Independent

The demographics are described in the chart below:

**Table 5:**

| | | | |
|---|---|---|---|
| Affected by Evidence? | Affected 70 | Non-affected 76 | |
| Political Affiliation | Democrat 72 | Independent 36 | Republican 38 |

The chart below summarizes the mean credibility ratings as well as the standard errors provided by each of the six groups.

**Table 6:**

| | Democrat | Independent | Republican | Total |
|---|---|---|---|---|
| Affected | 61.33 (3.33) | 46.28 (4.68) | 65.42 (7.39) | 58.16 (2.69) |
| Non-affected | 53.28 (4.33) | 54.89 (7.09) | 56.12 (5.10) | 54.63 (2.99) |
| Total | 57.75 (2.70) | 50.58 (4.25) | 59.05 (4.20) | 56.32 (2.02) |

The results produced by this question were puzzling. We would expect that the Affected Democrats would produce the highest credibility scores and the Affected Republicans and Affected Independents to produce the lowest credibility scores. However, though the Affected Independents produced the lowest average rating of 46.28, the Affected Republicans produced a higher mean rating than the Affected Democrats. This may be due to the small sample size of Republicans in this study, however the results are still confusing. Possible explanations for this result will be described in the following discussion section. We see that each of the three non-affected groups produced very similar average ratings, ranging from 53.28 (Democrats) to 56.12 (Independents), which is to be expected.

We also conducted a two-factor ANOVA analysis of the data described above. The affected vs. non-affected variable was not statistically significant (F=.449, p=.504), nor was the political affiliation variable (F=1.620, p=.202). The interaction between the two variables was not significant either (F=1.663, p=.193). This interaction did not occur in the predicted manner, so any interaction between the variables does not provide evidence for motivated reasoning. Additionally, it does not appear that subjects used normative reasoning to arrive at their credibility ratings. If they had, the Republicans would have, in all likelihood, found evidence suggesting a Democratic president is more likely to improve the unemployment rate to be uncredible. This, however, was not the case as we see Republicans actually produced a higher overall mean credibility rating than the Democrats.

The third prompt stated: “A new study conducted by researchers at Stanford University and published in the peer-reviewed journal "Topics in Social Science" found that, on average, men who identify as having a strong religious identity were more generous. Rate the degree to which you find this evidence to be credible.”

The four groups created by this prompt included the following:

Affected / Religious

Affected / Non-religious

Non-affected / Religious

Non-affected / Non-religious

The demographics are described in the chart below:

**Table 7:**

| | | |
|---|---|---|
| Affected by Evidence? | Affected 76 | Non-affected 70 |
| Religious or not? | Religious 57 | Non-religious 89 |

The chart below summarizes the mean credibility ratings as well as the standard errors provided by each of the four groups.

**Table 8:**

| | Religious | Non-religious | Total |
|---|---|---|---|
| Affected | 66.81 (3.80) | 54.57 (3.87) | 59.72 (2.82) |
| Non-affected | 69.68 (3.32) | 43.40 (3.41) | 52.79 (2.91) |
| Total | 68.07 (2.57) | 48.92 (2.63) | 56.40 (2.04) |

As is described in the chart, religious individuals rated the credibility of the statement as being much higher on average (68.07) as compared to non-religious individuals (48.92). However, we did not see the predicted motivated reasoning distribution of responses as the non-affected non-religious participants provided the lowest average rating (43.40). If subjects were using simple motivated reasoning, then the non-affected groups should have intermediate ratings of credibility while the positively affected group should have the highest and negatively affected group should have the lowest.

We also conducted a two-factor ANOVA analysis of the data described above. The affected vs. non-affected variable was not significant ($F=1.164$, $p=.282$). However, degree of

religious affiliation was a highly significant predictor of credibility rating (F=25.073, p=.000). The interaction between the two variables was marginally statistically significant (F=3.328, p=.070). However, this was not in the expected pattern of responses as the non-affected groups showed the highest (non-affected religious average = 69.68) and lowest (non-affected non-religious average = 43.40) ratings. As such, this does not suggest motivated reasoning taking place in the manner in which we predicted.

The fourth and final prompt stated: "The National Education Policy Center conducted an analysis of 500 public and private schools. It found that private schools provide a significantly more collaborative and supportive learning environment for female students compared to public schools. On the other hand, public schools provided an equally supportive and collaborative learning environment for male students. Rate the degree to which you find this evidence to be credible."

