THE BIAS BLIND SPOT AND MAKING

OBJECTIVE DECISIONS DESPITE IT

Katherine E. Hansen

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF PSYCHOLOGY

Advisor: Emily Pronin

June 2013
Abstract

People are quick to point out bias in others, yet are reticent to admit to bias in themselves. This dissertation explores the depths of this “bias blind spot” while putting forth a manipulation to overcome it. A series of studies first explores the psychological effects of knowingly using a biased decision-making strategy. Despite recognizing bias in their strategy, participants still saw their decision outcome as objective (even though they had actually been biased). After investigating an alternative mechanism, the current research suggests that this finding is due to bias occurring unconsciously.

When judging others, people focus on the person’s behavior for signs of bias. However, when judging the self, people look inwardly and mistakenly rely on introspection. This dissertation puts forth that this self-other asymmetry in bias attribution could be manipulated to increase objectivity. That is, when trying to make a decision that would appear objective to an observer, participants made more objective decisions than when asked to simply be objective or when left to make their own decision. In a series of studies, I demonstrate the efficacy of asking people to make a decision that appears objective. I then examine a downstream consequence of the manipulation to appear objective, presenting evidence that it improves interpersonal behavior. I then establish that this manipulation does not lead to reactance in private beliefs. Finally, a pair of studies establish how the appear-objective manipulation leads to greater willingness to blind oneself to potentially biasing information, indicating its usefulness in situations where there is not a clearly objective decision to make.

Despite showing people’s pervasive tendency to see themselves and the decisions that they make in an objective light, this dissertation puts forth a successful manipulation to increase the objectivity of people’s decisions. By shifting the focus of attention away from internal
thoughts and intentions and onto external behavior, the objectivity of people’s decisions can be increased despite the persistent existence of the bias blind spot. With a simple instruction, people can overcome a host of biases in their decisions, from sexism to self-serving bias.
# Table of Contents

Abstract iii

Table of Contents v

Acknowledgements vii

Preface viii

Introduction 1

Chapter 1: Awareness of Bias 10

  Study 1.1: Self-Serving Bias in Test Evaluation 12
    *Figure 1.1* 16
    *Figure 1.2* 17

  Study 1.2: Artistic Evaluation 18
    *Figure 1.3* 21
    *Figure 1.4* 22

  Study 1.3: The Possibility of Correction 23
    *Figure 1.5* 27
    *Table 1.1* 29

  Study 1.4: Conscious Cues 30
    *Figure 1.6* 32

Discussion 33

Chapter 2: Decreasing Bias through Instructions to Appear Objective 37

  Study 2.1a: Sexism in the Technology Industry 39
    *Figure 2.1* 43

  Study 2.1b: Specificity of Manipulation 44

  Study 2.1c: Toward Real-World Conditions 47
Acknowledgements

I first want to thank my advisors, