**Self- Awareness Processes in Motivated Reasoning**

by

Clint McKenna

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2024

Doctoral Committee:

        Professor David Dunning, Chair
        Associate Professor Allison Earl
        Professor Colleen Seifert
        Associate Professor Brian Weeks

Clint McKenna

cmck@umich.edu

ORCID iD:  0000-0002-9128-3581

## Dedication

This dissertation is dedicated to my mother, for the lifetime of love she has given me.

# Acknowledgements

I wish to thank my advisor Dr. David Dunning. Your perpetual support and endless patience over my time at Michigan will never be forgotten. To my advisor Dr. Allison Earl, thank you for the years of support and kindness, and challenging me to be better intellectually. To my committee members, Drs. Colleen Seifert and Brian Weeks, thank you for dedicating many hours of your support and expertise to help me in my research endeavors.

To my partner Susannah, thank you for all your love and support. You have nurtured my mind and guided me back to the light many times. To my cohort at Michigan: Sakura Takahashi, Nadia Vossoughi, Martha Berg, and Imani Burris, thank you for making me feel so welcome as a colleague still finding my way. To my labmates and friends at Michigan, especially Yuyan Han, Julia Smith, Zachary Reese, Qinggang Yu, Kaidi Wu, Koji Takahashi, Irene Melani, and Wilson Merrell, thank you for supporting me throughout the years and I am so proud to have known you all. To my research assistants, especially Ashley Anderson, Claire Goods, Heather Zimmerman, Carolyn Watson, Cate Madden, and Gabe Porter, thank you for all your hard work on our projects; it has been a privilege to mentor you all. To my mentors while I attended UC Irvine, especially Drs. Peter Ditto, Sarah Pressman, and Joey Cheng, thank you for the generous time and many opportunities you have given me over the years.

Finally, I would like to thank my dogs, Kubrick and Alaska, and cat Yoshimi, who have helped me get through some of the toughest years of my life.

# Table of Contents

# List of Tables

# List of Figures

# List of Appendices

**Abstract**

People have limited introspective access to many mental processes. In this dissertation, I test

whether people can report self-awareness of motivational influences. Specifically, I investigate

the phenomenon of motivated reasoning, where people's decisions and evaluations of

information are biased in favor or one's attitudes, beliefs, and identities. In six empirical studies,

people show limited ability to report the extent that their prior beliefs influence their reasoning.

In Studies 1 – 3, I demonstrate that people update their beliefs consistent with their motivations,

but do not report awareness of this influence consistently, or in a manner that aligns with a peer's

perception of them. In Studies 4-5, I show similar effects with a simulated political ballot

counting task. Awareness was generally underreported relative to peers, but those who were

greatly biased tended to report more perceived influence of political beliefs. Finally, in Study 6, I

introduce an incentive-compatible reward task that maximizes the likelihood of honest responses.

In this study, people evaluate the political ballots honestly, eliminating a possible counter-

explanation about normative views of motivated evaluation. In sum, this dissertation advances

our understanding of metacognitive processes related to motivated reasoning. I find that while

people engage in motivated reasoning, they often underreport awareness of how their desires are

influencing their decisions.

**CHAPTER 1**

**Introduction**

Does caffeine intake impact sleep quality? This may be a straightforward question. However, the response may depend on *who* is answering the question. For example, imagine that Stanley, a heavy coffee drinker, is asked about how caffeine consumption relates to sleep. He might spontaneously construct a lay theory about how caffeine reduces his likelihood of negative health outcomes. He might access relevant memories that justify this theory, like how coffee allowed him to wake up early enough to go to the gym before work last week, or how the green tea he drank this morning was filled with various nutrients and antioxidants. Likewise, Stanley might read an article about a study on caffeine and sleep, but only remember information that bolsters my theory, discarding information that invalidates it or scrutinizing the flaws in the study's methodology.

This illustration of Stanley's biased mental processes fall under the umbrella of *motivated reasoning* (Kunda, 1990). In motivated reasoning, these biases may unintentionally manifest through accessing self-serving cognitions (Dunning et al., 1995; Kunda, 1987), the biased assimilation of new information (Lord et al., 1979), and the skeptical evaluation of undesired information (Ditto et al., 1998; Ditto & Lopez, 1992). All these biases serve the function of reducing dissonant attitudes that hinder the optimistic desire for one to feel that good things will happen to me and bad things will not (Festinger, 1957; Kunda, 1990).

Consider Kunda's (1987) research about the biased evaluation of evidence. In one study, participants were given an article about the link between caffeine consumption among women and fibrocystic disease, related to painful breast lumps. Women, but not men, who read about this health issue demonstrated much more skepticism about the evidence, especially if they were heavy as opposed to low caffeine consumers. The importance (value relevance or social identification) of an attitude is critical in prompting motivational processes (Boninger et al., 1995). Furthermore, Kunda argues, these self-serving beliefs are motivational in nature, and not just based on prior belief estimates that happen to differ between groups.

To date, much of the research on motivated reasoning largely focuses on the underlying psychological processes that lead to differential outcomes as the result of people's prior attitudes, identities, or experiences. What is less understood is if people can report any self-awareness of motivated reasoning. Furthermore, it is currently unknown if people can accurately identify if and when motivated reasoning influenced their attitudes and actions. In summary, while past research has examined whether people have awareness of their mental processes in general and the processes behind motivated reasoning, there is a research gap: are people self-aware of their own motivated reasoning processes? In this chapter, I review the relevant psychological literature on self-awareness, attitudes, and motivated reasoning to better understand the gaps in knowledge of self-awareness processes in motivated reasoning, broadly defined. In the chapters that follow, I present six empirical studies that quantify self-reported awareness in two motivated reasoning paradigms. Additionally, I investigate possible normative and impression management explanations for discrepancies in self-awareness.

**Self-Awareness in Reasoning**

What does it mean to be aware or unaware of some cognitive process? Bargh (1994) describes three possible scenarios: awareness of stimuli, awareness of how stimuli are categorized or interpreted, and awareness of how stimuli influence judgments, states, and attributions. The present research focuses on the last scenario, where some stimulus influences higher reasoning about various aspects of self and social life. Following this, I discuss self-awareness literature in reasoning, generally and within motivated reasoning specifically.

**Awareness in General Reasoning**

People lack the ability to report awareness about their cognitive processes, dissonance reduction, or attitude changes (Bem & McConnell, 1970; Nisbett & Wilson, 1977). For instance, Cohen (2003) examined political party influence on attitude change, finding that people based their reasoning on the content of policy proposals, but when party information was present they generally aligned with the position of the party. More interestingly, participants denied the influence of party information and maintained that their attitudes were informed by the content of the proposal. People at the very least do not report awareness consistently when they are seemingly biased by their motivations.

Consider once again the case of Stanley, the caffeine consumer. When Stanley thinks about caffeine, he automatically activates a mental concept of coffee and positive evaluations. This attitude is solidified by years of positive associations between drinking coffee and feeling well; he can easily recruit these attitudes, memories, and beliefs about caffeine. As they are necessary to cognitive processes, these attitudes become more salient and accessible (Sanitioso et al., 1990). Indeed, attitudes generally may be thought of as object-evaluation associations (Fazio,

1995), and can be formed outside of conscious awareness (De Houwer et al., 1997; Olson & Fazio, 2001, 2002).

Do all attitudes work this way? While many attitudes are explicitly accessible through introspection, psychologists have argued that there are also implicit attitudes that are mostly observed through indirect measures rather than self-report (Greenwald et al., 2002; Greenwald & Banaji, 1995). Many studies describe these implicit attitudes as entirely unconscious, but people do appear to have access to them (Hahn et al., 2014; Hahn & Gawronski, 2019; Hahn & Goedderz, 2020). Hahn (2014) and colleagues had participants make explicit predictions about an implicit association task, where quick and automatic associations between preferred and nonpreferred social groups were made. The authors found robust evidence that explicit predictions were predictive of implicit attitude results. Morris and Kurdi (2023) further investigated this, finding that people predict their implicit attitudes even across different experimental tasks and attitude targets. The authors also found that people do not merely infer implicit attitudes from other factors like the knowledge of their demographic identities. Thus, although there are some attitudes that are not as easily accessible as explicit attitudes, one can report them through some combination of introspection and inference.

### Metacognition

What about when attitude objects are our own cognitions, as when Stanley thinks about his thoughts towards caffeine? We can describe this concept as *metacognition* (Petty, Briñol, Tormala, et al., 2007). Metacognition might be thought of as awareness of the presence (and self-evaluation) of these cognitions, but also awareness of the cognitive processes—the mental events that "add to, alter, or act upon representations in memory with detectable consequences"—that led to their formation (Wegener & Carlston, 2005). At least some cognitive processes are not

detectable: people lack awareness of reactance, for instance (Albarracín & Vargas, 2010). Metacognition is especially important when thinking about the validity of one's thoughts (Wegener et al., 2018), and is more likely when one is already motivated to think about this validity (Petty, Briñol, & DeMarree, 2007). As discussed in the following section, in the case of motivated reasoning, people appear to lack the motivation to evaluate the validity of their reasoning.

**Awareness in Motivated Reasoning**

Kunda (1990) describes how motivated reasoning can be driven by accuracy (a goal of arriving at a correct conclusion) or directional (a goal of arriving at a desired conclusion) motives. In some cases, accuracy motives can lead to the reduction of cognitive bias (Freund et al., 1985; Kruglanski & Freund, 1983; Pennycook & Rand, 2019; Ross et al., 2021; Tetlock, 1983). In other cases, we are driven by directional motives. As an example, Kahan (2012) investigated the political motives of factual assessment of a protest video, extending earlier work by Hastorf and Cantril (1954). Depending on whether the protest was described as an abortion or military recruitment center, people inferred different facts about the behavior of the crowd, like whether protestors blocked access or intimidated others. When assessing other factual evidence like numerical data, the presence of motivational cues also tends to inhibit accuracy (Kahan et al., 2017; Nurse & Grant, 2020; Persson et al., 2021; Van Boven et al., 2019). Thus, motivated reasoning may not necessarily prompt metacognitive evaluations of validity.

If not validity, what about other signals of awareness? Motivations can influence reasoning as early as the perceptual stages, and generally operate outside of awareness. For instance, individuals are capable of consciously allocating their visual attention towards a

stimulus (Meng & Tong, 2004), but processing desirable vs. undesirable stimuli may be constrained by motivational and environmental cues (Balcetis & Dunning, 2006, 2010).

It seems that when people are evaluating new evidence, they show limited access to introspection. There is evidence that people display a blindness to their levels of bias in reasoning or confidence (Dunning, 2012; Kruger & Dunning, 1999; Pronin et al., 2002; Scopelliti et al., 2015; Wilson, 2002; Zell et al., 2020), especially when comparing their own bias to other people (Pronin & Kugler, 2007; Van Boven et al., 1999). Through a process of selectively exploring new information and evidence (Hart et al., 2009), people shift their beliefs towards ones they find desirable (Darley & Gross, 1983). Darley and Gross (1983) found that people selectively recall, provide more weight to, and imbue additional meaning to evidence consistent with their expectations. All this dissonance reduction seems to also happen automatically, without conscious deliberation. For instance, Lieberman et al. (2001) found that amnesiacs incapable of retaining new memories will report attitude shifts similar to non-amnesiacs in a cognitive dissonance paradigm.

However, is it fair to say that motivated reasoning necessitates unawareness? Rosenzweig's (2016) review of this literature concludes that that it is not necessary to be unaware of a motivated process in order to benefit from one. This seems at odds with some literature on how awareness may dampen effects of awareness on self-affirmation (Sherman et al., 2009), biases in arbitrary decision making (McCaslin et al., 2010), or mood effects on well-being (Schwarz & Clore, 1983). Rosenzweig (2016) articulates that even if one is aware of the product of motivated reasoning, people may be blind to the process itself.

Schwarz & Clore's (1983) provide additional evidence to support this conclusion. In their studies, they explored motivational processes behind beliefs about how weather impacts mood.

In the first condition, participants were asked about the weather and explicitly told that the researchers were interested in the effects of weather on mood. In this condition, mood and life satisfaction were higher if bad weather could be blamed for their prima facie mood that would have been lower in a control condition where weather was not mentioned.

In the second condition, participants were primed to be aware of the weather and that there may be a link between the weather and mood. However, they were not necessarily aware that this link had any effect on them or that it provided a convenient reframing of what they likely wanted to believe. If they were, we might see participants lower their mood or life satisfaction when it was sunny out, since they would rationally correct for this bias in reasoning. Instead, if it were sunny outside, one could be aware that the weather is affecting my mood but might rationalize why this relation does not apply to them or that there might be some other process affecting their responses. They may even use the desired outcome (increased mood) as *evidence* that since the weather was consistent with their ideal state, no biased reasoning is afoot (Bem & McConnell, 1970). In sum, people can be aware of biased mental processes but may be blind to the fact that the conclusions they ultimately arrived at were a result of their influence.

To date, there is limited direct research on awareness during motivated reasoning. One exception is a recent investigation by Cusimano and Lombrozo (2023). The authors test whether people might appear "biased and proud" on some moral issues. This view suggests that rather than being oblivious to how evidence affects moral judgment, people will knowingly maintain their beliefs in spite of evidence (Cusimano & Lombrozo, 2021; Tetlock, 2002). Cusimano and Lombrozo (2023) showed that when evaluating a moral proposition, people were more likely to believe the propositions that were consistent with their beliefs and felt more justified in doing so. Additionally, participants evaluated evidence that was either moralized or not. Moralized

7

evidence prompted discounting of evidence and more awareness that participants were discounting evidence, as measured by a self-report response to how much their concerns about the evidence results affected their decision. This paper seems to contradict much of the research on how people are limited in their own metacognitive evaluations of their own beliefs.

I propose that we are aware of cognitive processes to the extent that these meta-beliefs about our own biases are detectable or noticeable during a reasoning or evaluation task. When prompted, one should be able to reflect on their own metacognition and have some understanding of how much their motivations are driving their reasoning of behavior. I will argue that in the case of motivated reasoning however, it is often in our best interest to make sure they remain undetectable.

**The Function of Unawareness**

Why might people be unaware that they engage in motivated reasoning? Are there any advantages, psychological or otherwise, from a lack of self-knowledge? I propose that there are functional ends met by constraining awareness or attention to cognitive processes that may be influenced by one's prior attitudes, beliefs, and identities.

Attitudes have long been seen as helpful in meeting functional goals. Katz (1960) articulated several functions of attitudes, including adjustment (regulating external rewards and punishments), ego defense (removal of conflicts or threats), value expression (enhancing communicating a positive self-image), or knowledge (appraisal of an attitude object). Many of these functions help people when engaging in motivated reasoning. Consider once again the case of Stanley, the coffee consumer. If Stanley wants to avoid cognitive dissonance, suppressing dissonant attitudes about his coffee consumption (or adding consonant cognitions) may help prompt a self-evaluation of how his seemingly incoherent behavior is acceptable. Similarly,

8

suppose that a political leader that Stanley admires was caught up in a tax scandal. He might express support for the politician's continued dedication to family values to express his moral superiority. The framework introduced by Katz is a convenient way of describing the benefits of maintaining our self-concepts through cognitive defenses.

A cognitive defense might manifest via knowledge avoidance, such as avoidance of comprehension, attention, or exposure (Greenwald, 1997). To take the example of a heavy coffee drinker evaluating the caffeine-health relation, Stanley may consciously be aware of a magazine article critiquing caffeine as unhealthy, but it would be in his best interest to obscure his comprehension by scanning paragraphs quickly, or his attention by turning on the television as he reads. In this case, Stanley is certainly aware of the article and content of the arguments within, but he might not be privy to the cognitive process of avoiding deliberate evaluation. This process is similar to what is called *bounded awareness*, where people demonstrate an inattentional blindness or failure to focus on an important issue (Bazerman & Sezer, 2016; Chugh & Bazerman, 2007).

As an extension of the ego defense motive, people also tend to align socially with those that share consonant (as opposed to dissonant) beliefs (Golman et al., 2016). This preference for belief consonance may be because the presence of adversaries can prompt reevaluation of one's own beliefs, which itself is uncomfortable from a cognitive dissonance perspective. Self-selection into social groups based on beliefs is ideal from this dissonance reduction perspective. Furthermore, one can rationalize this behavior as a way of improving subsequent reasoning: if someone is never confronted with untruths from outgroups, they can be more certain that the logic of their own beliefs is sound (Cushman, 2020). As a further possibility, people may also deliberately fail to rehearse thinking about threatening information in order to forget it over time

(Bénabou & Tirole, 2002). By selective forgetting, this likely avoids any dissonant thoughts about one's identity as an impartial thinker.

Is this avoidance of information pervasive? Golman et al. (2017) review the literature on information avoidance. Although it is counterintuitive to not seek out as much information as possible for better decision making, many people will avoid acquiring new knowledge. This behavior can manifest in a number of ways, including physical avoidance of information, inattention, biased evaluation of evidence, or selective forgetting (Golman et al., 2017). In a classic example, a rational actor would benefit from improved health or knowledge about their body, but people are often reluctant to engage with such information (Earl et al., 2009; Earl & Albarracín, 2007; Lerman, 1996). Presumably because of not wanting to think about threats, one might disengage entirely from the possibility of learning uncomfortable truths about themselves.

People lacking information in memory, such as an organizing schema or mental concept, may also benefit from the lack of discomfort in social privilege. For instance, a man (but not a woman) might receive deference at an auto mechanic (Wu & Dunning, 2018, 2020). As a man, being unaware of the concept of male privilege means that one is not burdened by the knowledge that a woman might be treated differently in the moment. When people do engage in biased seeking of information, they do so even arbitrarily, perhaps in anticipation of enjoyment at being correct (Scherer et al., 2013). For instance, investors tend to be biased towards recruiting memories of financial decisions as more lucrative than reality (Walters & Fernbach, 2021). Much of information avoidance thus seems to come from anticipated dissonance reduction.

In interpersonal negotiations, it can also be advantageous to avoid information (Auster & Dana, 2021). Concealment of information from oneself can help people avoid discomfort at profiting from another's loss (Dana et al., 2007). It can also prevent the possibility of leaking

10

information to another party, hindering persuasion attempts. (Smith et al., 2017; Trivers, 2011; von Hippel & Trivers, 2011). In bargaining games, asymmetric information between two parties can also be an advantage, even leading a party to deliberately avoid gaining information (Hermalin & Katz, 2009; Lau, 2008; Rogerson, 1992). More information also can lead to divergence in perceptions and expectations of fairness, which might be best to avoid in situations where parties must arrive at a mutual decision (Babcock et al., 1995). In sum, there are many psychological and economics benefits conferred by the absence of knowledge in social domains.

From the evidence, it is clear that people not only have limited introspective access to their own psychological processes, but often will willingly avoid awareness. It is also apparent that people may benefit from this lack of awareness, either psychologically or materially. Whether or not people can access the state of their own motivated reasoning processes remains unclear, but it is likely that a state of unawareness would maximize social fitness in several ways.

**Dissertation Overview**

People may be unaware of the extent to which their reasoning is affected by their own desires. There is some difficulty however, in quantifying this awareness. Much of the research outlined above does not explicitly ask whether participants were influenced by motivations when making a judgment, but only infers this by drawing a participant's attention to stimuli or a stimulus-influence link. A parsimonious method of investigating this is to simply ask: if participants can reliably report awareness of a desire affecting their judgments, we should see greater awareness reported when a decision is consistent their motivations, but lower awareness when a decision is inconsistent with one's motivations (directional goals). A second question in understanding awareness effects is whether a lack of reported awareness can be explained away by normative claims about what biased behavior should look like. That is, if people maintain a

11

lay theory about when motivations should bias one's judgment, any patterns in self-reported awareness should be consistent with what it means to be "aware" of motivated reasoning. My dissertation seeks to empirically investigate these in the following chapters.

In Chapter 2, I present three empirical studies that investigate self-awareness and motivated reasoning in a belief updating task. In both general beliefs about everyday life and political beliefs about an upcoming election, participants reported initial beliefs about a future event and were given the opportunity to update their beliefs. While people display motivated belief updating, they do not display patterns of awareness when considering their own biases. However, participants evaluating another person's responses do infer motivated biases that are not reported by the reasoner.

Chapter 3 investigates motivated evaluations of political ballots in a simulated ballot counting task across two studies. People again engage in biased evaluation based on their political leanings, but display small patterns of awareness, generally driven by extremely biased respondents. Republicans, relative to Democrats, evaluated ballots in a more biased manner, but also displayed more awareness. Peer participants also infer more political biases in this task relative to oneself.

In Chapter 4, I introduce an incentive-compatible procedure for detecting authentic beliefs and apply this to Chapter 3's ballot paradigm. I use the procedure to rule out the possibility of inauthentic responses being influenced by normative expectations of a ballot's acceptability. That is, if people provided evaluations of ballots that deviated from how they expected others to evaluate the same ballots, any subsequent awareness measures could be unreliable as an honest indication of perceived influence. In one study, this incentive task reveals similar behavior as those in Chapter 3.

Finally, Chapter 5 presents a discussion of this dissertation research and future directions in understanding the role of self-awareness in motivated reasoning.

# CHAPTER 2
## Awareness of Motivated Belief Updating

## Study 1

In Study 1, we examined how people update their beliefs differentially based on whether an event is "good news" (a desired outcome relative to currently held beliefs) or "bad news" (an undesired outcome relative to currently held beliefs). This study was modeled on well-established paradigms on optimistic updating (e.g. Sharot et al., 2011). In this type of experiment, participants are typically presented with an event that is either positive or negative in valence and prompted to estimate the likelihood that the event occurs. For instance, participants might be asked the likelihood from 0 to 100% whether they will "attend a party in the next month." Depending on whether the participant views the event as a positive (as an extrovert might see a party) or negative (as an introvert might see a party), the event is categorized as positive or negative. After providing an estimate, participants are provided a base rate, such as the percentage of people who attend a party in a month. Depending on whether this base rate is optimistic or pessimistic for a given person, the information is then categorized as "good news" or "bad news." For instance, a person who overestimates the likelihood of going to a party might find the base rate discouraging if they are an extrovert (bad news) but encouraging if they are an introvert (good news). After participants see the base rate, they then provide an updated

likelihood percentage that may or may not deviate from the initial estimate. The extent to which they update their belief is of main interest.

An asymmetry between good and bad news has been well-established in the psychological literature (Barron, 2020; Garrett et al., 2014, 2018; Garrett & Sharot, 2017; Kappes & Sharot, 2019; Sharot et al., 2011). When faced with bad news, people are tempted to not budge in the face of evidence and provide a second estimate that is similar to their first estimate. In this situation, participants should report high levels of influence. As they update more in the face of bad news, they should report lower levels of influence, reflecting a willingness to conform to evidence despite it being undesirable. The opposite tends to be true for good news: when updating little, there should be low reports of influence as most seize on more optimistic evidence, and when updating a lot, there should be higher perceptions of influence.

The belief updating paradigm provides a useful task for observing both motivated reasoning and awareness perceptions. When individuals evaluate information, it is often done incrementally and may be difficult to detect without obvious qualitive change, such as switching from accepting evidence to rejecting it (Bechler et al., 2019, 2020; Clark & Wegener, 2013). In an ideal study, we would measure the amount of cognitive processing to see if a greater quantity of processing predicted more awareness. While we cannot quantify the cognitive processing going on, belief updating allowed for differential measurement of belief change with precision.

**Method**

*Participants*

Our sample consisted of undergraduate participants who took part in our study in exchange for course credit. After eliminating 7 people who responded negatively to data quality questions[1], we were left with 159 participants ($M_{age}$ = 18.9, 52.9% female).

*Procedure*

Figure 1 depicts the presentation of each event in 2 blocks of questions for 67 life events for optimistic updating task. In the first block, participants provide a probability estimate estimating the likelihood of the event occurring. These events were generally positive ("Have a supervisor/teacher praise you in the next month"), negative ("Die from heart disease"), or neutral ("Get a haircut in the next month"). After providing an estimate on this item for themselves, our college participants were presented a base rate, which was determined from a separate sample of college-aged individuals, detailed below.

Prior to the study, we created base rates determined from a sample of college-aged Amazon MTurk participants, or from public sources. For example, 51% of our college-aged stated that they were praised by their supervisor or teacher in the next month. Participants then provide a new estimate that may or may not differ from their first. For each trial, after these estimates were provided, participants were asked "Did whether or not you found the average

---

[1] The data quality question for all studies was a "yes or no" question phrased as follows: "As researchers, the quality of our data is very important to us, so we want to make sure that your responses are valid and authentic. In your honest opinion, should we use your data?" Any response marked "no" was removed from further analysis.

16

rates optimistic vs. pessimistic influence your estimates?" on a scale from 1 (Did not influence me) to 5 (Did influence me). Appendix A shows all items and base rates used in Study 1.

**Figure 1**

*Stimulus Presentation for Study 1 Trials*
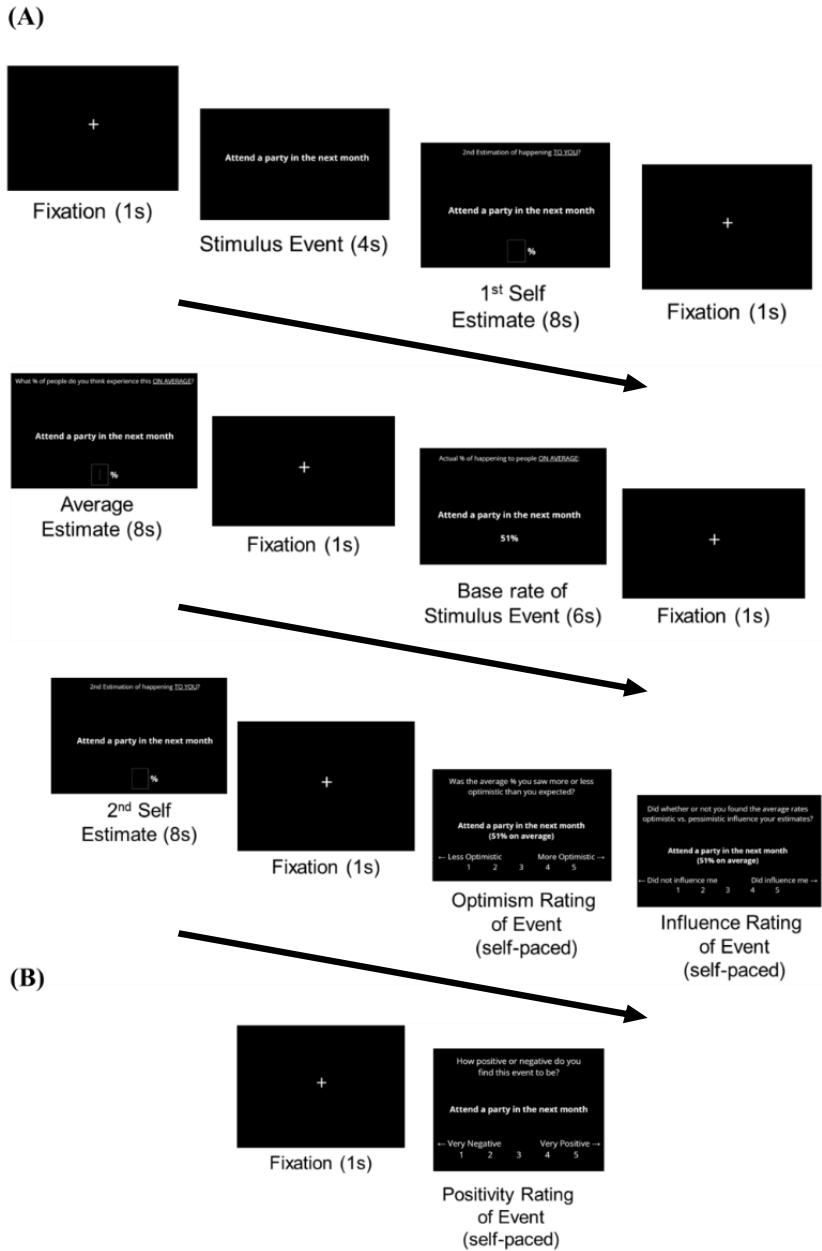
*Note.* Panel A depicts a typical trial for Study 1's first block. Participants provided an estimate for each event and were presented with a base rate of a similar sample. They then provided an updated estimate and their self-perception of how much the desirability of an event influenced them. Panel B shows the 2nd block for each event where we estimate whether participants viewed a given event as positive or negative so that we can classify events as good news or bad news.

In a second block, participants rated each event from 1 (Very Negative) to 5 (Very Positive). Following previous research, for each participant, events were considered desirable if they were rated above the midpoint and undesirable if they were rated below the midpoint.

**Good News vs. Bad News.** All trials were categorized based on each participant's initial estimates in relation to the base rate (see Table 1): "good news" is when a desirable event was estimated to be *less likely* for the self than the base rate or an undesirable event was estimated to be *more likely* than the base rate; "bad news" is when a desirable event was estimated to be *more likely* for the self than the base rate or an undesirable event was estimated to be *less likely* than the base rate. We measured the extent to which people update in response to good vs. bad news, and whether there were any patterns of awareness of influence.

A pattern of motivated reasoning would predict that in the case of bad news, people will reject new information and update less, but in the case of good news, they will willingly accept new information and update more. A pattern of awareness would predict a positive relation between awareness of influence and updating for good news (as one updates towards a desired outcome, they report more self-awareness) and a negative relation between awareness of influence and updating for bad news (as one is resistant and updates less to undesired information, they should report more self-awareness).

**Table 1**

*Classification of Good vs. Bad News*

|  | Desired or Positive Event | Undesired or Negative Event |
| --- | :---: | :---: |
| % higher than expected? | Good News | Bad News |
| % lower than expected? | Bad News | Good News |

*Note.* Each participant provides a percentage estimate of the likelihood an event occurs, then is provided a base rate. Percentage higher or lower than expected refers to the base rate of an event compared to a 1st estimate for the self. For instance, if one estimated a 20% chance of attending a party and was presented with a base rate of 51%, this would be considered good news if a party was desired, or bad news if a party was undesired.

One could argue that any large amount of updating towards the base rate should be consistent with low influence, as conforming to a base rate is probably a more rational choice than any other method of estimating when things are going to happen to the self. At the very least, one should observe higher levels of awareness for bad news compared to good news at low levels of updating and reversed or similar levels of awareness at higher levels of updating.

Update ratings were calculated as the absolute difference between the 1st and 2nd estimate and signed so that positive values are estimates moving towards the base rate and negative values are moving away from the base rate (Sharot & Garrett, 2022). We then calculated the average update amount and awareness ratings separately for both good news and bad news for each participant. This allowed us to average out extreme cases where one may have given an initial estimate that was far from the baseline. If analyzing data on the item level, these extreme cases might misrepresent movement on beliefs a result of a different starting points. Sharot and Garrett (2022) recommend adjusting for the difference between an initial estimate and the correct base rate as a control variable when modeling this, which yields the same pattern of results as what we present below. For simplicity, we present the participant-averaged update scores.

19

**Results**

Consistent with past research, participants updated more in response to good news ($M = 6.789$, $SD = 4.045$) when compared to bad news ($M = 3.91$, $SD = 3.116$), paired $t(157) = 9.396$, $p < .001$, $d = .747$. The grand mean of awareness on all items was 2.574 on our 5-point scale. However, the modal response, occurring on 37.7% of all items, was 1 (did not influence me). On average, participants reported estimates for more bad news trials ($M = 21.475$, $SD = 6.797$) when compared to good news trials ($M = 20.589$, $SD = 6.216$).

Next, we regressed average awareness on average update amount, $F(1, 145) = 59.168$, $p < .001$, $R^2_{adjusted} = .285$. The relation between update amount and awareness (regardless of news) revealed a positive relation, $\beta = .122$, $SE = .016$, $p < .001$, 95% CI [.09, .153]. Using a mixed-effects model with participant as a random intercept (to account for the within-subjects news type predictor), we further investigated the interaction between these predictors, finding no moderating effect of news, $\beta = -.013$, $SE = .011$, $p = .253$, 95% CI [-.035, .009]. These results are depicted in Figure 2.

Finally, I probed the data to address some possible limits of the present results. I first looked at extremity influencing a null interaction, insofar as awareness is related to attitude accessibility (e.g. Giner-Sorolla, 2001). Using what Sharot & Garrett (2022) call the "estimation error" in belief updating, I take the absolute value of the difference between the first estimate and the base rate, and average this for each participant. When including this as a variable in a update amount x news x estimation error interaction predicting awareness, there was no significant moderating effect, $\beta = .002$, $SE = .003$, $p = .485$, 95% CI [-.003, .007].

To address the possibility that practice effects influenced update and awareness reports, I also examined the update x news interaction predicting awareness by only examining the first 10 trials from every participant. Using a similar mixed-effects model adjusting for participant as a random intercept, this interaction was not significant, suggesting that participants did not respond differently over the course of the experiment, $\beta = -.002$, $SE = .011$, $p = .869$, 95% CI [-.024, .02].

**Figure 2**

*Study 1 Moderating Effect of News Type on the Update-Awareness Relation*
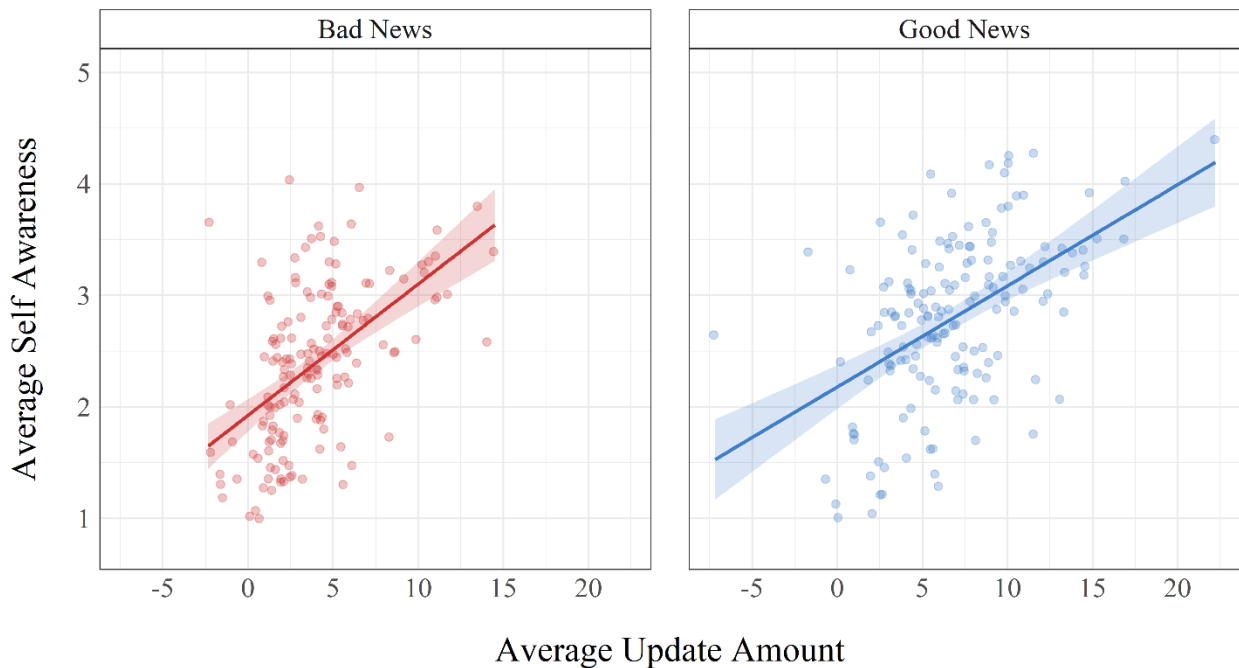


*Note.* Each point represents the average update amounts and influence ratings for each participant. A participant has one average for bad news and one average for good news. Positive values translate to average movement towards the base rate, negative values translate to average movement away from the base rate.