As in the first question, the four groups created by this prompt included the following:

Affected / public school

Affected / private school

Non-affected / public school

Non-affected / private school

The demographics are described in the chart below:

**Table 9:**

| | | |
|---|---|---|
| Affected by Evidence? | Affected 70 | Non-affected 76 |
| Type of School Attended | Public 125 | Private 21 |

The chart below summarizes the mean credibility ratings as well as the standard errors provided by each of the four groups.

**Table 10:**

| | Public School | Private School | Total |
|---|---|---|---|
| Affected | 55.81 (2.85) | 60.18 (8.02) | 56.50 (2.70) |
| Non-affected | 55.50 (3.00) | 64.80 (6.65) | 56.72 (2.76) |
| Total | 55.65 (2.07) | 62.38 (5.24) | 56.62 (1.93) |

These results suggest that the affected vs. non-affected variable was not predictive of average credibility ratings as the non-affected private school group produced the highest mean rating of 64.80. The public school attendees produced average ratings of 55.50 (non-affected) and 55.81 (affected), suggesting that this distinction was not relevant to making credibility judgments.

We also conducted a two-factor ANOVA analysis of the data described above. As expected, whether an individual was affected by the evidence was not predictive of credibility ratings ($F=.152$, $p=.697$). The type of school one attended was more predictive ($F=1.530$, $p=.218$), but still far from statistically significant. The interaction between the two variables was not significant ($F=.199$, $p=.656$). Thus, it is evident that this prompt did not produce the expected motivated reasoning results.

*Discussion*:

Of the four prompts, only one produced the predicted distribution of responses that would categorize motivated reasoning. In the first question, we saw that the positively affected group gave the highest credibility ratings, the negatively affected group produced the lowest credibility ratings, and the two non-affected groups provided intermediate scores. This follows the results of Kunda's study (1987) and suggests that at least for this particular question, motivated reasoning took place.

The second question, which involved political affiliation, produced results that are perhaps the most difficult to interpret. On one hand, we saw that the three non-affected groups (Democrat, Republican, Independent) all provided credibility ratings within a very tight range (53.28-56.12), which follows the predicted pattern. However, the affected groups provided very puzzling results. We saw that the Republicans provided the highest credibility ratings while Independents provided the lowest ratings. The low rating from the Independent group may be explained by the idea that Independents may have little faith in the ability of *either* major party to provide meaningful change. They may also be more skeptical of government and its efficacy in general. However, interpreting why Republicans, on average, did not produce lower credibility ratings than Democrats is difficult to say the least. There are a few possible explanations for this finding. On one hand, it could simply be that there was a small sample size of Affected Republicans (12) in the study, making the results difficult to generalize in any meaningful fashion. It is also possible, however, that some of these individuals are single-issue voters or social conservatives who are chiefly concerned with abortion, gun control, or immigration, for example, regardless of their opinions on economic issues. If this were the case, then a

Republican might vote against his or her perceived economic self-interest in the hopes of electing a president who will uphold his or her socially conservative values. A third possibility exists in which Republicans were tangentially aware of the purpose of the study and decided to inflate their credibility ratings on this question in order to seem logical and unbiased. We attempted to control for this by not asking about political affiliation until the conclusion of the study, but it is possible that these individuals foresaw this demographic question. Certainly more work needs to be conducted to examine this important distinction and thought-provoking result.

While the third question did not produce effects of motivated reasoning in the manner in which we intended, the results were significant and logical upon close examination. We found that the degree of religious identity that one holds was predictive of his or her credibility rating for the prompt provided, regardless of whether or not the individual was affected by the evidence ($F=25.073$, $p=.000$). This effect could be caused by a number of reasons. First, it is possible that religious people have more religious friends and family members who they view as generous, so they use the availability heuristic to arrive at their high credibility ratings. Similarly, it is possible that religious people want to believe that their faith is integral to making them a generous member of society, which leads them to provide high credibility scores. This question, it could be argued, *did* in fact produce evidence for motivated reasoning in a community-based, non-gendered manner. Though this is different from our prediction and the prediction from Kunda's study (1987), religious people may have wanted, in a sense, to believe that positive evidence about their *community*, not necessarily themselves, was true. However, it is difficult to determine which of these specific explanations is most accurate in describing the results

produced. A further investigation of this experiment may ask participants to elaborate on how they arrived at their given credibility ratings in order to explain this outcome.