**Study 1 Discussion**

While we see a classic account of motivated updating, there is a positive relation between updating and awareness for both news types. This indicates that people report more influence of desirability as they update more, regardless of if the information is desired or not. The positive effect for good news might be interpreted as consistent with a pattern of awareness: people are more willing to accept information that validates the likelihood of a desirable outcome and participants are able to recognize this by reporting more influence. For bad news, we should see a negative relation (or no relation), suggesting that people know when they do not conform to negative information. These results are not consistent with a pattern of awareness. Rather, any awareness of influence could be that people are simply inferring awareness just because they are updating a lot.

**Study 2**

In Study 1, participants reported higher awareness of influence in their desires as they updated their beliefs more, regardless of whether it was good or bad news. This was not consistent with a pattern of motivated reasoning, where awareness should be greater when people update in response to good news but lesser when updating in response to bad news.

Study 2 used a similar optimistic updating task as Study 1. To extend generalizability beyond a student population, we examined an online sample of participants making judgments about the 2022 US midterm elections. Prior beliefs and attitudes interfere with how people discount political information (Taber et al., 2009; Taber & Lodge, 2006). We therefore expect that participants will be more willing to accept politically convenient information when updating beliefs, but not politically inconvenient information. If the results follow from Study 1, good/bad news should not moderate the relation between updating and awareness ratings.

22

**Method**

*Participants*

Participants were recruited from Prolific the week leading up to the 2022 US midterm elections. We initially recruited 200 self-identified Republicans and 200 self-identified Democrats and eliminated 7 participants who responded negative to our data quality check and 15 who identified as politically independent, leaving a final sample of 378 participants ($M_{age}$ = 42.7, 48% female).

*Procedure*

Our participants evaluated 14 different election races featuring a competitive senate or governor seat. Among the 14 elections, 7 favored Democrat candidates and 7 favored Republican candidates. Like Study 1, participants made an estimate of a candidate winning, were presented with a base rate, then made an updated estimate. For instance, in the Michigan governor race, they first saw a page asking "What is the chance you think that Gretchen Whitmer (Democrat) or Tudor Dixon (Republican) <u>will win</u> the Michigan Governor race?" After providing an estimate (that adds up to 100%), they were then provided our base rate, which was gathered from fivethirtyeight.com, a website that provides a model estimate of a candidate winning based on an aggregate of polls and other data. The base rate was a percentage forecast based on fivethirtyeight's model ranging from 0 – 100% (e.g. "538 predicts that Gretchen Whitmer (88% chance) is more likely to defeat Tudor Dixon (12% chance). See Appendix B for base rates from all 14 elections. They then had an opportunity to provide a new estimate that either differed or stayed the same as their first estimate. After providing estimates for each race, participants were asked "Do you think your political preferences might have influenced how much or how little

you revised your estimate in light of the 538.com information?" and rated this on a scale of 1 (did not influence at all) to 5 (greatly influenced my responses).

Similar to Study 1, the results did not differ when fitting a mixed-effects model to account for initial estimates affecting the amount of room for updating. For simplicity, we report models that examine the participant averages for updating and awareness ratings.

**Good News vs. Bad News.** Each of the 14 elections considered were again categorized into good or bad news. Rather than rating each candidate as preferred or not, we inferred good vs. bad news based on the participant's political party. For example, a Republican participant who initially overestimated the likelihood of Tudor Dixon (an "inpartisan" Republican) defeating Gretchen Whitmer (an "outpartisan" Democrat) relative to the 538.com base rate would be considered a case of "bad news." Patterns of motivated reasoning and awareness would look similar to the case of Study 1: people should update more in response to good news when compared to bad news, and aware participants should show a negative relation between updating and awareness for bad news and a positive relation between updating and awareness for good news.

**Results**

We again aggregate good and bad news within an individual, such that each participant has an average updating amount and awareness rating for good news and bad news. Based on our categorization scheme, participants were presented with more bad news trials ($M = 7.468$, $SD = 1.496$) when compared to good news trials ($M = 5.946$, $SD = 1.586$). As with Study 1 and past research, participants updated more in response to good news ($M = 11.525$, $SD = 9.009$) than bad news ($M = 9.887$, $SD = 7.763$), paired $t(370) = 3.605$, $p < .001$, $d = .187$. The grand mean of

awareness on all items was also nearly identical to Study 1: 2.586 on our 5-point scale. The modal response was 1 (did not influence at all), occurring on 33.5% of all responses.

As with Study 1, we fitted a linear model predicting average participant awareness as a function of updating amount, $F(1, 372) = 2.597$, p = .108, $R^2_{adjusted} = .004$. Unlike Study 1, there was no relation between updating and awareness, $\beta = -.013$, $SE = .008$, $p = .108$, 95% CI [-.029, .003]. Looking at a mixed-effects model adjusting for participant as a random intercept, there was also no moderating effect of news on this relation, $\beta < .001$, $SE = .005$, $p = .971$, 95% CI [-.009, .009]. These results are shown in Figure 3.

**Figure 3**

*Study 2 Moderating Effect of News Type on the Update-Awareness Relation*
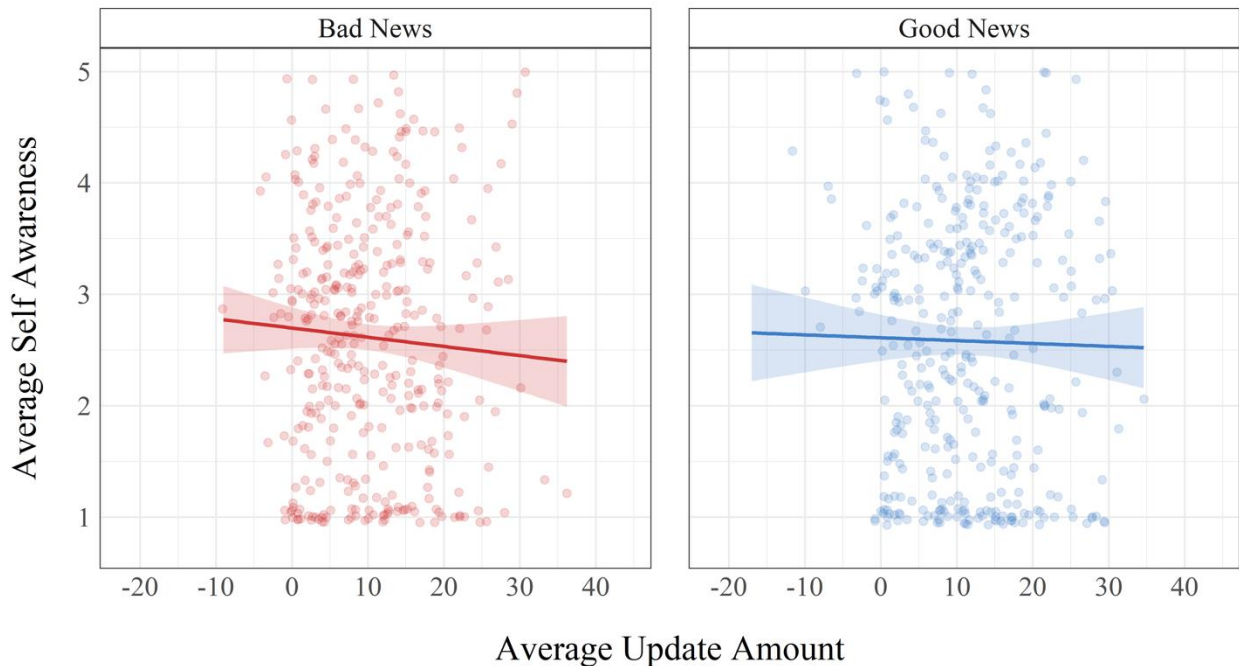


*Note.* As with Study 1, each point represents the average update amounts and influence ratings for each participant: one for bad news and one for good news. Positive values translate to average movement towards the base rate, negative values translate to average movement away from the base rate.

**Study 2 Discussion**

In Study 2, we presented a belief updating task to Republican and Democrat participants in an online sample. Participants updated their beliefs when presented base rate information about an upcoming election. The suggest that in aggregate, people are biased towards accepting information consistent with desired outcomes, but they generally do not report differential awareness of the influence of their political preferences regardless of whether it is good or bad news. Unlike Study 1, there was generally a static rating of awareness of political influence near the midpoint of 3 out of 5 ($M = 2.57$ for all awareness reports). Similar to Study 1's results, this suggests that there is no pattern of awareness for results consistent with motivated reasoning.

The results of this study and Study 1 suggest that motivated reasoning about political and everyday life events do not reveal a pattern of awareness. Does one's own self-perceptions cohere with what others think of us? While people generally believe themselves to be rational about adapting their beliefs to evidence, this does not always reflect one's perceptions of another person (Ehrlinger et al., 2005; Pyszczynski & Greenberg, 1987). If people understand low reports of awareness as normative, they should rate themselves a similar amount as another person.

**Study 3**

Previous studies established that participants generally do not show awareness when completing a motivated reasoning task related to belief updating. Study 3 extends this to peer judgments of political bias using data from Study 2. People can generally recognize that motivational influences affect the attitudes of others but fail when examining themselves (Cohen, 2003; Epley & Dunning, 2000; Pronin et al., 2002). Additionally, when people's expectations of a perceived actor are violated, perceivers will find arguments more effective and persuasive

26

(Eagly et al., 1978; Walster et al., 1966). Study 3 therefore examines whether a participant's own awareness of bias is consistent with a third party judge.

In addition to serving as an investigation of normative perceptions of the nature of awareness, this study acts as a check on our awareness questions. It is important that participants comprehend the goal addressed by a give question, and that a question is not influenced by extraneous context effects (Schwarz, 1999; Wilson & Hodges, 1992). This study allows us to check that awareness questions are understandable and straightforward when applied to another person.

**Method**

*Participants*

In this study, we take data from Study 2 and use them as stimuli for peer judgments. Study 2 data were filtered to contain complete data on all 14 election races from Democrat or Republicans and contained no updating data outside 3 standard deviations. This left us with data from 281 participants. We recruited an equal number of participants (N = 281, $M_{age}$ = 41.02, 51.6% female) from Prolific and yoked these participants to the sample from Study 2, such that each new participant evaluated the responses from a single Study 2 participant (referred to as "Participant A").

*Procedure*

We informed the new participants of Participant A's political preferences, their first estimate, the 538.com estimate, and their revised estimate. We then asked them "Do you think Participant A's political preferences might have influence how much or how little they revised their estimate in light of the 538.com information? Political preferences may have influenced

Participant A regardless of if they shifted a lot or very little." Participants rated this on a scale from 1 (Did not influence at all) to 5 (Greatly influenced their responses). We averaged all trials for an index of perceived bias. Average update amount for good and bad news was consistent with Study 2.

**Results**

When compared to original participants ($M = 2.557$, $SD = 1.106$), current participants ($M = 2.834$, $SD = .736$) reported more perceived bias than self-reported bias, Welch's t-test $t(487.39) = 3.497$, $p < .001$, $d = .295$.

To test perceived bias as predicted by a good/bad news x update amount interaction, we fitted a mixed effects model with participant as a random intercept. Unlike the Study 2 data, participants in Study 3 rated these peers in a pattern consistent with a pattern of awareness of political bias: there was a negative relation between updating amount and perceived bias for bad news, and a positive relation for good news, $\beta = .043$, $SE = .008$, p < .001, 95% CI [.027, .059]. This interaction is depicted in Figure 4. This suggests that people are sensitive when others do not accept information for bad news and perceive more bias, but perceive more bias when people update a lot for good news.

We also checked to see if inpartisans (e.g. Republicans evaluating Republicans) or outpartisans (e.g. Republicans evaluating Democrats) rated perceived bias differently. Participants were aware if the peer they were yoked to shared the same party, but there were no main or moderating effects (3-way interaction with news type x update amount) on perceived bias, $\beta = .008$, $SE = .017$, $p = .64$, 95% CI [-.026, .042].

**Figure 4**

*Study 3 Moderating Effect of News Type on the Update-Perceived Bias Relation*
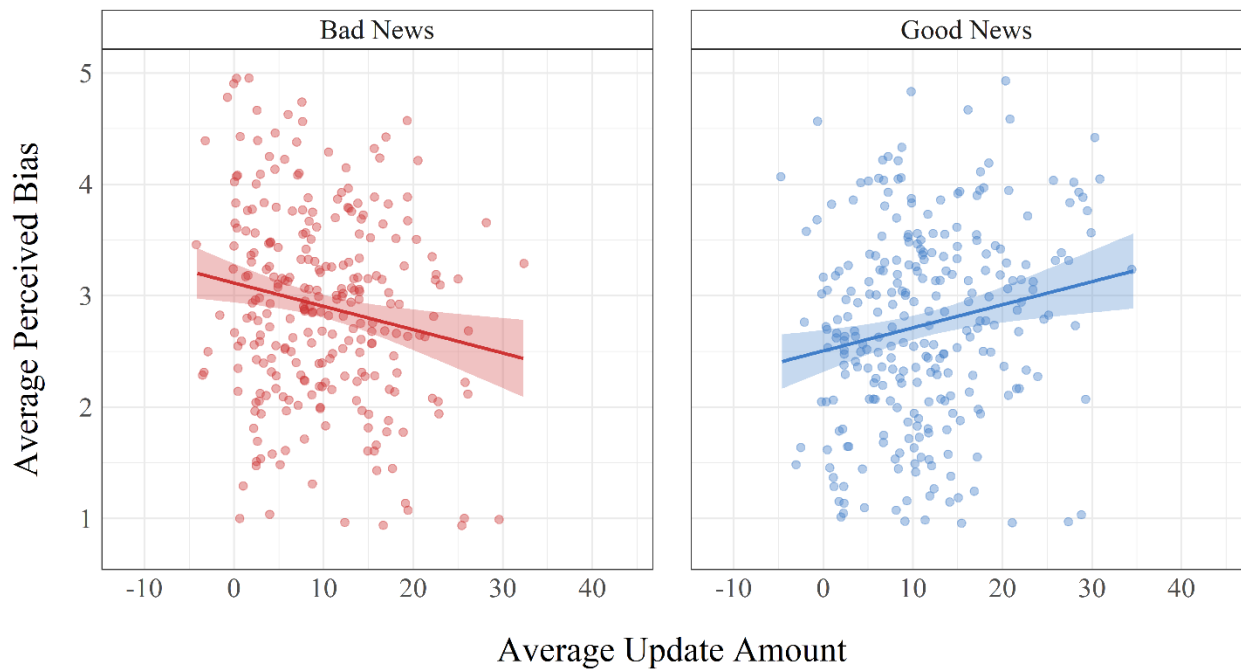


*Note.* As with Studies 1 and 2, each point represents the average update amounts and influence ratings for each participant being evaluated: one for bad news and one for good news. The y-axis for Study 3 is a third-party peer judge of participants in Study 2 based on their 1st and 2nd updates (reflected on the x-axis).

**Study 3 Discussion**

It is difficult to describe a "ground truth" for the quantity of bias in motivated reasoning, but perceptions of others from a similar sample is a good approximation. Indeed, peers can rate one on personality traits with similar or even superior accuracy to oneself, depending on the trait (Funder, 2012; Nater & Zell, 2015; Vazire, 2010; Vazire & Mehl, 2008). In Study 3, we presented belief updating responses from participants in a previous study sample to new participants of the same population. The results demonstrate that there is a noticeable difference between how individuals evaluate motivated reasoning in others vs. ourselves. Participants rated

perceived levels of bias as greater when compared to an original participant's perceived bias. Additionally, data revealed a good/bad news x update amount interaction predicting perceived bias, such that an expected pattern of motivated reasoning was apparent. Overall, these results reduce the likelihood of the explanation that differences in awareness ratings are normative and indicative of a shared lay theory about awareness.

# CHAPTER 3

## Simulated Ballot Counting Task

### Study 4

After Studies 1 through 3, we identify two aims to extend our research program: account for the possibility that participants conflate the *quantity* of belief change with awareness and examine awareness in motivated reasoning using a different experimental paradigm. Study 4 seeks to address both aims using another motivated reasoning paradigm introduced by Kopko et al. (2011). In this study, participants complete a simulated ballot counting task where they evaluate several hand-written ballots that may or may not contain stray marks that would invalidate them. Unlike belief change, participants make a binary judgment about each ballot (keep or discard). This allows us to ensure that they are not using the quantity of action or inaction as a heuristic for perceiving their own political influence.

We varied the extent to which our stimuli varies by ambiguity. By including unambiguous ballots, we address the possibility that participants are merely expressive responding (e.g. Schaffner & Luks, 2018). That is, people may be aware when they are providing motivated responses but simply not admitting it. This may be inferred with completely unambiguous ballots that, given no other information, a dispassionate observer would find objectionable. If a participant discards such a ballot when congenial, but reports no awareness of this, we can infer that expressive responding may be taking place. These unambiguous trials also included ballots that should be obviously discarded. For instance, some ballots had multiple candidates' bubbles clearly filled in, yet were still counted for either Biden or Trump, evenly

distributed. Since no voter intention could possibly be inferred, we expected that these ballots be unanimously discarded.

## Method

### *Participants*

We initially recruited 200 Democrats and 200 Republicans from Prolific. We excluded a total of 16 participants: 1 participant answered negatively to our data quality question, 4 participants reported technical errors, and 11 participants reported a political affiliation that was not Democrat or Republican. Our final sample was 384 participants ($M_{age}$ = 45.5, 50.4% female).

### *Procedure*

Participants were first given some brief and vague instructions about the key factors in what to consider when deciding if a ballot should be counted (see Appendix C; adapted from Kopko et al., 2011). They were to evaluate 40 total ballots, divided evenly in favor or Donald Trump or Joe Biden. The ballots varied in ambiguity, or the extent to which a ballot contains stray or abnormal markings, as shown in Figure 4. Half of the ballots were ambiguous—designed a priori to be plausibly acceptable to a dispassionate observer—and half were unambiguous—designed to be invalid by any reasonable participant. Among unambiguous ballots, we included 10 trials of unambiguous ballots that should be kept, and 10 trials of ballots unambiguous ballots that should be discarded. These were designed a priori and generally displayed either no egregious markings (see Figure 4, left), or completely egregious markings (for example, a large "X" across the entire ballot page). See Appendix D for all ballots used.

For each ballot, participants were shown the ballot and told to which candidate the ballot was counted in favor. Participants responded to a binary decision of "Don't Keep Ballot" by

pressing the D key or "Keep Ballot" by pressing the "K" key. They also responded to our measure of awareness "Did whether or not you favored certain candidates influence your response?" using the number keys for a 1 (Did not influence me) – 5 (Did influence me) Likert scale. The candidate order on the ballot was counterbalanced.

Participants also reported information about their political party preference and items related to election integrity. These items were adapted from a recent Pew Research survey (Pew Research Center, 2022) and included questions like "The next elections in the United States will be run and administered well" and "How much do you trust the officials who oversee elections where you live?" on a 1– 5 Likert scale. An exploratory factor analysis revealed that all items loaded on the same factor ($\alpha = .91$).

**Figure 5**

*Study 4 Ballot Counting Stimuli*
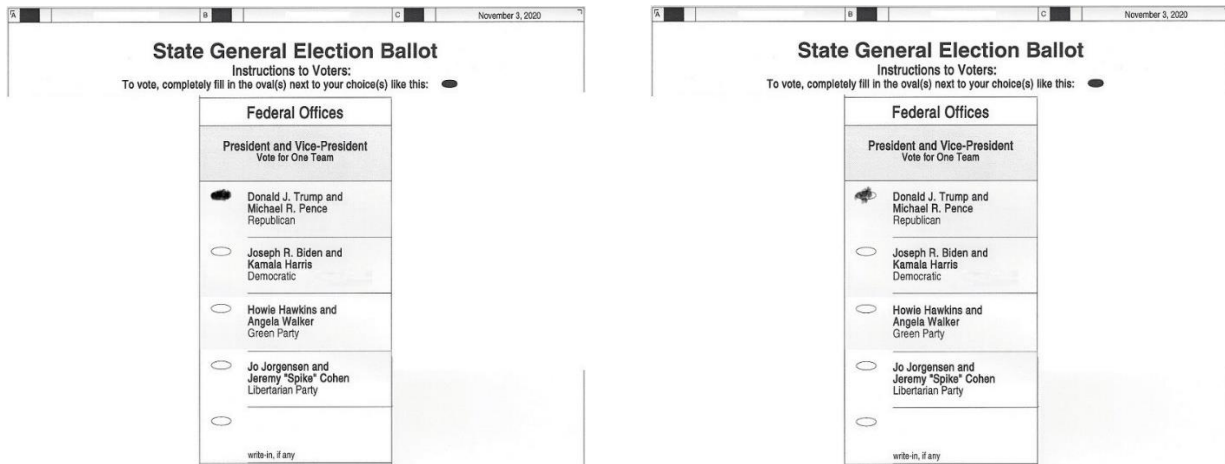


*Note.* The ballots used in Study 4 feature different levels of ambiguity. Both these ballots favor Donald Trump if kept, but the right ballot features a smudge that would license one from throwing it out if they oppose Trump, or keeping it if they favor Trump.

**Motivated Evaluation.** What might a pattern of motivated reasoning look like with ballot responses? This would be consistent with a congeniality bias towards discarding or keeping ballots. For instance, this might look like a Republican discarding a ballot that favored Biden but not Trump, and a Democrat discarding a ballot that favored Trump but not Biden. A pattern of awareness would be similar: when compared to choices that are uncongenial to one's political beliefs, there should be higher levels of awareness when making choices congenial with one's political preferences. We determine this by creating a difference score of outpartisan ballots minus inpartisan ballots. Thus, a positive value suggests that a participant was more biased in favor of their preferred candidate. A negative value suggests that they were biased in favor of their nonpreferred candidate. A value of 0 suggests that they did not preferentially evaluate one candidate more often than the other. If a pattern of awareness is present, we should see a linear relation between this discard index and reported influence. If these variables are not related, we can infer that awareness of political influence is not present.

**Results**

*Motivated Evaluation*

Among ambiguous ballots, there was a moderate effect of biased evaluation, with outpartisan ballots ($M = 3.398$, $SD = 3.135$) being discarded more often the inpartisan ballots ($M = 2.293$, $SD = 2.626$), paired t-test $t(381) = 8.114$, $p < .001$, $d = .382$.

Among unambiguous ballots, we examined partisan differences among "unambiguous discard" and "unambiguous keep" trials. Similar to ambiguous ballots, there was a moderate effect for unambiguous discard ballots, with outpartisan ballots ($M = 4.781$, $SD = .696$) also being discarded more than inpartisan ballots ($M = 4.549$, $SD = 1.123$), paired t-test $t(383) = 4.846$, $p < .001$, $d = .247$. Among unambiguous keep trials, outpartisan ballots ($M = .492$, $SD = $

1.27) were still discarded more than inpartisan ballots ($M = .172$, $SD = .675$), paired t-test $t(383)$ = 5.05, $p < .001$, $d = .258$.

After collapsing inpartisan and outpartisan discarding into our biased discard index difference score, we looked at differences in biased discarding by ambiguity. As expected, ambiguous ballots ($M = 1.094$, $SD = 2.659$) elicited more biased discarding when compared to unambiguous ballots ($M = .552$, $SD = 1.988$), paired t-test $t(383) = 5.68$, $p < .001$, $d = .29$.

*Awareness*

The modal response for perceived influence was once again a 1 out of 5, occurring in 82.8% of all trials. We regressed awareness as a function of the biased discard index, $F(1, 382)$ = 121.497, p $< .001$, $R^2_{adjusted} = .239$. Across all trials, we found that an increase in ballot discard bias predicted greater awareness, $\beta = .098$, $SE = .009$, $p < .001$, 95% CI [.081, .116]. When looking at awareness differences between ambiguity, ambiguous ballots ($M = 1.46$, $SD = .881$) yielded more awareness when compared to unambiguous ballots ($M = 1.368$, $SD = .858$), paired t-test $t(383) = 6.948$, $p < .001$, $d = .355$.

To look at the discard-awareness relation as moderated by ambiguity, we used a mixed effects model regressing average awareness on ambiguity, with subject as a random intercept. More ambiguous trials prompted more reporting of awareness, $\beta = .071$, $SE = .014$, $p < .001$, 95% CI [.043, .098]. After examining an interaction term, ambiguity did not moderate the relation between discard bias and awareness, $\beta = -.008$, $SE = .007$, $p = .236$, 95% CI [-.021, .005]. Specifically, the relation between discarding and awareness was slightly stronger among unambiguous trials (simple slopes $\beta = .054$, p $< .001$) when compared to ambiguous trials ($\beta = .046$, p $< .001$), but the slopes did not significantly differ from each other. This relation is depicted in Figure 5.

35

**Figure 6**

*Study 4 Relation of Self-Perceived Bias by Biased Evaluation x Ambiguity*
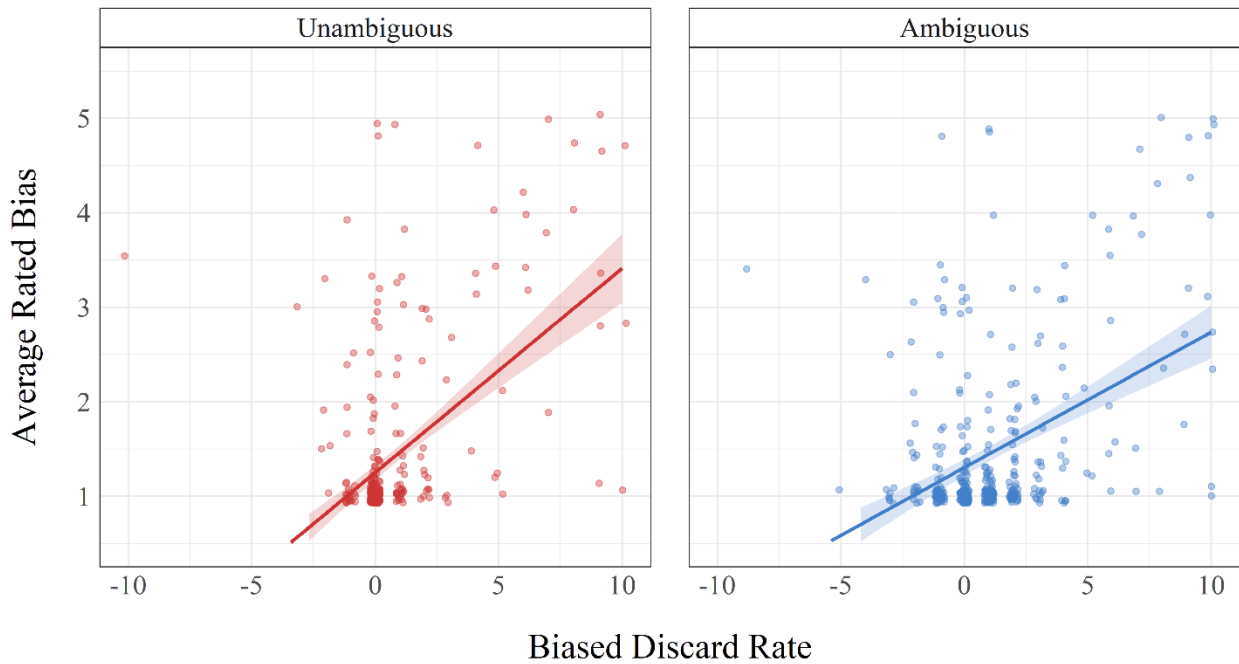


*Note.* Ambiguous trials elicited more awareness when compared to unambiguous trials. This relation did not vary as a function of biased discarding, as the rates of awareness as a function of biased discarding were similarly positive for both ambiguous and unambiguous trials. Participants who were not biased in discarding generally showed low awareness regardless of ambiguity.

### *Role of Political Identity*

Election integrity has become an increasingly partisan topic, with much disagreement over confidence in polling accuracy (Pew Research Center, 2022). As such, we were interested in an exploratory analysis of partisanship as a predictor and moderator in the present effects. Republicans ($M = 2.606$, $SD = 5.388$) were more likely to engage in biased discarding when compared to Democrats ($M = .675$, $SD = 2.486$), Welch's t-test $t(270.8) = 4.516$, $p < .001$, $d = .459$. Using mixed-effects models to test moderation, political affiliation did not moderate the relation between ambiguity and discard bias, $\beta = .171$, $SE = .191$, $p = .369$, 95% CI [-.202, .545].

Specifically, the effect of increased ambiguity on discard bias was similarly positive for

Republicans (simple slopes $\beta$ = .627, p < .001) when compared to Democrats (simple slopes $\beta$ =

.455, p = .001).

Additionally, Republicans reported more overall awareness ($M$ = 1.57, $SD$ = 1.218) when

compared with Democrats ($M$ = 1.257, $SD$ = .753), Welch's t-test $t(304.54)$ = 3.632, $p < .001$, $d$

= .37. Finally, we examined politics as a moderator of discard bias effects on awareness in a

mixed-effects model adjusting for participant intercept. For Republicans (simple slopes $\beta$ = .063,

p < .001), there were similar awareness reported as the discard bias index increased when

compared to Democrats (simple slopes $\beta$ = .045, p < .001), $\beta$ = .018, $SE$ = .013, $p$ = .172, 95% CI

[-.008, .044]. This relation between discard bias and awareness is depicted in Figure 6.

**Figure 7**

*Study 4 Relation Between Self-Perceived Bias Biased Evaluation by Politics*
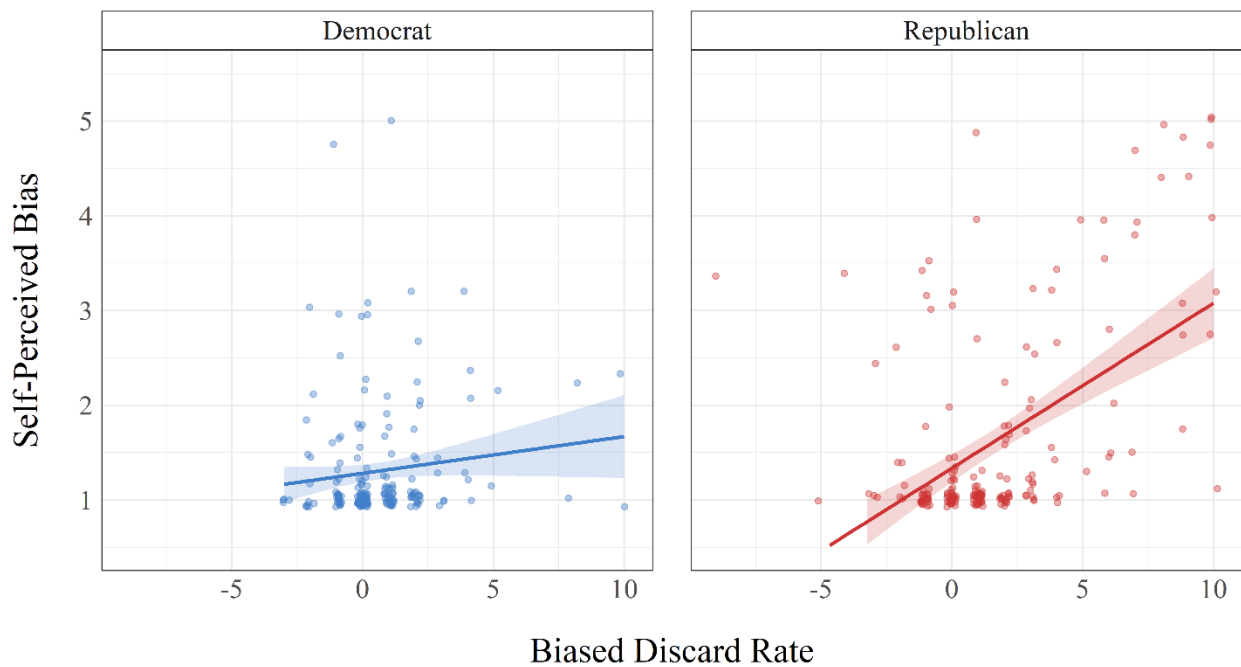
*Note.* Republican participants demonstrate more bias on the ballot counting task relative to Democrats. The plot shows similar effects of political party on the bias-awareness relation: both Republicans and Democrats show positive slopes for the effect.

Given that there were so many trials in which participants did not report more than the minimal amount of awareness, an exploratory analysis investigated political bias differences between Democrat and Republicans based only on these minimal awareness trials. For each participant, we calculated a percentage of ballots discarded out of inpartisan and outpartisan ballots evaluated where participants reported a 1 out of 5 on political influence. Using a mixed-effect model adjusting for participant as a random intercept, we tested whether politics moderated the relation between inpartisan and outpartisan ballot discarding. The discrepancy between how much Republicans discarded more outpartisan ballots ($M = .452$, $SD = .226$) when compared to inpartisan ballots ($M = .408$, $SD = .227$) was greater than Democrats discarding outpartisan ($M = .393$, $SD = .184$) or inpartisan ballots ($M = .389$, $SD = .181$), but this interaction only trended towards significance, $\beta = .037$, $SE = .02$, $p = .065$, 95% CI [-.002, .076].

### Election Integrity Beliefs

An exploratory analysis revealed a strong effect of party on election integrity beliefs. Republicans reported much lower endorsement of these beliefs ($M = 3.259$, $SD = .973$) when compared to Democrats ($M = 4.102$, $SD = .733$), Welch's t-test $t(356.73) = 9.6$, $p < .001$, $d = .979$. Regressing ballot discard index on election integrity beliefs, $F(1, 382) = 7.185$, p = .008, $R^2_{adjusted} = .016$, revealed a significant effect of beliefs on discarding, such that an increase in endorsement of election integrity beliefs resulted in lower discard bias, $\beta = -.611$, $SE = .228$, $p = .008$, 95% CI [-1.059, -.163]. A separate simple regression model also revealed a small effect of election integrity beliefs on overall awareness, $F(1, 382) = 4.978$, p = .026, $R^2_{adjusted} = .01$.

Specifically, an increase in election integrity led to a small reduction in awareness, $\beta$ = -.102, *SE* = .046, *p* = .026, 95% CI [-.192, -.012].

**Study 4 Discussion**

In Study 4, we presented participants with ambiguous and unambiguous ballots that favored either a preferred political candidate or a nonpreferred political candidate. Participants engaged in biased evaluation of the ballots. Consistent with Kopko et al. (2011), people were more biased when evaluating ballots in favor of their preferred political party. Although biased discarding occurred among both ambiguous and unambiguous ballots, the effect seemed to be stronger among ambiguous ballots.