The final question regarding the learning environment of public and private schools did not produce significant results, nor evidence of motivated reasoning. This is initially surprising as the first question, which did produce evidence of motivated reasoning, similarly covered the topic of public vs. private schooling. There is, however, a major distinction between these two prompts that may have created this disparity. In the first prompt, the evidence spoke of an objective measure, a graduation rate. The fourth prompt, on the other hand, covered a much more subjective topic. As such, it is likely that the results produced in the fourth question reflect participants' personal experiences in a manner that is not true of the first question. We intended to produce a result in which affected private school students would find evidence that their education was fostered by a supportive, collaborative learning environment to be credible. Conversely, we intended to produce a result in which affected public school graduates would find evidence that their schools were less supportive and collaborative to be less credible. This design has one major flaw, however. If a private school graduate found his or her school to be particularly unsupportive or uncollaborative, then he or she would be likely to provide a very low credibility rating, regardless of what the evidence provided suggests. Additionally, if a public school graduate found his or her school to be very unsupportive, then he or she would find this evidence to be quite credible, even though the evidence provided negatively affects his or her group. As such, we mistakenly assumed that all individuals would want to believe that their school experience was positive in nature, however this may not be the case for someone who desires confirmation that their negative experiences were not unique to him or herself. The first

question, however, avoids this issue by providing an empirical data point that would be difficult to conflate with personal experience. It is thus possible that this oversight in design is the cause of the lack of significant results for this specific question.

Overall, we found evidence of motivated reasoning in one of the four questions and found evidence of what we may categorize as *community-based* motivated reasoning in another. The two questions that failed to produce significant results or evidence of motivated reasoning had design flaws that may have prevented this from occurring. In one case, the nature of one's political affiliation may be too complex of a topic to produce clean, "logical" results. In the other, we did not account for the inevitable possibility that participants would incorporate personal experiences into their credibility judgments which likely rendered the results of this question unconvincing. This experiment suggests that while motivated reasoning does occur at times, the conditions under which this process takes place may be relatively specific in nature.

## Experiment 3

*Introduction:*

In Experiment 1, we investigated how the type of evidence and type of statement (subjective vs. objective) affected individuals' willingness to update their beliefs. Experiment 2 produced mixed results, however in at least one of the prompts we found evidence of motivated reasoning. As such, Experiment 3 sought to build on these findings by examining motivated belief updating. This experiment investigated how one's reflectiveness may relate to his or her belief updating practices. Additionally, this experiment evaluated whether or not the order in which different topics (neutral vs. charged) are presented may impact our judgments. This

experiment was motivated by Kahan (2017) whose research indicated that when evaluating emotionally charged topics, highly numerate individuals become more polarized in their beliefs. The present study used reflectiveness as a proxy for numeracy. We predicted that among highly reflective individuals, those who evaluated charged topics first would be less likely to change their beliefs. We predicted that evaluating neutral evidence first would lead subjects to use "cold" reasoning skills which would prime these individuals to apply this same approach when evaluating charged topics as well. When evaluating charged topics without this priming effect, we hypothesized that individuals would be less likely to update their beliefs as they may have a personal connection to these charged beliefs and thus should be motivated to uphold them. We also examined whether or not backfire effects would be present.

*Methods*:

One hundred forty nine participants (44% female) were recruited using Amazon's Mechanical Turk and compensated a small amount for completing the study. As in Experiment 1, subjects were asked to provide initial belief responses about four topics on a scale of 0-100. The four topics included cookies, pop music (neutral topics), abortion, and climate change (charged topics). Specifically, the questions asked whether cookies were the most popular American dessert, whether most Americans enjoyed pop music, whether a woman is better off after receiving an abortion, and whether climate change is the single most important issue of our time. Participants were given one of two unique versions of the study. In one version, participants evaluated neutral prompts first and in the other the order was reversed in order to counterbalance the sample. Questions from Experiment 2 served as filler items for Experiment 3 and vice versa.

We further counterbalanced the two forms of Experiment 2 with the two forms of Experiment 3 in order to assure that no crossover effects were taking place, thus creating four overall versions of the combination of experiments. After providing initial belief responses, new evidence was provided to the subjects and they were asked to re-evaluate their beliefs on the 0-100 scale. See Appendix 3 for the specific questions and evidence from Form 1 given to subjects. Subjects were then asked to complete the three question Cognitive Reflection Task (CRT), which measures individuals' ability to reflect upon initial judgments in order to limit errors based on a "gut" reaction. See Appendix 4 for the specific questions that comprise the CRT. As in Experiments 1 and 2, the evidence provided was created by the researchers for the purpose of the study. This was explained to the participants in a disclaimer following completion of the study.