When looking at how participants reported awareness of their pollical beliefs influencing their evaluations, there was a tendency to report low influence across the board, with 82.8% of all trials yielding a minimum awareness rating. More ambiguous ballots prompted more awareness overall. This relation was not moderated by discard bias, as both ambiguous and unambiguous trials shared a similarly positive relation between discard bias and awareness. Even though there were many minimum ratings among bias participants evaluating ambiguous trials, those who were highly biased did not underreport awareness. This suggests that ambiguity did not sufficiently license people to report low awareness relative to unambiguous ballots, even if it did license additional discard bias.

Additionally, we investigated political identity and election integrity beliefs as exploratory analyses. When compared to Democrats, Republicans were more likely to engage in biased evaluation of ballots, but also were more aware of their behavior. Republicans were also much more likely to report lower beliefs in election integrity when compared to Democrats. Political identity did not seem to moderate the relation between discard bias and awareness, as

both Democrats and Republicans reported more awareness as they became more biased. This might be qualified by a low variation of bias among Democrats, as seen in Figure 6. The ballot task is therefore more likely to elicit biased responding among Republicans when compared to Democrats, but this was also reflected in differences in self-awareness.

## Study 5

Study 5 adapts participant responses from Study 4, testing whether peers report similar levels of perceived influence as participants in Study 4 when evaluating political ballots. Perceptions of political influence can be compared against self and other ratings, which allow us to determine the extent to which there are normative standards that we imbue in others' biases vs. our own. This study follows the pattern of Study 3, where participants reviewed peer responses to elections as peers updated their beliefs, and rated the peer on perceived political bias. The present study follows a similar procedure, but with the simulated ballot task from Study 4.

**Method**

*Participants*

Participant data from Study 4 were used as stimuli for this study. Participants were yoked to new participants from Prolific. We recruited participants until we had a matching final sample of 384 participants ($M_{age}$ = 43.3, 52% female).

*Procedure*

Once again, we referred to original participants as "Participant A" and for each ballot, we presented Participant A's political beliefs and preferred party, the ballot image, and Participant A's decision. Perceived bias ratings and Participant A self-perceived awareness ratings were

40

averaged, and Participant A discard bias index was again calculated as the difference between outpartisan ballots discarded minus inpartisan ballots discarded.

**Results**

Participants rated Participant A's bias across all trials as greater ($M = 2.004$, $SD = 1.415$) than original participant data ($M = 1.414$, $SD = 1.026$), paired $t(381) = 10.789$, $p < .001$, $d = .552$. These two ratings were moderately correlated across all trials, $r(380) = .256$, $p < .001$. Using a simple regression model, we looked at whether Participant A's biased discard index predicted overall perceived bias, $F(1, 380) = 12.308$, $p < .001$, $R^2_{adjusted} = .051$. As Participant A's discard bias increased, peers accurately perceived more bias, $\beta = 1.459$, $SE = .316$, $p = .001$, 95% CI [.838, 2.08].

*Ambiguity*

All ambiguity models considered only ambiguous and unambiguous discard trials, as reported in Study 4. Ambiguity was significantly predictive of perceived bias, such that ambiguous responses ($M = 2.173$, $SD = 1.446$) yielded more perceived bias when compared to unambiguous responses ($M = 1.835$, $SD = 1.363$), paired $t(379) = 12.992$, $p < .001$ $d = .666$.

Using a mixed-effects model adjusting for participant intercepts, we looked at ballot ambiguity as a moderating variable in the relation between discard evaluation bias and perceived bias, which was nonsignificant, $\beta = .024$, $SE = .013$, $p = .059$, 95% CI [-.001, .05]. Specifically, the relation between discard index and perceived bias was similar between unambiguous ballots (simple slopes $\beta = .138$, p < .001) and ambiguous ballots (simple slopes $\beta = .114$, p < .001). These associations are depicted in Figure 7.

**Figure 8**

*Study 5 Self-Perceived and Peer Bias Similarities of Biased Evaluation-Bias Rating Relation*
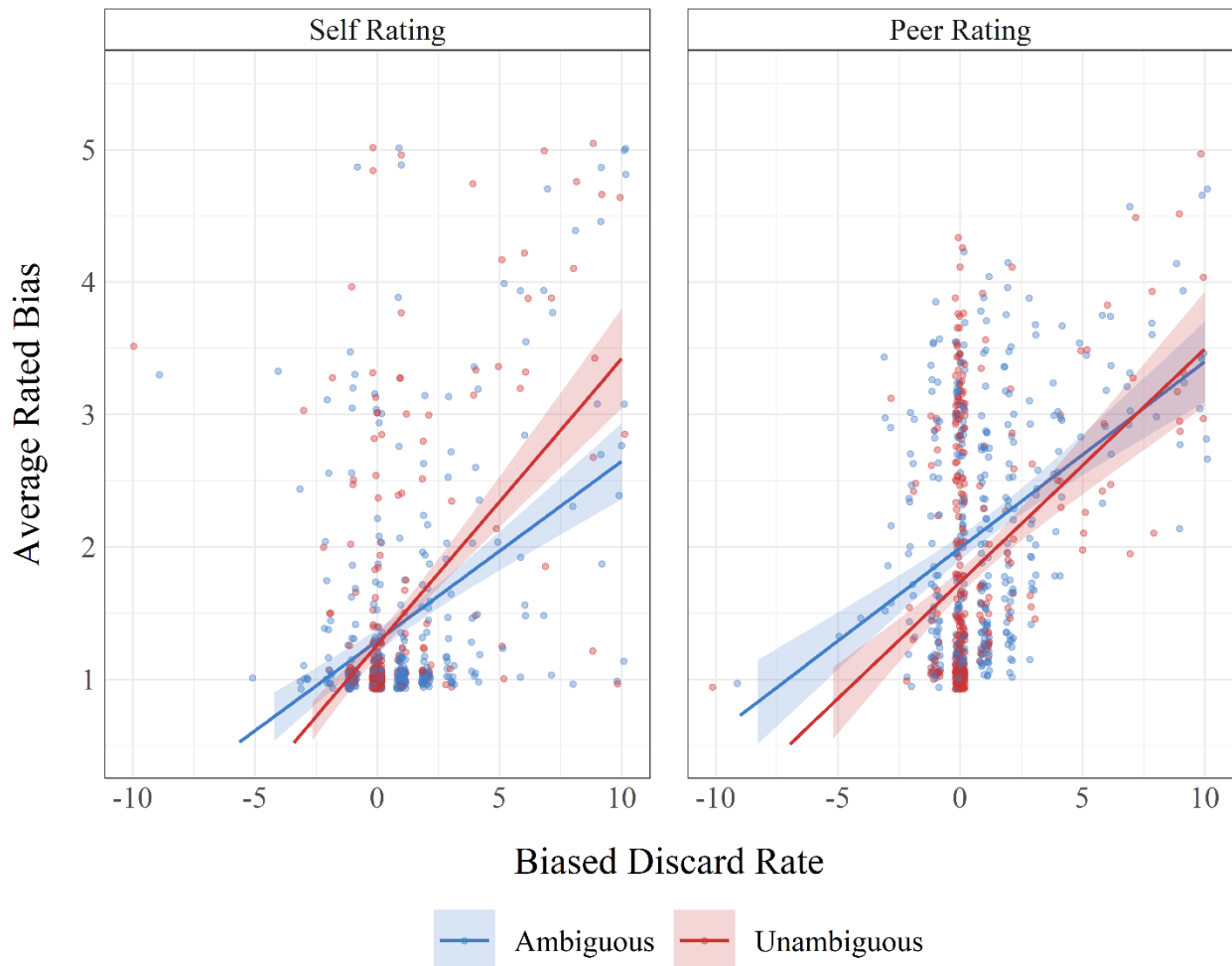


*Note.* Self and peer ratings depict similar relations between biased discarding and bias ratings for both ambiguous and unambiguous discard trials. When compared to original participant self ratings, peers reported higher ratings of bias across all scenarios but this did not vary as a function of ambiguity.

Finally, we looked at a mixed-effects model to test the whether the relation between the relation between self vs. peer perceived bias was moderated by ambiguity. This was also nonsignificant, suggesting that peers did not perceive bias differently when compared to original subjects themselves, depending on whether the ballot was ambiguous or not, $\beta = .039$, $SE = .03$, $p = .2$, 95% CI [-.021, .098].

**Study 5 Discussion**

Study 5 examined whether participants who viewed a peer's evaluations of a simulated ballot task reported similar inferences of bias. Participants were presented with responses from a previous participant named "Participant A" who has been yoked to the present participant's stimuli. Perceived bias was overall higher among these observers when compared to yoked Participant As, and higher for ambiguous responses when compared to unambiguous responses. Observers also perceived when Participant As were discarding in a more biased manner. The relation between self and peer perceptions of bias were not moderated by ambiguity. Given that we see similar awareness levels across biased discarding could suggest that original participants from Study 4 saw themselves as an "observer" that inferred bias from the pattern of responses they reported. That is, they had no insight beyond what an observer would attribute to someone given the same situational factors (Taylor & Koivumaki, 1976).

In Study 3, we demonstrated that people inferred more bias than participants did themselves on a belief updating task, and that more biased belief updating bias was predictive of more perceived bias. This study extends the results of Study 3 to a new paradigm, illustrating that participants underreport the extent to which political beliefs influence decisions on evaluating ballots relative to other people. While it is not possible to discern from the present data just how much one's beliefs are biasing people, this is further evidence that normative standards of influence are not accounting for our self-perceptions in motivated reasoning.

# CHAPTER 4

## Incentive-Compatible Test of Ballot Task Motivation

### Study 6

In Study 4, we introduced a ballot counting motivated reasoning task. Study 6 addressed the possibility that social desirability inhibited true responding to the simulated ballot task, in that people were aware that their responses were biased but were not willing to publicly admit it. Thus, we used an incentive-compatible task based on the Becker-Degroot-Marschak technique (BDM; Becker et al., 1964) to address this issue.

Incentive-compatible tasks are used in behavioral economics experiments as a way to place participants in a situation where the most rewarding outcome follows from the participant stating their authentic preferences (Groves & Ledyard, 1987). When motivated to interpret information in a way that is congenial to one's political values, cognitive inference processes may terminate quickly once people convince themselves that a congenial answer is acceptable (Chaiken et al., 1996; Epley & Gilovich, 2016; Kruglanski & Webster, 1996). If it is plausible that one's political belief is an acceptable interpretation, people may pounce on it as a response instead of reflecting on it more deeply. When used to reduce congenial responses on political belief surveys, incentives have generally been successful (Bullock et al., 2015; Prior et al., 2015). By incentivizing responses, we attempted to draw attention away from the expressive utility of providing a congenial response and replace it with an incentive to provide an authentic response.

**Method**

*Participants*

We recruited 400 participants from Prolific, including 200 Republicans and 200 Democrats. We eliminated 8 participants who did not identify as Republican or Democrat, leaving a final sample of 392 participants ($M_{age}$ = 42.9, 49.2% female).

*Procedure*

As with Study 4, participants evaluated ballots and were tasked with either keeping them or discarding them. This study presented only 10 ambiguous ballots: half benefitting Joe Biden and half benefiting Donald Trump. For each trial, participants made an initial decision to keep or discard the ballot. We create the biased discard rate index as a difference score of outpartisan ballots minus inpartisan ballots. Positive scores indicated that one engaged in more biased evaluation, rejecting outpartisan ballots at a greater rate than inpartisan ballots.

As mentioned, a BDM procedure is often used in experimental economics to facilitate honest reporting of preferences of beliefs (Becker et al., 1964; Groves & Ledyard, 1987). In a BDM procedure, participants state a preference for a monetary reward that is then compared to another value, such as a random value or some specific amount. Rather than asking for a specific preferred value, this experiment presented multiple tradeoffs between a potential reward and different fixed amounts. For each tradeoff, participants made a choice between a fixed amount and an amount that was tied to perceptions of what others thought about the ballot they are evaluating. By making a series of incentivized decisions alongside the ballot valuations, we were able to probe normative beliefs about what others thought about a particular ballot. If a participant thought that many of their peers will discard a ballot, this measure captures that normative belief.

Modeled after recent research by Graham (2023), participants made a series of tradeoffs between two bonus "lottery tickets." For each ballot, participants decided between either Lottery Ticket A or B in nine choices. Lottery Ticket A paid out an amount pegged to the sample discard rate for each ballot. For instance, if 50% of participants discarded the seen trial, this ticket would pay out $0.50. If 60% discarded, this would pay out $0.60, and so on. Lottery Ticket B paid a set bonus amount, ranging from $0.10 to $0.90 and increasing in increments of $0.10. For example, participants chose between Ticket A (matching % of discard rate) vs. Ticket B ($0.10), increasing until the choice between Ticket A (matching % of discard rate) vs. Ticket B ($0.90). This decision process is shown in Figure 8.

**Figure 9**

*Study 6 Incentivized Decisions*
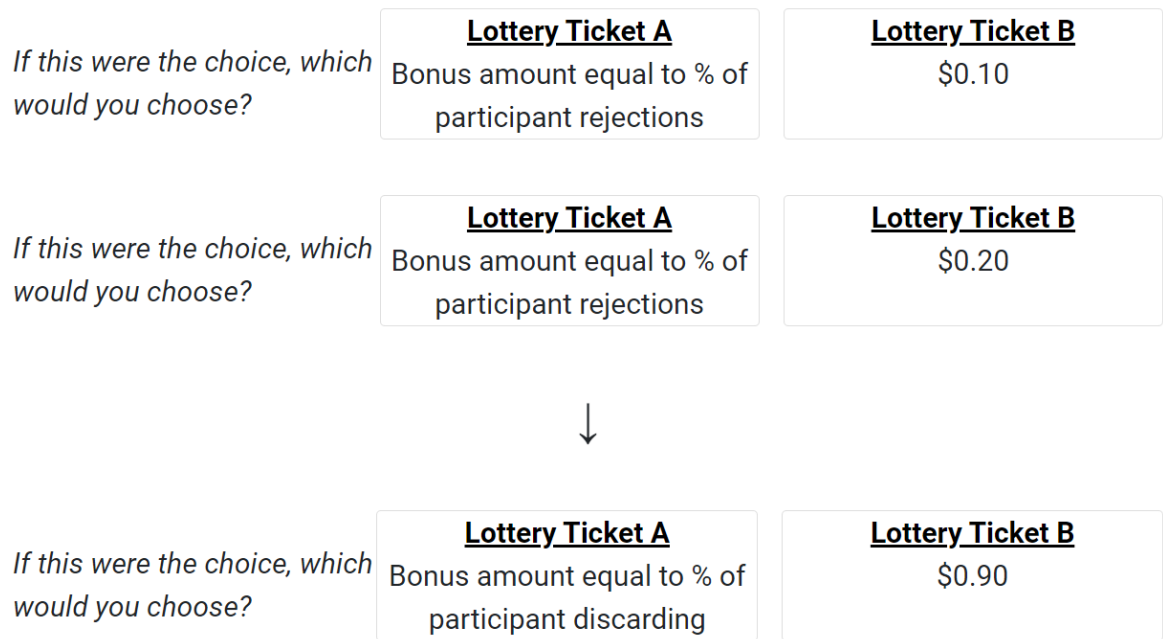


*Note.* This figure depicts the series of incentivized decisions in Study 6. For each ballot, participants made a series of 9 choices. Lottery Ticket A is always pegged to the % of sample participants who discard the ballot (e.g. always $0.50 if 50% of the sample discards the ballot).

Lottery Ticket B increases from $0.10 to $0.90 in increments of $0.10. A rational actor would choose Lottery Ticket A until they believe B exceeds the discard rate for the given ballot, at which point, they should always pick Ticket B for the rest of the series.

We considered the participant's normative threshold as the amount when participants switched from Ticket A to Ticket B. For instance, if one thought 41% of participants discarded the current ballot, they should have picked Ticket A for each tradeoff until Ticket B pays off $0.50, at which point they should have switched and chose Ticket B the rest of the way. To create an index of biased favorability, we averaged threshold values where respondents switched for ballots facing outpartisans and subtracted the average threshold value for inpartisans. Thus, a positive value represents a tendency to think that more of the sample is going to reject outpartisan ballots when compared to inpartisan ballots. We refer to this as a *favorability index* in the present results.

Finally, each ballot had a question about whether one's choices were influenced by their political beliefs on a 1 to 5 Likert scale, as in our previous studies.

**Results**

*Motivated Evaluation*

Participants on average discarded more outpartisan ballots ($M = 1.464$, $SD = 1.623$) when compared to inpartisan ballots ($M = 1.076$, $SD = 1.334$), paired t-test $t(391) = 5.843$, $p < .001$, $d = .295$.

*Favorability Index*

Did participants report a higher threshold for inpartisans than outpartisans when deciding on the normative-based incentive questions? To examine our favorability index, we first excluded 10.2% of all trials where participants responded to the series of incentive questions by

47

switching tickets in an irrational way, suggesting they did not understand the task. For instance, if someone took Ticket B instead of Ticket A for $0.10, but then took Ticket A instead of Ticket B for $0.20, this would suggest that they thought the normative average was above $0.20, but opted for a ticket worth only $0.10. We first looked at these threshold differences between inpartisan and outpartisan ballots. Overall, inpartisan ballots ($M = .373$, $SD = .252$) did not have a greater or lower threshold when compared to outpartisan ballots ($M = .385$, $SD = .26$), paired $t(379) = .994$, $p = .321$, $d = .051$.

After determining the favorability index difference score (with higher values indicating a belief that more of the sample will discard the outgroup vs. ingroup ballot), we next correlated it with the biased discard rate, revealing a moderate relation, $r(378) = .398$, $p < .001$. That is, as participants believed outpartisan (vs. inpartisan) ballots were to be seen as less legitimate by the sample, they were also more biased in rejecting the ballots themselves.