*Results*:

Six participants were excluded from the analysis as they completed the entire experiment in under 2 minutes, which was deemed to be inconsistent with evaluating the evidence with sufficient care.

Questions with evidence intended to move participants' responses closer to zero (negative evidence) were reversed scored in the same manner as Experiment 1 using the following equation:

$$|FB - IB| = Difference\ in\ belief$$

where FB is final belief and IB is initial belief.

As such, a change from 50 to 80 in a question with positive evidence would be scored 30 in the same manner that a change from 50 to 20 in a question with negative evidence would be scored 30.

The following chart summarizes the average change in beliefs for each of the four topics covered:

**Table 11:**

| Neutral Topics - Mean and (Standard Error) | Charged Topics - Mean and (Standard Error) |
|---|---|
| Cookies: 5.15 (2.59) | Abortion: 7.38 (1.66) |
| Music: 17.50 (2.35) | Climate change: 8.94 (2.23) |
| Total: 11.32 (1.97) | Total: 8.16 (1.61) |

Subjects were categorized into one of four groups based on the order in which they saw the prompts and how they performed on the CRT evaluation. If a subject saw the prompts about neutral topics first they were classified as “neutral first” and if they saw the charged topic questions first they were classified as “charged first.” If a subject scored 2-3 (out of 3) on the CRT evaluation, they were classified as a “high” performer and if they scored 0-1 (out of 3) on the CRT evaluation, they were classified as a “low” performer. Thus, the four groups were as follows:

Neutral first / High CRT

Neutral first / Low CRT

Charged first / High CRT

Charged first / Low CRT

The demographic composition of subjects was as follows:

**Table 12:**

| CRT Performance | High 86 | Low 57 |
|---|---|---|
| Neutral or Charged first condition? | Charged 70 | Neutral 73 |

For the purposes of data analysis, we grouped responses to the two neutral items together and responses to the two charged items together.

The following chart summarizes the average changes in belief for the neutral first (both high and low CRT groups) and charged (both high and low CRT groups) conditions.

**Table 13:**

| Neutral First Condition - Mean and (Standard Error) | Charged First Condition - Mean and (Standard Error) |
|---|---|
| Neutral topics: 13.87 (2.80) | Neutral topics: 8.73 (2.80) |
| Charged topics: 10.38 (2.22) | Charged topics: 5.84 (2.37) |
| Total: 12.13 (2.17) | Total: 7.28 (1.97) |

Thus, it is evident that the group which evaluated neutral evidence first updated their beliefs to a greater extent for both neutral (F=1.1690, p=.196) and charged (F=1.969, p=.163) topics, regardless of CRT scores. These results are not significant, however. We also conducted an ANOVA analysis which found that the condition (neutral first or charged first) was marginally significant overall (F=3.590, p=.059) in predicting belief updating. The type of question (neutral or charged) was not statistically significant (F=1.556, p=.213), nor was the interaction between the condition and the type of question (F=.013, p=.908).

Similarly, the following chart summarizes the average changes in belief for the high CRT (both neutral and charged conditions) and low CRT individuals (both neutral and charged conditions).

**Table 14:**

| High CRT - Mean and (Standard Error) | Low CRT - Mean and (Standard Error) |
|---|---|
| Neutral topics: 15.24 (2.71) | Neutral topics: 5.49 (2.67) |
| Charged topics: 10.02 (2.24) | Charged topics: 5.34 (2.25) |
| Total: 12.63 (2.02) | Total: 5.42 (1.98) |

Thus, we see that across neutral (F=6.00, p=.016) and charged topics (F=2.00, p=.159), individuals with high CRT scores updated their beliefs more on average. However, the result is only significant for neutral topics. Thus, though Kahan predicts that more numerate people are better at self-deception on charged topics, we do not see significant evidence of this as high CRT

individuals updated their charged beliefs more, not less, on average. We do, however, see that overall high CRT individuals change their beliefs significantly more (F=7.775, p=.006). On the other hand, we do not see a significant effect of the type of question (F=1.075, p=.301) or interaction between the two variables - question type and CRT performance (F=.958, p=.328).

The following chart summarizes the total mean changes in belief (with standard errors) for each of the four groups.

**Table 15:**

| | Neutral First Condition | Charged First Condition | Total |
|---|---|---|---|
| High CRT | 17.85 (2.97) | 6.90 (2.45) | 12.63 (2.02) |
| Low CRT | 2.93 (2.09) | 7.82 (3.31) | 5.42 (1.98) |
| Total | 12.13 (2.16) | 7.28 (1.97) | 9.76 (1.48) |

The more important distinction arises when additionally including the neutral vs. charged question types to conduct an analysis. Thus, the chart below details the changes in belief for each of the four groups on both neutral and charged topics.

**Table 16:**

| Group | Neutral First Condition/ High | Neutral First Condition/ Low | Charged First Condition/ High | Charged First Condition/ Low |
|---|---|---|---|---|
| Mean Change in Belief for Neutral Topics (Standard Error) | 19.64 (3.78) | 4.59 (3.04) | 10.40 (3.60) | 6.36 (4.26) |
| Mean Change in Belief for Charged Topics (Standard Error) | 16.06 (2.94) | 1.27 (2.25) | 3.40 (2.99) | 9.28 (3.62) |
| Total Mean Change in Beliefs (Standard Error) | 17.85 (2.87) | 2.93 (2.09) | 6.90 (2.42) | 7.82 (3.20) |

We conducted one-factor ANOVA analyses to compare the following pairs of groups: When comparing high CRT individuals, we see that those who were exposed to the neutral items first updated their beliefs to a greater extent overall ($F=7.908$, $p=.006$). The result was marginally significant for neutral topics ($F=2.966$, $p=.0887$) and significant for charged topics

(F=8.658, p=.004). Among those in the neutral first condition, we see that those with a high CRT score updated their beliefs more (F=13.142, p=.0005). Interestingly, we see little difference in the total belief updating between high and low CRT scorers in the charged first condition (F=.052, p=.821). However, we see that those with a high CRT score updated their beliefs more on the neutral items (F=.502, p=.481) and less on the charged items compared to their low CRT counterparts (F=1.507, p=.224), though these results were not statistically significant. Finally, when comparing the low CRT individuals, we see that those in the charged first condition actually updated their beliefs on charged items more than those in the neutral first condition (F=3.299, p=.075), though this result was only marginally significant.

We also conducted a two-factor ANOVA analysis of the data described above. The charged vs. neutral first condition did not independently predict a significant pattern (F=1.104, p=.295). The high vs. low CRT score, however was a significant variable (F=5.904, p=.016) with high CRT individuals updating their beliefs more than their low CRT counterparts. Furthermore, the interaction between the two variables was significant (F=7.551, p=.007).

*Discussion*:

In this experiment, we found that among high CRT individuals, subjects in the neutral first condition updated their beliefs significantly more than those in the charged first condition (F=7.908, p=.006). This was especially true of the charged topics in which the average changes in belief were 16.1 for the neutral group and 3.4 for the charged group (F=8.658, p=.004). This follows our prediction as we expected that a sort of priming effect would occur in which individuals who evaluated neutral evidence first would use “cold” reasoning to evaluate both the

neutral and charged items. The high CRT individuals in the charged condition were not primed to use this emotionally independent reasoning technique as they did not evaluate neutral topics first.

We see that among those in the charged first condition, the total changes in belief were similar for the high and low CRT groups. Belief updating trends on the neutral and charged topics were not statistically significant.

We also found that among individuals in the neutral first condition, those with a higher CRT score updated their beliefs more ($F=13.142$, $p=.0005$). This suggests that perhaps high CRT individuals, on average, are more willing to incorporate novel evidence into their beliefs and judgments. This is supported by the ANOVA analysis which found that CRT scores were a significant predictor of belief updating ($F=5.904$, $p=.016$). Additionally, we found that high CRT individuals update their beliefs on neutral topics ($F=6.00$, $p=.016$) and charged topics ($F=2.00$, $p=.159$) more than low CRT individuals, though only the result for neutral topics was significant.

Kahan (2017) found that more numerate individuals are better at self-deception on charged topics as they became more polarized in their views. The present study suggests that high CRT individuals, on average, update their beliefs more than low CRT individuals, however the effect was only significant for neutral topics. It is possible that the CRT, which tests reflectiveness, is not in fact an adequate proxy for numeracy. However, if we assume that the CRT is a sufficient predictor of numeracy, then this experiment fails to replicate Kahan's findings. Further research should investigate this important remaining question.