*Awareness*

As with previous studies, participants strongly underreported any amount of self-perceived bias, with a 1 out of 5 on our Likert scale accounting for 79.6% of all trials. To assess a main effect of motivated responding predicting more awareness, we fit a simple linear model, $F(1, 390) = 124.73$, $p < .001$, $R^2_{adjusted} = .24$. Participants who engaged in more motivated responding reported being more aware on average, $\beta = .37$, $SE = .033$, $p < .001$, 95% CI [.305, .435]. Much like biased discarding, the favorability index was also predictive of awareness (although less so) in a simple regression, $F(1, 378) = 14.143$, $p < .001$, $R^2_{adjusted} = .034$, with an increase in favorability index predicting greater awareness, $\beta = .962$, $SE = .256$, $p < .001$, 95% CI [.459, 1.465].

*Role of Political Identity*

We also examined a political difference between liberals and conservatives with respect to the relation between biased discarding, favorability, and awareness.

Using a multiple regression model, $F(3, 388) = 46.704$, $p < .001$, $R^2_{adjusted} = .26$, We tested the interaction of politics x biased discard index on the outcome of awareness. Political party significantly moderated the relation between biased discarding and awareness, $\beta = .159$, *SE* $= .066$, $p = .017$, 95% CI [.028, .29]. Whereas Democrats significantly showed this effect (simple slopes $\beta = .271$, $p < .001$), the effect was stronger for Republicans (simple slopes $\beta = .43$, $p < .001$). This relation is depicted in Figure 9.

When comparing political party differences on the favorability index, Democrats (*M* $= .004$, *SD* $= .178$) and Republicans (*M* $= .015$, *SD* $= .192$) did not differ significantly, Welch's $t(372.05) = .59$, $p = .555$, $d = .061$. Unlike biased discarding, a multiple regression, $F(3, 376) = 7.317$, $p < .001$, $R^2_{adjusted} = .048$, investigating a politics x favorability index interaction predicting awareness was nonsignificant, $\beta = .181$, *SE* $= .509$, $p = .723$, 95% CI [-.82, 1.181].

**Figure 10**

*Study 6 Effect of Biased Ballot Evaluation on Self-Perceived Awareness*
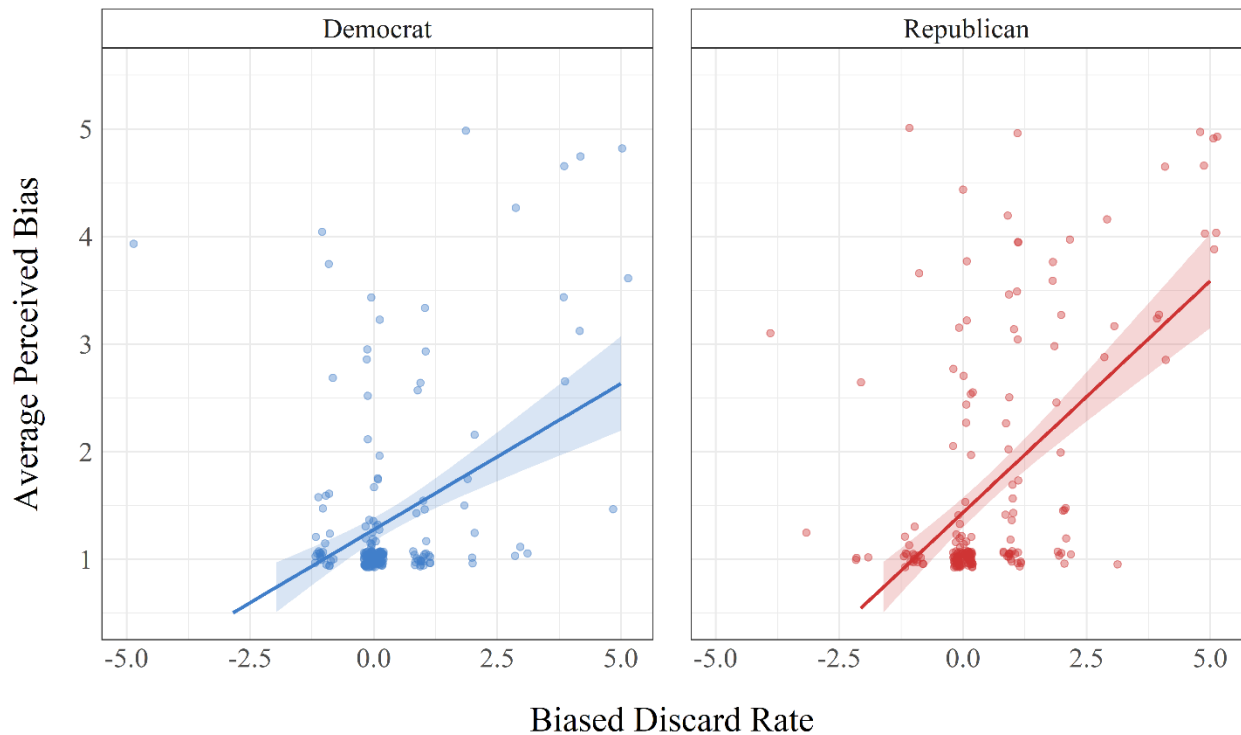


*Note.* Facets depict average self-perceived bias predicted by a participant's average biased discarding rate. Modal responses indicate low reporting of awareness. Similar to Study 4, data reveal a positive relation for both Democrats and Republicans. The relation is stronger for Republicans when compared to Democrats.

**Study 6 Discussion**

Study 6 examined the relation between biased discarding in a simulated ballot task, and whether people could report more awareness of the influence of personal political beliefs. To extend previous studies, we used an incentive-compatible task that revealed a similar pattern as Study 4 of biased responding on self-reported awareness. Participants discarded outpartisan ballots at a greater rate than inpartisan ballots. The modal response of perceived self-awareness is 1 out 5 of, suggesting that despite demonstrating some amount of motivated reasoning, participants maintain that they are seldom biased, overall. Participants who were more biased

when evaluating ballots reported higher levels of awareness when compared to less biased participants.

The favorability index measured the normative beliefs of participants, which did not differ by politics. Furthermore, the average threshold for incentive capture did not differ when evaluating inpartisan vs. outpartisan ballots. The favorability index was predictive of biased ballot evaluation, suggesting that when participants were more biased, they were honestly so. That is, when people were biased, they seemed to feel that others would see ballots similarly in the same situation. While both biased discarding and the favorability index were predictive of awareness, this relation did not differ between political identities. This differs from Study 4, where Republican—but not Democrat—participants showed more awareness when they were more biased.

In sum, the results of this study complement those of Study 4. Participants reported low self-awareness overall, but greater self-awareness when more biased. The incentive choices suggest that people generally believed that other participants would make similar evaluations of ballots. This minimizes the likelihood that participants gave responses that were expressive or inauthentic.

# CHAPTER 5

## Discussion

In this dissertation, I investigated whether people show awareness of the extent to which their decisions and reasoning processes are driven by their desires. After reviewing the psychological evidence, I hypothesized that people would generally lack the introspective access to realize the influence of their beliefs on their ability to reason objectively about information. In four chapters, I presented empirical evidence that although people engage in motivated reasoning, they do not consistently report awareness that their beliefs are influencing their evaluations. Specifically, in two different motivated reasoning paradigms, participants universally underreport the extent to which they are engaging in any biased reasoning when compared to peers judging their behavior, even when bias exists to some extent. The data show that people assert a modal self-evaluation of no bias present in all studies, despite consistently engaging in motivated reasoning.

In Chapter 2, two studies revealed a lack of any pattern for motivated belief updating in either beliefs about everyday life or beliefs about an upcoming political election. Additionally, peer ratings of these beliefs indicated that people infer patterns of motivated reasoning among others, despite there being no patterns reported about oneself. That is, while there was some association between awareness and bias within the self, this was not greater than an observer viewing the behavior. Observers also reported more perceived bias generally, while original participants gave a modal response of the lowest self-perceived bias, even when there was bias present.

In Chapter 3, I presented a simulated ballot counting task to see if people reported the influence of politics when evaluating political ballots in a biased manner. As expected, participants discarded outpartisan ballots (Democrats discarding pro-Trump ballots or Republicans discarding pro-Biden ballots) at a greater rate then inpartisan ballots (Democrats discarding pro-Biden ballots or Republicans discarding pro-Trump ballots).

Results also indicate that people reported more awareness when discarding ambiguous ballots—ballots that have clear misleading or confusing stray pencil marks—when compared to unambiguous ballots—ballots that are filled in with valid pencil marking. As biased discarding increased, the rate of awareness was positive for both ambiguous and unambiguous ballots. This suggests that at levels of high bias, there was evidence of a reality constraint: extremely biased participants were honest about their political influence when not doing so would be seen as too outlandish (Kunda, 1990). If someone discarded an outpartisan ballot that was clearly marked and seemingly valid, they mostly seemed less willing to concede that they were being biased. Curiously, this happened with ambiguous ballots at a similar rate. The lack of difference in ambiguity among biased participants does not rule out any attribution to expressive responding, where people assert themselves as lacking bias despite being evaluating the ballots favorably towards their political preferences. If it were the case that a biased participant wanted to cheerlead for their party, it would be expected that they deny any bias (Peterson & Iyengar, 2021). Although there is less awareness reported among ambiguous trials among biased participants, there was not a significant difference when compared to unambiguous trials. Similar to the results in Chapter 2, in Chapter 4 I report that peers inferred more bias than participants did themselves. Consistent with self ratings, I find that peers infer similar rates of bias on ambiguous and unambiguous ballots when looking at another person's responses.

Finally, I introduced an incentive compatible task to the ballot experiment in Chapter 5 to induce more honest responding in case there are normative pressures to not report on one's own biases. There were no differences between inpartisan and outpartisan thresholds on these normative beliefs. A favorability index—assessing the belief that the study sample would overall discard a given outpartisan vs. inpartisan ballot—revealed that as people viewed a ballot as less legitimate, they were more willing to discard it themselves. This suggests that participants made a given choice honestly, assuming that their peers would act similarly in the same situation. Our favorability index was also correlated with biased discarding, suggesting that, at the individual level, participants were honest about discarding in a biased manner.

I did not make any predictions about political party affecting biased evaluations of ballots, but exploratory analyses revealed a greater bias among Republicans when compared to Democrats. This difference was not found in the belief updating biases reported in Chapter 2. It could be the case that elections themselves have become overly politicized with strong differences between parties on election integrity (Pew Research Center, 2022). This will be an important consideration for researchers using paradigms that might be constrained by a participant's prior beliefs (Tappin et al., 2020) or assuming their measure necessarily results in symmetrical biases between political parties (Baron & Jost, 2019; Ditto et al., 2019).

This dissertation contributes to the study of motivated reasoning and how biased self-assessment plays a role in our beliefs and evaluations. The results are consistent with the many findings demonstrating lack of self-insight and introspective access to our reasoning processes. The present findings also demonstrate that those who are extremely biased in their evaluations report higher levels of bias in the ballot task, but not the belief updating task. I believe that while readers can generalize a lack of self-awareness in motivated reasoning, there might be some

domains that prompt more awareness that others. The two paradigms also demonstrate differences in extremity of bias: biased belief updating was not associated with awareness at any level, but biased ballot evaluation was acknowledged among the most exceptionally biased partisans.

<div align="center">

**Limitations and Future Considerations**

</div>

This dissertation adopted a deliberately narrow scope by establishing self-awareness through participant self-report. On the one hand, this is desirable, since how else can researchers study self-awareness other than self-report? In the past couple decades, various pressures have pushed the field of psychology to focus on conducting quick surveys as opposed to studying observed behavior in social situations (Baumeister et al., 2007). How individuals rate their beliefs on a 101-point scale might be drastically different than how they might reason in everyday situations, for instance about their health or social life. Thus, it would be useful to study motivated processes through ecological self-assessment, economic games, qualitative study, or a field experiment. Another useful extension of the present research could quantify and understand discrepancies in the decision process using indirect methods like eye- or mouse-tracking (Franco-Watkins et al., 2019). Self-awareness might not be explicitly found outside of self-report measures, but one can look for indications that people might inadvertently leave traces of reasoning in their behavioral and nonverbal tendencies.

Additionally, five of the six studies presented concern political beliefs and evaluation, specifically focused on US politics. Scholars have pointed out that for psychological findings to be generalized, they should be extended and replicated in diverse populations (Henrich et al., 2010). Research can also investigate some of the most motivating domains in society outside of politics, including health, self-esteem, or social relationships.

**Considerations for Self Deception**

I see the present research as a stepping stone to the harder problem of understanding the nature of *self deception*, or the process of intentionally misleading oneself while simultaneously concealing this process from one's awareness somehow. Self deception is often conflated with motivated reasoning in the social science literature (e.g. a deflationary account of self deception; Mele, 1997). Scholars who purport to demonstrate self deception often limit themselves to persuasion paradigms (Smith et al., 2017) or questionable indirect measures (Gur & Sackeim, 1979). In future research, I would desire to study self deception that approaches motivated decisions temporally. Philosophers describe how the self might acquire a belief in the future as a result of the present self deceiving the future self. This is called temporal partitioning (Sorensen, 1985). In temporal partitioning, the future self has no recollection of the deception, and thus is misled unknowingly.

As Balcetis (2008) describes it, "memories need time to marinate in the motivational juices." I suggest that people engage in self deception as an anticipatory dissonance reduction strategy for a future threatening event, rewriting their mental history that put them in their situation. If one anticipates some threat at a future time, they can position themselves in an environment where one either avoids a threatening stimulus or reduces its impact somehow. They then may engage in strategies that maximize the likelihood that they forget their behavior and could plausibly deny their past selves' involvement in the deception. This might function similar to situational self control, where one might avoid procrastination during study by leaving their cell phone home when going to the library (Duckworth et al., 2016). Such a phenomenon would likely follow from coherence processes found in motivated reasoning where one's cognitions about a stimulus are processed as a complex system that seeks to balance competing

thoughts and emotions (Converse, 1964; Dunning, 2007; Simon et al., 2015). I believe understanding self deception and motivated reasoning processes across a temporal dimension is crucial for understanding our biased processing and any chance of correcting it (e.g. by inducing more reflection; Milkman et al., 2009).

## Implications

Motivated reasoning likely accounts for a nontrivial amount of errors in reasoning, disagreement, and conflict. I propose that a better understanding of our reasoning processes will be especially relevant to understand given the proliferation of digital media. As social media often lets one determine a social group to surround themselves with (or provides an algorithm to do the same), people may lack any prompt for self-reflection or consideration of their potential errors. Another consideration is how our reasoning will interact with new technology like generative artificial intelligence. For instance, if I want to reason my way to a conclusion, this will be much easier if a chatbot is asked to generate arguments for (but not against) my prior beliefs. Even if artificial intelligence can be prompted to challenge our beliefs, if one lacks self-awareness, they will not think to ask for arguments against oneself. In sum, I believe that given the contentious political climate nationally and globally, a better understanding of the fundamental nature of motivated reasoning will be helpful for improving civil discourse.

# APPENDIX A

## Study 1 Scenarios and Base Rate Percentages

**Table A.1**

*Study 1 Scenario List*

| Item | Scenario | Base Rate (%) |
|------|----------|---------------|
| 1 | Attend a party in the next month | 51 |
| 2 | Cook dinner for friends in the next month | 45 |
| 3 | Donate money to a needy person or cause in the next month | 34 |
| 4 | Average 7 or more hours of sleep in a week in the next month | 86 |
| 5 | Exercise at least twice in a week in the next month | 72 |
| 6 | Finish reading a book in the next month | 40 |
| 7 | Fix a broken possession in the next month | 31 |
| 8 | Find or receive a gift of a dollar or more in the next month | 51 |
| 9 | Get a haircut in the next month | 46 |
| 10 | Get invited to a party in the next month | 51 |
| 11 | Get taken out for dinner in the next month | 52 |
| 12 | Have a sexual encounter that you enjoy in the next month | 47 |
| 13 | Have a supervisor or teacher praise your work in the next month | 51 |

| 14 | Have an out of town friend visit you in the next month | 31 |
|---|---|---|
| 15 | Have your photo taken in the next month | 73 |
| 16 | Invite a non-family member to a meal in the next month | 41 |
| 17 | Learn a new skill related to work or school in the next month | 53 |
| 18 | Make a purchase in excess of $50 for your personal enjoyment in the next month | 60 |
| 19 | Meet with your supervisor or teacher in the next month | 53 |
| 20 | Participate in a game of sport in the next month | 35 |
| 21 | Play a board game in the next month | 37 |
| 22 | Play with a pet in the next month | 72 |
| 23 | Run into an old friend that you haven't seen in a long time in the next month | 37 |
| 24 | Receive a pay check in the next month | 74 |
| 25 | Receive a compliment about how you dress in the next month | 42 |
| 26 | Shop for clothes in the next month | 55 |
| 27 | Successfully teach someone a new skill or concept in the next month | 46 |
| 28 | Take a day or more of (non-holiday) break from school or work in the next month | 65 |
| 29 | Try out a new food or dish in the next month | 63 |
| 30 | Try out a new hobby, craft, or sport in the next month | 39 |
| 31 | Go out of town for leisure in the next month | 37 |
| 32 | Wish a friend a happy birthday in the next month | 65 |

| 33 | Win a competitive game of sport in the next month | 27 |
| 34 | Burn something that you are cooking in the next month | 37 |
| 35 | Embarrass yourself in the next month | 61 |
| 36 | Have a family or friend get ill in the next month | 42 |
| 37 | Find out that someone you know personally has died in the next month | 16 |
| 38 | Get lost in the next month | 19 |
| 39 | Get rejected by someone in the next month | 15 |
| 40 | Get sick or suffered a physical illness in the next month | 35 |
| 41 | Have a disagreement with a friend in the next month | 40 |
| 42 | Have a headache in the next month | 80 |
| 43 | Hear about a natural disaster in the next month | 60 |
| 44 | Learn of a terrorist attack in the next month | 12 |
| 45 | Hurt someone's feelings in the next month | 48 |
| 46 | Become hungover from overdrinking in the next month | 27 |
| 47 | Received a phone call from a telemarketer in the next month | 57 |
| 48 | Saw a dead animal/human in the next month | 31 |
| 49 | Stay up past 2 AM for school or work in the next month | 47 |
| 50 | Was stuck in traffic in the next month | 54 |
| 51 | Was teased at or made fun of in the next month | 35 |
| 52 | Get lied to in the next month | 46 |
| 53 | Receive an unexpected bill in the next month | 26 |
| 54 | Clean the bathroom in the next month | 80 |

| 55 | Make a purchase using cash in the next month | 59 |
|----|----------------------------------------------|----|
| 56 | Bet money on a sporting event in the next year | 19 |
| 57 | Travel internationally in the next year | 71 |
| 58 | Have heart disease as your cause of death | 20 |
| 59 | Get cataracts by age 75 | 50 |
| 60 | Donate blood in the next year | 10 |
| 61 | Get a sexually transmitted infection at some point | 20 |
| 62 | Graduate from U of M with a college degree | 93 |
| 63 | Get cancer at some point in your life | 6 |
| 64 | Get cheated on in a relationship | 18 |
| 65 | Fail a class in college at some point | 30 |
| 66 | Switch political parties at some point | 9 |
| 67 | Get diabetes at some point | 11 |

# APPENDIX B

## Study 2 Election Candidates and Base Rate Percentages

Candidate A (the frontrunner) was always presented first. Items 1 – 7 have Democrat frontrunners and items 8 – 14 have Republican frontrunners. The presentation of all trials were randomly ordered.