One result that remains puzzling is why the neutral first, low CRT group changed their beliefs the least of all groups. When comparing the total belief updating of this group and the

charged first, low CRT group, we find a result that is not statistically significant (F=1.534, p=.221). However, when comparing the two low CRT groups, we see that the charged first group actually updated their beliefs on charged topics marginally significantly more (F=3.299, p=.075). This result is very confusing given the opposite effect among high CRT individuals (F=8.658, p=.004). This suggests that the priming system we proposed to explain the differences in charged belief updating among high CRT individuals may not hold true for low CRT individuals. It is challenging to speculate as to how this result may have arisen. One possible explanation is that among low CRT individuals, the topics evaluated first, be them neutral or charged, were given more attention and thus the changes in belief depended greatly on whether neutral or charged evidence was presented first. Certainly, further research should seek to explain this interesting phenomenon in which the patterns exhibited by high and low CRT individuals are opposite when evaluating charged topics first.

As in Experiment 1, Experiment 3 did not produce evidence of backfire effects. This further supports assertions that the backfire effect is either only produced under very specific conditions or is not a robust phenomenon capable of consistent replication.

## Discussion

The present study investigated the conditions under which we are more likely to update our beliefs as well as the approaches that we may use to make these decisions. Experiment 1 provided evidence that we are more likely to update our objective beliefs as compared to our subjective beliefs in the face of novel evidence, which confirms our prediction. Because our subjective views may be more intrinsically linked to our perception of who we are, this result

may suggest a motivated identity-protection mechanism. Additionally, this finding indicates that the way in which we frame topics may be relevant in persuading or convincing others of our opinions (or facts). For example, convincing a climate change denier of the perils of global warming may be more effective if one frames the issue through empirical facts and relationships between those facts as compared to emotional pleas for support. In practice this may look like detailing the overwhelming evidence that the earth is, in fact, warming and that this warming will lead to sea level changes that will be catastrophic for coastal cities. This technique may be more effective in changes one's mind than blanket statements about how the world will soon end if climate change is not addressed (Cummings 2018).

Experiment 2 provided mixed results, suggesting that instances of motivated reasoning may be specific to certain conditions. This study replicated Kunda's (1987) findings in only one of the four scenarios. However, another of the prompts (that which was concerned with religiosity) produced results that may be interpreted as evidence for *community-based* motivated reasoning. Because motivated reasoning can directly lead to confirmation bias, it is, in some sense, reassuring that we do not solely rely on this process when making credibility judgments about new information. However, motivated reasoning and confirmation bias do still present a significant risk to the acceptance of facts, both scientific and otherwise. For example, if an individual has a particular distrust for physicians and medicine, then he or she may be motivated to believe that vaccines cause autism, despite the wealth of research that suggests that this link is nonsensical. As such, it is important to further investigate the instances in which motivated reasoning is utilized as compared to the cases in which individuals use normative reasoning approaches to arrive at conclusions.

Experiment 3 provided evidence that contradicts Kahan's (2017) claim that more numerate individuals are better at self-deception. Our study found that individuals who performed better on the CRT were actually *more* likely to change their beliefs compared to low scoring individuals, though the result was only significant for neutral topics. Among these high scoring subjects, those who viewed charged topics first updated their beliefs to charged topics significantly less ($F=8.658$, $p=.004$). This suggests that there is a mechanism by which priming highly reflective individuals' normative reasoning systems may lead them to use this approach when evaluating emotionally charged topics as well. This is an important note for politicians or scientists hoping to convince the public or their peers of the validity of a controversial statement. For example, while 98% of scientists agree that humans did, in fact, evolve through evolution, nearly 40% of white evangelical Protestants are not convinced (Masci 2019). As such, it may be possible to convince some of these individuals by asking them to evaluate neutral topics in which there is near 100% consensus and then introducing the overwhelming scientific consensus that believes in evolution.

In both Experiments 1 and 3, we found no evidence of backfire effects, which is in keeping with the many failures to replicate Nyhan and Reifler's findings (see Haglin, 2017, Weeks et al., 2014, Garrett et al., 2013). This further suggests that backfire effects are highly elusive and may only appear under very specific conditions.

In sum, the ways in which we make decisions and judgments about the world around is a complex process that involves both motivated and normative reasoning approaches. In order to foster a world in which everyone can at least agree upon the same set of facts, it will be important to continue to investigate the conditions under which we are able to elicit

emotionally-independent reasoning approaches. This could provide a first step towards bridging the divide between the two realities in which we live today.

**References**

Cummings, W. (2019, January 22). 'The world is going to end in 12 years if we don't address climate change,' Ocasio-Cortez says. Retrieved from https://www.usatoday.com/story/news/politics/onpolitics/2019/01/22/ocasio-cortez-climate-change-alarm/2642481002/

Cunningham, W. & Raye, C. & Johnson, M. (2005). Implicit and Explicit Evaluation: fMRI Correlates of Valence, Emotional Intensity, and Control in the Processing of Attitudes. Journal of cognitive neuroscience. 16. 1717-29.

Dawson, E., Gilovich, T., & Regan, D. T. (2002). Motivated Reasoning and Performance on the Wason Selection Task. *Personality and Social Psychology Bulletin*, *28*(10), 1379–1387.

Doll, B.B., Hutchison, K., & Frank, M.J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *The Journal of neuroscience: The Official Journal of the Society for Neuroscience, 31 16*, 6188-98.

Fernbach, P. M., Light, N., Scott, S. E., Inbar, Y., & Rozin, P. (2019). Extreme opponents of genetically modified foods know the least but think they know the most. *Nature Human Behaviour, 3*(3), 251-256.

Garrett R.K., Nisbet E.C. and Lynch E.K. (2013). Undermining the corrective effects of media based political fact checking? The role of contextual cues and nave theory. *Journal of Communication* 63(4): 617–637.

Gilovich, T. (1991). How we know what isn't so: The fallibility of human reason in everyday life. New York: Free Press.

Haglin, K. (2017). The limitations of the backfire effect. *Research & Politics, 4*(3), 1-5.

Jones, B. (2019, February 05). Republicans and Democrats have grown further apart on what the nation's top priorities should be. Retrieved from http://www.pewresearch.org/fact-tank/2019/02/05/republicans-and-democrats-have-grown-further-apart-on-what-the-nations-top-priorities-should-be/

Kahan, D. (2015). Climate-Science Communication and the Measurement Problem. *Advances in Political Psychology* 36, 1-43.

Kahan, D.M., Jenkins-Smith, H. & Braman, D. (2011). Cultural Cognition of Scientific Consensus. *Journal of Risk Research.* **14**, 147-174.

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

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing ones own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Kunda, Z. (1987). Motivation and inference: Self-serving generation and evaluation of evidence. Journal of Personality and Social Psychology, 53, 636-647.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin, 108 3*, 480-98.

Litman, L., Robinson, J., & Abberbock, T. (2016). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. Behavior Research Methods, 1 - 10.

Masci, D. (2019, February 11). For Darwin Day, 6 facts about the evolution debate. Retrieved from http://www.pewresearch.org/fact-tank/2019/02/11/darwin-day/

Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior, 32*(2), 303-330.

Sloman, S. A., & Fernbach, P. (2018). *Knowledge illusion: Why we never think alone*. New York, NY: Riverhead Books.

Weeks B.E. and Garrett R.K. (2014). Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 US presidential election. *International Journal of Public Opinion Research* 26(4): 401–422.

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election. *Journal of Cognitive Neuroscience,18*(11), 1947-1958.

Wyer, R. S. & Frey, D. (1983). The effects of feedback about self and others on the recall and judgments of feedback-relevant information. Journal of Experimental Social Psychology, 19, 540-559.

Yousif, S. & Keil, F. (2018, July). *When in doubt: Using confidence and consensus as 'summary statistics' of collective knowledge*. Poster session presented at the 40th Annual Meeting of the Cognitive Science Society, Madison, WI.

## Appendix 1

In a recent survey, 24 out of 32 NFL owners ranked the Patriots as above average in honesty. Given this information, rate the degree to which you believe the New England Patriots NFL football team to be an honest franchise that plays by the rules.

According to the market research company Statista, the United States' Gross Domestic Product (GDP) was over $19 trillion in 2017, leading second place China by about $7.5 trillion. Given this information, rate the degree to which you believe the United States to have the largest economy in the world.

An independent research team conducted a study that found that women who received an abortion are not better off financially or in terms of mental well-being as compared to women who considered an abortion before ultimately deciding against it. Given this information, rate the degree to which you believe a woman is better off after an abortion.

Among a survey of 364 economists, 272 stated that the unemployment rate does worse when a Democrat is President. Given this information, rate the degree to which you believe the unemployment rate does better (is lower) when a Democrat is President.

Among a survey of 280 anthropologists, 207 stated that religious people, on average, are more generous than their nonreligious counterparts. Given this information, rate the degree to which you believe being a religious person makes one, on average, more generous.

According to comScore TV Essentials, the viewership of the three major cable news networks (CNN, Fox News, and MSNBC) increased 12% in 2017. Given this information, rate the degree to which you believe viewership of cable news networks to be increasing.