**Table B.1**

*Study 2 Candidate Base Rates*

| Item | Election | Candidate A | Candidate B | Base Rate Candidate A (%) |
|------|----------|-------------|-------------|---------------------------|
| 1 | MI Governor | Gretchen Whitmer | Tudor Dixon | 88 |
| 2 | KS Governor | Laura Kelly | Derek Schmidt | 67 |
| 3 | NM Governor | Michelle Lujan Grisham | Mark V. Ronchetti | 84 |
| 4 | AZ Senate | Mark Kelly | Blake Masters | 65 |
| 5 | PA Senate | John Fetterman | Mehmet Oz | 54 |
| 6 | WA Senate | Patty Murray | Tiffany Smiley | 91 |

| 7 | NH Senate | Maggie Hassan | Donald C. Bolduc | 73 |
|---|---|---|---|---|
| 8 | AZ Governor | Kari Lake | Katie Hobbs | 62 |
| 9 | GA Governor | Brian Kemp | Stacey Abrams | 93 |
| 10 | OH Senate | J.D. Vance | Tim Ryan | 83 |
| 11 | GA Senate | Herschel Junior Walker | Raphael Warnock | 58 |
| 12 | NC Senate | Ted Budd | Cheri Beasley | 81 |
| 13 | WI Senate | Ron Johnson | Mandela Barnes | 78 |
| 14 | FL Senate | Marco Rubio | Val Demings | 95 |

# APPENDIX C


# Ballot Evaluation Instructions Used in Study 4, 5, and 6


## Figure C.1

*Ballot Evaluation Instructions for Studies 4, 5, and 6*

<u>**Standards for Determining What Constitutes a Vote and What Will Be Counted As a Vote**</u>

The State Board of Elections has developed these standards and procedures to define what is a vote and when that vote should be counted in circumstances in which voting systems are unable to determine the voter's intent with respect to a marked ballot. This could occur if a voter uses an improper marking instrument, marks in an inappropriate manner, places the marks in the wrong location on the ballot, or otherwise acts in a manner that causes the voter's ballot to be unreadable by the voting system for which the ballot was designed.

<u>**Definitions**</u>

- **Overvote:** A voter marks more choices than number of positions allowed in the contest.
- **Undervote:** A voter marks fewer choices than the number of positions allowed in the contest.
- **Voter Intent:** A scanner or other counting machine is unable to determine the voter's choice on an official ballot, but human counters can clearly determine the voter's choice by hand count the official ballot.

<u>**General Standards**</u>

Only the official ballots shall be counted according to the principles and rules. These principle include, but are not limited to, determining the voter's intent, counting all ballot items in which the voter's intent can be determined, and not rejecting an official ballot because of technical errors in the marking unless it is impossible to clearly determine the voter's intent. If a voter's intent can be determined, it shall be considered to be a valid vote. If a voter has done anything other than to mark the ballot properly according to ballot marking instructions provided, the ballot shall be counted unless it is impossible to determine the voter's intent. If the voter has shown consistency in marking choices on the ballot, then those choices of the voter shall be counted, but shall not be counted if it results in an overvote.

# APPENDIX D

## Sample Ballots Used in Study 4, 5, and 6

Two forms were used to counterbalance order effects: one with Trump presented first and one with Biden presented first. The Biden-first form is identical in all respects, except the order of candidates are swapped. "Ballot counted for" column refers to which candidate participants were told was awarded with a vote. All 40 ballots were presented to participants in Studies 4 and 5. Study 6 used only 10 ambiguous ballots, specified in a column below.

**Table D.1**

*Ballot Stimuli for Studies 4, 5, and 6*

| Item | Ballot | Ballot Counted For | Ballot Type | Used in Study 6? |
|---|---|---|---|---|

| | | | |
|---|---|---|---|
| 1 | State General Election Ballot — President and Vice-President, Vote for One Team. Donald J. Trump and Michael R. Pence (Republican) oval filled. | Donald Trump | Unambiguous Keep |
| 2 | State General Election Ballot — President and Vice-President, Vote for One Team. Donald J. Trump and Michael R. Pence (Republican) oval filled. | Donald Trump | Unambiguous Keep |
| 3 | State General Election Ballot — President and Vice-President, Vote for One Team. Donald J. Trump and Michael R. Pence (Republican) oval filled. | Donald Trump | Unambiguous Keep |

| | | | |
|---|---|---|---|
| 4 |  | Donald Trump | Unambiguous Keep |
| 5 |  | Donald Trump | Unambiguous Keep |
| 6 |  | Donald Trump | Unambiguous Discard |

| | | |
|---|---|---|
| 7 |  | Donald Unambiguous<br><br>Trump Discard |
| 8 |  | Donald Unambiguous<br><br>Trump Discard |
| 9 |  | Donald Unambiguous<br><br>Trump Discard |

| | | | |
|---|---|---|---|
| 10 |  | Donald Trump | Unambiguous Discard |
| 11 |  | Joe Biden | Unambiguous Keep |
| 12 |  | Joe Biden | Unambiguous Keep |

| | | | |
|---|---|---|---|
| 13 |  | Joe Biden | Unambiguous<br><br>Keep |
| 14 |  | Joe Biden | Unambiguous<br><br>Keep |
| 15 |  | Joe Biden | Unambiguous<br><br>Keep |

| | | |
|---|---|---|
| 16 |  Joe Biden | Unambiguous<br><br>Discard |
| 17 |  Joe Biden | Unambiguous<br><br>Discard |
| 18 |  Joe Biden | Unambiguous<br><br>Discard |

| | | | |
|---|---|---|---|
| 19 |  | Joe Biden | Unambiguous  Discard |
| 20 |  | Joe Biden | Unambiguous  Discard |
| 21 |  | Donald Trump | Ambiguous    X |

| | | | |
|---|---|---|---|
| 22 |  | Donald Trump | Ambiguous |
| 23 |  | Donald Trump | Ambiguous X |
| 24 |  | Donald Trump | Ambiguous |

| | | | | |
|---|---|---|---|---|
| 25 |  | Donald Trump | Ambiguous | |
| 26 |  | Donald Trump | Ambiguous | |
| 27 |  | Donald Trump | Ambiguous | X |

| # | Ballot | | | |
|---|---|---|---|---|
| 28 | State General Election Ballot — President and Vice-President | Donald Trump | Ambiguous | X |
| 29 | State General Election Ballot — President and Vice-President | Donald Trump | Ambiguous | X |
| 30 | State General Election Ballot — President and Vice-President | Donald Trump | Ambiguous | |

31

State General Election Ballot

Instructions to Voters:
To vote, completely fill in the oval(s) next to your choice(s) like this:

**Federal Offices**

**President and Vice-President**
Vote for One Team

- Donald J. Trump and Michael R. Pence — Republican
- Joseph R. Biden and Kamala Harris — Democratic
- Howie Hawkins and Angela Walker — Green Party
- Jo Jorgensen and Jeremy "Spike" Cohen — Libertarian Party
- write-in, if any

Joe Biden    Ambiguous

32

State General Election Ballot

Instructions to Voters:
To vote, completely fill in the oval(s) next to your choice(s) like this:

**Federal Offices**

**President and Vice-President**
Vote for One Team

- Donald J. Trump and Michael R. Pence — Republican
- Joseph R. Biden and Kamala Harris — Democratic
- Howie Hawkins and Angela Walker — Green Party
- Jo Jorgensen and Jeremy "Spike" Cohen — Libertarian Party
- Lizard People — write-in, if any

Joe Biden    Ambiguous

33

State General Election Ballot

Instructions to Voters:
To vote, completely fill in the oval(s) next to your choice(s) like this:

**Federal Offices**

**President and Vice-President**
Vote for One Team

- Donald J. Trump and Michael R. Pence — Republican
- Joseph R. Biden and Kamala Harris — Democratic
- Howie Hawkins and Angela Walker — Green Party
- Jo Jorgensen and Jeremy "Spike" Cohen — Libertarian Party
- write-in, if any

Joe Biden    Ambiguous

| | | | | |
|---|---|---|---|---|
| 34 |  | Joe Biden | Ambiguous | X |
| 35 |  | Joe Biden | Ambiguous | X |
| 36 |  | Joe Biden | Ambiguous | X |

| | | | | |
|---|---|---|---|---|
| 37 | State General Election Ballot — President and Vice-President | Joe Biden | Ambiguous | X |
| 38 | State General Election Ballot — President and Vice-President | Joe Biden | Ambiguous | |
| 39 | State General Election Ballot — President and Vice-President | Joe Biden | Ambiguous | |

40

**State General Election Ballot**

Instructions to Voters:
To vote, completely fill in the oval(s) next to your choice(s) like this:

**Federal Offices**

**President and Vice-President**
Vote for One Team

○ Donald J. Trump and
Michael R. Pence
Republican

● Joseph R. Biden and
Kamala Harris
Democratic

○ Howie Hawkins and
Angela Walker
Green Party

○ Jo Jorgensen and
Jeremy "Spike" Cohen
Libertarian Party

○

write-in, if any

November 3, 2020

Joe Biden     Ambiguous          X

# Bibliography

Albarracín, D., & Vargas, P. (2010). Attitudes and persuasion: From biology to social responses to persuasive intent. In *Handbook of social psychology, Vol. 1, 5th ed* (pp. 394–427). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470561119.socpsy001011

Auster, S., & Dana, J. (2021). Utilizing Strategic Ignorance in Negotiations. In *Deliberate ignorance: Choosing not to know*. The MIT Press. https://direct.mit.edu/books/edited-volume/5010/chapter/2657414/Utilizing-Strategic-Ignorance-in-Negotiations

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased Judgments of Fairness in Bargaining. *The American Economic Review*, *85*(5), 1337–1343.

Balcetis, E. (2008). Where the Motivation Resides and Self-Deception Hides: How Motivated Cognition Accomplishes Self-Deception. *Social and Personality Psychology Compass*, *2*(1), 361–381. https://doi.org/10.1111/j.1751-9004.2007.00042.x

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology*, *91*(4), 612–625. https://doi.org/10.1037/0022-3514.91.4.612

Balcetis, E., & Dunning, D. (2010). Wishful seeing: More desired objects are seen as closer. *Psychological Science*, *21*(1), 147–152. https://doi.org/10.1177/0956797609356283

Bargh, J. A. (1994). The Four Horsemen of Automaticity: Intention, Awareness, Efficiency, and Control as Separate Issues. In R. Wyer & T. Srull (Eds.), *Handbook of Social Cognition* (pp. 1–1). Lawrence Erlbaum.

Baron, J., & Jost, J. T. (2019). False Equivalence: Are Liberals and Conservatives in the United States Equally Biased? *Perspectives on Psychological Science*, *14*(2), 292–303. https://doi.org/10.1177/1745691618788876

Barron, K. (2020). Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Experimental Economics*. https://doi.org/10.1007/s10683-020-09653-z

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. https://doi.org/10.1111/j.1745-6916.2007.00051.x

Bazerman, M. H., & Sezer, O. (2016). Bounded awareness: Implications for ethical decision making. *Organizational Behavior and Human Decision Processes*, *136*, 95–105. https://doi.org/10.1016/j.obhdp.2015.11.004

Bechler, C. J., Tormala, Z. L., & Rucker, D. D. (2019). Perceiving attitude change: How qualitative shifts augment change perception. *Journal of Experimental Social Psychology*, *82*, 160–175. https://doi.org/10.1016/j.jesp.2019.02.001

Bechler, C. J., Tormala, Z. L., & Rucker, D. D. (2020). Choosing persuasion targets: How expectations of qualitative change increase advocacy intentions. *Journal of Experimental Social Psychology*, *86*, 103911. https://doi.org/10.1016/j.jesp.2019.103911

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232. https://doi.org/10.1002/bs.3830090304

Bem, D. J., & McConnell, H. K. (1970). Testing the Self-Perception Explanation of Dissonance
Phenomena: On the Salience of Premanipulation Attitudes. *Journal of Personality and Social Psychology*, *14*(1), 23–31. https://doi.org/10.1037/h0020916

Bénabou, R., & Tirole, J. (2002). Self-Confidence and Personal Motivation. *The Quarterly Journal of Economics*, *117*(3), 871–915. https://doi.org/10.1162/003355302760193913

Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality and Social Psychology*, *68*(1), 61–80. https://doi.org/10.1037/0022-3514.68.1.61

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. https://doi.org/10.1561/100.00014074

Chaiken, S., Giner-Sorolla, R., & Chen, S. (1996). Beyond accuracy: Defense and impression motives in heuristic and systematic information processing. In *The psychology of action: Linking cognition and motivation to behavior* (pp. 553–578). The Guilford Press.

Chugh, D., & Bazerman, M. H. (2007). Bounded awareness: What you fail to see can hurt you. *Mind & Society*, *6*(1), 1–18. https://doi.org/10.1007/s11299-006-0020-4

Clark, J. K., & Wegener, D. T. (2013). Message Position, Information Processing, and Persuasion. In *Advances in Experimental Social Psychology* (Vol. 47, pp. 189–232). Elsevier. https://doi.org/10.1016/B978-0-12-407236-7.00004-8

Cohen, G. L. (2003). Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs. *Journal of Personality and Social Psychology*, *85*(5), 808–822. https://doi.org/10.1037/0022-3514.85.5.808

Converse, P. E. (1964). The nature of belief systems in mass publics. In *Ideology and Its Discontent* (pp. 206–261). Free Press of Glencoe. http://www.tandfonline.com/doi/abs/10.1080/08913810608443650

Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, *43*, e28. https://doi.org/10.1017/S0140525X19001730

Cusimano, C., & Lombrozo, T. (2021). Morality justifies motivated reasoning in the folk ethics of belief. *Cognition*, *209*, 104513. https://doi.org/10.1016/j.cognition.2020.104513

Cusimano, C., & Lombrozo, T. (2023). People recognize and condone their own morally motivated reasoning. *Cognition*, *234*, 105379. https://doi.org/10.1016/j.cognition.2023.105379

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*(1), 20–33. https://doi.org/10.1037/0022-3514.44.1.20

De Houwer, J., Hendrickx, H., & Baeyens, F. (1997). Evaluative Learning with "Subliminally" Presented Stimuli. *Consciousness and Cognition*, *6*(1), 87–107. https://doi.org/10.1006/ccog.1996.0281

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., Celniker, J. B., & Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science*, *14*(2), 273–291. https://doi.org/10.1177/1745691617746796

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria

    for preferred and nonpreferred conclusions. *Journal of Personality and Social*

    *Psychology*, *63*(4), 568–584. https://doi.org/10.1037/0022-3514.63.4.568

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998).

    Motivated sensitivity to preference-inconsistent information. *Journal of Personality and*

    *Social Psychology*, *75*(1), 53–69. https://doi.org/10.1037/0022-3514.75.1.53

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational Strategies for Self-Control.

    *Perspectives on Psychological Science*, *11*(1), 35–55.

    https://doi.org/10.1177/1745691615623247

Dunning, D. (2007). Self-Image Motives and Consumer Behavior: How Sacrosanct Self-Beliefs

    Sway Preferences in the Marketplace. *Journal of Consumer Psychology*, *17*(4), 237–249.

    https://doi.org/10.1016/S1057-7408(07)70033-5

Dunning, D. (2012). Confidence considered: Assessing the quality of decisions and performance.

    In *Social metacognition* (pp. 63–80). Psychology Press.

    https://doi.org/10.4324/9780203865989

Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference:

    Are self-serving theories of success a product of motivational forces? *Journal of*

    *Personality and Social Psychology*, *69*(1), 58–68. https://doi.org/10.1037/0022-

    3514.69.1.58

Eagly, A. H., Wood, W., & Chaiken, S. (1978). Causal inferences about communicators and their

    effect on opinion change. *Journal of Personality and Social Psychology*, *36*(4), 424–435.

    https://doi.org/10.1037/0022-3514.36.4.424

Earl, A., & Albarracín, D. (2007). Nature, decay, and spiraling of the effects of fear-inducing

arguments and HIV counseling and testing: A meta-analysis of the short- and long-term

outcomes of HIV-prevention interventions. *Health Psychology*, *26*(4), 496–506.

https://doi.org/10.1037/0278-6133.26.4.496

Earl, A., Albarracín, D., Durantini, M. R., Gunnoe, J. B., Leeper, J., & Levitt, J. H. (2009).