An independent research team conducted a study of 28 charter schools and 28 public schools and concluded that public schools provide a more supportive, collaborative learning environment, citing teacher engagement, classroom atmosphere, and curricular patterns. Given this information, rate the degree to which you believe charter schools provide a more supportive, collaborative learning environment for students compared to traditional public schools.

Among a survey of 229 climate scientists, 167 stated that the CO2 levels are not the highest they have been in three million years. Given this information, rate the degree to which you believe the carbon dioxide levels in the atmosphere to be the highest in three million years.

Since 2007, the Yankees have won only one World Series title, while the San Francisco Giants have won three and the Red Sox have won two. Given this information, rate the degree to which you believe the New York Yankees to be the winningest MLB team since 2007.

Among a survey of 332 anthropologists, 251 ranked the United States to be outside the top 25 most free countries in the world. Given this information, rate the degree to which you believe the United States to be one of the most free countries in the world.

Among a survey of 312 reproductive biologists, 236 stated that a fetus could survive outside of the womb at 20 weeks gestational age. Given this information, rate the degree to which you believe a fetus to be a viable life form before 20 weeks gestational age.

An independent research team conducted an analysis of wage patterns and cost of living in America. They concluded that the average American worker is financially secure, with 67% of Americans being secure or extremely secure. Given this information, rate the degree to which you believe the average American worker is financially secure.

According to a recent Gallup poll, the percentage of Protestant Americans who attend church at least once a week has remained between 44% and 45% over the last decade. Given this information, rate the degree to which you believe church attendance has decreased in the United States in the last decade.

Among a survey of 47 media specialists, 35 rated presenting a representative sample of what is going on in the world as falling outside of the top 3 goals of the news media. Given this

information, rate the degree to which you feel the news media's main goal is to present a representative sampling of what is going on in the world.

Among a survey of 81 education experts, 59 stated that low-income charter schools graduates are more likely to go on to graduate from college. Given this information, rate the degree to which you believe low-income charter school graduates are more likely to go on to graduate from college.

An independent research team concluded that climate change is the most important global problem today, citing the impact it may soon have on food availability, stock market volatility, infrastructure damage, and conflict for water, among other factors. The research team ranked it 1st out of 25 global problems, ahead of terrorism, illegal immigration, and violent crime. Given this information, rate the degree to which you believe climate change to be one of the three most important issues of our time.

## Appendix 2

Brookings Institute conducted a study which found that male college students are 2.7x more likely to graduate college if they attended a private high school compared to a public high school. Female graduates of private high schools did not show a higher probability of graduating from college compared to female public school graduates. Rate the degree to which you find this evidence to be credible.

A recent study conducted by the economic think tank, National Bureau of Economics Research, found that when a Democrat is president, the unemployment rate is significantly lower for women (p=.002). The unemployment rate for men, however, is generally unaffected by the political party of the president. Rate the degree to which you find this evidence to be credible.

A new study conducted by researchers at Stanford University and published in the peer-reviewed journal "Topics in Social Science" found that, on average, men who identify as having a strong religious identity were more generous (pRate the degree to which you find this evidence to be credible.

The National Education Policy Center conducted an analysis of 500 public and private schools. It found that private schools provide a significantly more collaborative and supportive learning environment for female students compared to public schools. On the other hand, public schools

provided an equally supportive and collaborative learning environment for male students. Rate the degree to which you find this evidence to be credible.

## Appendix 3

A study conducted by Pew Research found ice cream overwhelmingly to be the most popular dessert in America based on purchase frequency and likability ratings. Given this information, rate the degree to which you believe cookies to be the most popular dessert in America.

A recent Gallup Poll found that, ironically, only 37% of Americans enjoy listening to "Pop"/Top-40 music. Given this information, rate the degree to which you believe that most Americans enjoy "Pop"/Top-40 music.

According to a recent study conducted by the National Institutes of Health (NIH), women who received an abortion are better off financially after the procedure as compared to women who considered an abortion before ultimately deciding against it. Additionally, women who received an abortion reported, on average, a higher degree of mental well-being compared to women who decided against the procedure. Given this information, rate the degree to which you believe a woman is better off after an abortion.

A recent study conducted by the International Public Policy Group (IPPG) concluded that while climate change is an important issue, there are more pressing problems for overall global prosperity. These include terrorism, political polarization, illegal immigration, income inequality and violent crime. The research group ranked climate change 7th out of 25 global problems in

terms of importance. Given this information, rate the degree to which you believe climate change to be the single most important issue of our time.

## Appendix 4

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? _____ cents

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days