Participation in counseling programs: High-risk participants are reluctant to accept HIV-

prevention counseling. *Journal of Consulting and Clinical Psychology*, *77*(4), 668–679.

https://doi.org/10.1037/a0015763

Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering Into the Bias Blind Spot: People's

Assessments of Bias in Themselves and Others. *Personality and Social Psychology

Bulletin*, *31*(5), 680–692. https://doi.org/10.1177/0146167204271570

Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments

produced by errors in self- or social prediction? *Journal of Personality and Social

Psychology*, *79*(6), 861–875. https://doi.org/10.1037/0022-3514.79.6.861

Epley, N., & Gilovich, T. (2016). The Mechanics of Motivated Reasoning. *Journal of Economic

Perspectives*, *30*(3), 133–140. https://doi.org/10.1257/jep.30.3.133

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences,

and correlates of attitude accessibility. In *Attitude strength: Antecedents and

consequences* (pp. 247–282). Lawrence Erlbaum Associates, Inc.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Franco-Watkins, A. M., Hickey, H. K., & Johnson, J. G. (2019). Comparing process tracing

paradigms: Tracking attention via mouse and eye movements. In *A Handbook of Process*

*Tracing Methods* (2nd ed., pp. 96–110). Routledge.

https://doi.org/10.4324/9781315160559-8

Freund, T., Kruglanski, A. W., & Shpitzajzen, A. (1985). The Freezing and Unfreezing of

Impressional Primacy: Effects of the Need for Structure and the Fear of Invalidity.

*Personality and Social Psychology Bulletin*, *11*(4), 479–487.

https://doi.org/10.1177/0146167285114013

Funder, D. C. (2012). Accurate Personality Judgment. *Current Directions in Psychological

Science*, *21*(3), 177–182. https://doi.org/10.1177/0963721412445309

Garrett, N., González-Garzón, A. M., Foulkes, L., Levita, L., & Sharot, T. (2018). Updating

Beliefs under Perceived Threat. *The Journal of Neuroscience*, *38*(36), 7901–7911.

https://doi.org/10.1523/JNEUROSCI.0716-18.2018

Garrett, N., & Sharot, T. (2017). Optimistic update bias holds firm: Three tests of robustness

following Shah et al. *Consciousness and Cognition*, *50*, 12–22.

https://doi.org/10.1016/j.concog.2016.10.013

Garrett, N., Sharot, T., Faulkner, P., Korn, C. W., Roiser, J. P., & Dolan, R. J. (2014). Losing the

rose tinted glasses: Neural substrates of unbiased belief updating in depression. *Frontiers

in Human Neuroscience*, *8*.

https://www.frontiersin.org/articles/10.3389/fnhum.2014.00639

Giner-Sorolla, R. (2001). Affective Attitudes Are not Always Faster: The Moderating Role of

Extremity. *Personality and Social Psychology Bulletin*, *27*(6), 666–677.

https://doi.org/10.1177/0146167201276003

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information Avoidance. *Journal of

Economic Literature*, *55*(1), 96–135. https://doi.org/10.1257/jel.20151245

Golman, R., Loewenstein, G., Moene, K. O., & Zarri, L. (2016). The Preference for Belief

Consonance. *Journal of Economic Perspectives*, *30*(3), 165–188.

https://doi.org/10.1257/jep.30.3.165

Graham, M. H. (2023). Measuring Misperceptions? *American Political Science Review*, *117*(1),

80–102. https://doi.org/10.1017/S0003055422000387

Greenwald, A. G. (1997). *Self-knowledge and self-deception: Further consideration*. Lawrence

Erlbaum Associates, Inc.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and

stereotypes. *Psychological Review*, *102*(1), 4–27. https://doi.org/10.1037/0033-

295X.102.1.4

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S.

(2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept.

*Psychological Review*, *109*(1), 3–25. https://doi.org/10.1037/0033-295X.109.1.3

Groves, T., & Ledyard, J. (1987). Incentive compatibility since 1972. *Information, Incentives,

and Economic Mechanisms: Essays in Honor of Leonid Hurwicz*, 48–111.

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon.

*Journal of Personality and Social Psychology*, *37*(2), 147–169.

Hahn, A., & Gawronski, B. (2019). Facing one's implicit biases: From awareness to

acknowledgment. *Journal of Personality and Social Psychology*, *116*(5), 769–794.

https://doi.org/10.1037/pspi0000155

Hahn, A., & Goedderz, A. (2020). Trait-Unconsciousness, State-Unconsciousness,

Preconsciousness, and Social Miscalibration in the Context of Implicit Evaluation. *Social

Cognition*, *38*(Supplement), s115–s134. https://doi.org/10.1521/soco.2020.38.supp.s115

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*(3), 1369–1392. https://doi.org/10.1037/a0035028

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

Hastorf, A. H., & Cantril, H. (1954). They saw a game; a case study. *The Journal of Abnormal and Social Psychology*, *49*(1), 129–134. https://doi.org/10.1037/h0057880

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hermalin, B. E., & Katz, M. L. (2009). Information and the hold-up problem. *The RAND Journal of Economics*, *40*(3), 405–423. https://doi.org/10.1111/j.1756-2171.2009.00071.x

Kahan, D. M., Hoffman, D. A., Braman, D., & Evans, D. (2012). They Saw a Protest: Cognitive Illiberalism and the Speech-Conduct Distinction. *Stanford Law Review*, *64*(4), 851–906.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86. https://doi.org/10.1017/bpp.2016.2

Kappes, A., & Sharot, T. (2019). The automatic nature of motivated belief updating. *Behavioural Public Policy*, *3*(1), 87–103. https://doi.org/10.1017/bpp.2017.11

Katz, D. (1960). The Functional Approach to the Study of Attitudes. *Public Opinion Quarterly*, *24*(2, Special Issue: Attitude Change), 163. https://doi.org/10.1086/266945

Kopko, K. C., Bryner, S. M., Budziak, J., Devine, C. J., & Nawara, S. P. (2011). In the Eye of the Beholder? Motivated Reasoning in Disputed Elections. *Political Behavior*, *33*(2), 271–290. https://doi.org/10.1007/s11109-010-9133-x

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

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, *19*(5), 448–468. https://doi.org/10.1016/0022-1031(83)90022-7

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*(2), 263–283. https://doi.org/10.1037/0033-295X.103.2.263

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*(4), 636–647. https://doi.org/10.1037/0022-3514.53.4.636

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Lau, S. (2008). Information and bargaining in the hold-up problem. *The RAND Journal of Economics*, *39*(1), 266–282. https://doi.org/10.1111/j.1756-2171.2008.00013.x

Lerman, C. (1996). BRCA1 testing in families with hereditary breast-ovarian cancer. A prospective study of patient decision making and outcomes. *JAMA: The Journal of the*

*American Medical Association*, *275*(24), 1885–1892.

https://doi.org/10.1001/jama.275.24.1885

Lieberman, M. D., Ochsner, K. N., Gilbert, D. T., & Schacter, D. L. (2001). Do Amnesics

Exhibit Cognitive Dissonance Reduction? The Role of Explicit Memory and Attention in

Attitude Change. *Psychological Science*, *12*(2).

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The

effects of prior theories on subsequently considered evidence. *Journal of Personality and

Social Psychology*, *37*(11), 2098–2109. https://doi.org/10.1037/0022-3514.37.11.2098

McCaslin, M. J., Petty, R. E., & Wegener, D. T. (2010). Self-enhancement and theory-based

correction processes. *Journal of Experimental Social Psychology*, *46*(5), 830–835.

https://doi.org/10.1016/j.jesp.2010.05.002

Mele, A. R. (1997). Real self-deception. *Behavioral and Brain Sciences*, *20*(1), 91–102.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences

between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7), 2.

https://doi.org/10.1167/4.7.2

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How Can Decision Making Be

Improved? *Perspectives on Psychological Science*, *4*(4), 379–383.

https://doi.org/10.1111/j.1745-6924.2009.01142.x

Morris, A., & Kurdi, B. (2023). Awareness of implicit attitudes: Large-scale investigations of

mechanism and scope. *Journal of Experimental Psychology: General*.

https://doi.org/10.1037/xge0001464.supp

Nater, C., & Zell, E. (2015). Accuracy of Social Perception: An Integration and Review of Meta-Analyses. *Social and Personality Psychology Compass*, *9*(9), 481–494. https://doi.org/10.1111/spc3.12194

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Nurse, M. S., & Grant, W. J. (2020). I'll see it when I believe it: Motivated numeracy in perceptions of climate change risk. *Environmental Communication*, *14*(2), 184–201. https://doi.org/10.1080/17524032.2019.1618364

Olson, M. A., & Fazio, R. H. (2001). Implicit Attitude Formation Through Classical Conditioning. *Psychological Science*, *12*(5).

Olson, M. A., & Fazio, R. H. (2002). Implicit Acquisition And Manifestation Of Classically Conditioned Attitudes. *Social Cognition*, *20*(2), 89–104. https://doi.org/10.1521/soco.20.2.89.20992

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Persson, E., Andersson, D., Koppel, L., Västfjäll, D., & Tinghög, G. (2021). A preregistered replication of motivated numeracy. *Cognition*, *214*, 104768. https://doi.org/10.1016/j.cognition.2021.104768

Peterson, E., & Iyengar, S. (2021). Partisan Gaps in Political Information and Information-Seeking Behavior: Motivated Reasoning or Cheerleading? *American Journal of Political Science (John Wiley & Sons, Inc.)*, *65*(1), 133–147. https://doi.org/10.1111/ajps.12535

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta–Cognitive Model (MCM) of

    Attitudes: Implications for Attitude Measurement, Change, and Strength. *Social*

    *Cognition*, *25*(5), 657–686. https://doi.org/10.1521/soco.2007.25.5.657

Petty, R. E., Briñol, P., Tormala, Z. L., & Wegener, D. T. (2007). The role of metacognition in

    social judgment. In *Social psychology: Handbook of basic principles, 2nd ed* (pp. 254–

    284). The Guilford Press.

Pew Research Center. (2022, October 31). Two Years After Election Turmoil, GOP Voters

    Remain Skeptical on Elections, Vote Counts. *Pew Research Center - U.S. Politics &*

    *Policy*. https://www.pewresearch.org/politics/2022/10/31/two-years-after-election-

    turmoil-gop-voters-remain-skeptical-on-elections-vote-counts/

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. https://doi.org/10.1561/100.00014127

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection

    illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*,

    *43*(4), 565–578. https://doi.org/10.1016/j.jesp.2006.05.011

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus

    others. *Personality and Social Psychology Bulletin*, *28*(3), 369–381.

Pyszczynski, T., & Greenberg, J. (1987). Toward an Integration of Cognitive and Motivational

    Perspectives on Social Inference: A Biased Hypothesis-Testing Model. In L. Berkowitz

    (Ed.), *Advances in Experimental Social Psychology* (Vol. 20, pp. 297–340). Academic

    Press. https://doi.org/10.1016/S0065-2601(08)60417-7

Rogerson, W. P. (1992). Contractual Solutions to the Hold-Up Problem. *The Review of Economic Studies*, *59*(4), 777–793. https://doi.org/10.2307/2297997

Rosenzweig, E. (2016). With Eyes Wide Open: How and Why Awareness of the Psychological Immune System Is Compatible With Its Efficacy. *Perspectives on Psychological Science*, *11*(2), 222–238. https://doi.org/10.1177/1745691615621280

Ross, R. M., Rand, D. G., & Pennycook, G. (2021). Beyond "fake news": Analytic thinking and the detection of false and hyperpartisan news headlines. *Judgment and Decision Making*, *16*(2), 484–504.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology*, *59*(2), 229–241. https://doi.org/10.1037/0022-3514.59.2.229

Schaffner, B. F., & Luks, S. (2018). Misinformation or expressive responding? What an inauguration crowd can tell us about the source of political misinformation in surveys. *Public Opinion Quarterly*, *82*(1), 135–147. https://doi.org/10.1093/poq/nfx042

Scherer, A. M., Windschitl, P. D., & Smith, A. R. (2013). Hope to be right: Biased information seeking following arbitrary and informed predictions. *Journal of Experimental Social Psychology*, *49*(1), 106–112. https://doi.org/10.1016/j.jesp.2012.07.012

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*(2), 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. https://doi.org/10.1037/0022-3514.45.3.513

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias Blind Spot: Structure, Measurement, and Consequences. *Management Science*, *61*(10), 2468–2486. https://doi.org/10.1287/mnsc.2014.2096

Sharot, T., & Garrett, N. (2022). A guideline and cautionary Note: How to use the belief update task correctly. *Methods in Psychology*, *6*, 100091. https://doi.org/10.1016/j.metip.2022.100091

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), Article 11. https://doi.org/10.1038/nn.2949

Sherman, D. K., Cohen, G. L., Nelson, L. D., Nussbaum, A. D., Bunyan, D. P., & Garcia, J. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology*, *97*(5), 745–764. https://doi.org/10.1037/a0015451

Simon, D., Stenstrom, D. M., & Read, S. J. (2015). The coherence effect: Blending cold and hot cognitions. *Journal of Personality and Social Psychology*, *109*(3), 369–394. https://doi.org/10.1037/pspa0000029

Smith, M. K., Trivers, R., & von Hippel, W. (2017). Self-deception facilitates interpersonal persuasion. *Journal of Economic Psychology*, *63*, 93–101. https://doi.org/10.1016/j.joep.2017.02.012

Sorensen, R. A. (1985). Self-Deception and Scattered Events. *Mind*, *94*(373), 64–69.

Taber, C. S., Cann, D., & Kucsova, S. (2009). The Motivated Processing of Political Arguments. *Political Behavior*, *31*(2), 137–155. https://doi.org/10.1007/s11109-008-9075-8

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs.
*American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Tappin, B. M., Pennycook, G., & Rand, D. G. (2020). Bayesian or biased? Analytic thinking and political belief updating. *Cognition*, *204*, 104375. https://doi.org/10.1016/j.cognition.2020.104375

Taylor, S. E., & Koivumaki, J. H. (1976). The perception of self and others: Acquaintanceship, affect, and actor-observer differences. *Journal of Personality and Social Psychology*, *33*(4), 403–408. https://doi.org/10.1037/0022-3514.33.4.403

Tetlock, P. E. (1983). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, *46*(4), 285–292. https://doi.org/10.2307/3033716

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471. https://doi.org/10.1037/0033-295X.109.3.451

Trivers, R. (2011). *The folly of fools: The logic of deceit and self-deception in human life*. Basic Books.

Van Boven, L., Kamada, A., & Gilovich, T. (1999). The perceiver as perceived: Everyday intuitions about the correspondence bias. *Journal of Personality and Social Psychology*, *77*(6), 1188–1199. https://doi.org/10.1037/0022-3514.77.6.1188

Van Boven, L., Ramos, J., Montal-Rosenberg, R., Kogut, T., Sherman, D. K., & Slovic, P. (2019). It depends: Partisan evaluation of conditional probability importance. *Cognition*, *188*, 51–63. https://doi.org/10.1016/j.cognition.2019.01.020

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300. https://doi.org/10.1037/a0017908

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, *95*(5), 1202–1216. https://doi.org/10.1037/a0013314

von Hippel, W., & Trivers, R. (2011). Reflections on self-deception. *Behavioral and Brain Sciences*, *34*(1), 41–56. https://doi.org/10.1017/S0140525X10003018

Walster, E., Aronson, E., & Abrahams, D. (1966). On increasing the persuasiveness of a low prestige communicator. *Journal of Experimental Social Psychology*, *2*(4), 325–342. https://doi.org/10.1016/0022-1031(66)90026-6

Walters, D. J., & Fernbach, P. M. (2021). Investor memory of past performance is positively biased and predicts overconfidence. *Proceedings of the National Academy of Sciences*, *118*(36), e2026680118. https://doi.org/10.1073/pnas.2026680118

Wegener, D. T., & Carlston, D. E. (2005). Cognitive Processes in Attitude Formation and Change. In *The handbook of attitudes* (pp. 493–542). Lawrence Erlbaum Associates Publishers.

Wegener, D. T., Clark, J. K., & Petty, R. E. (2018). Cognitive and Metacognitive Processes in Attitude Formation and Change. In *The Handbook of Attitudes* (2nd ed.). Routledge.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious* (pp. viii, 262). Belknap Press/Harvard University Press.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In *The construction of social judgments* (pp. 37–65). Lawrence Erlbaum Associates, Inc.

Wu, K., & Dunning, D. (2018). Hypocognition: Making Sense of the Landscape beyond One's

    Conceptual Reach. *Review of General Psychology*, *22*(1), 25–35.

    https://doi.org/10.1037/gpr0000126

Wu, K., & Dunning, D. (2020). Hypocognition and the Invisibility of Social Privilege. In S. R.

    Thye & E. J. Lawler (Eds.), *Advances in Group Processes* (pp. 1–23). Emerald

    Publishing Limited. https://doi.org/10.1108/S0882-614520200000037001

Zell, E., Strickhouser, J., Sedikides, C., & Alicke, M. (2020). The Better-Than-Average Effect in

    Comparative Self- Evaluation: A Comprehensive Review and Meta-Analysis.

    *Psychological Bulletin*, *146*. https://doi.org/10.1037/bul0000218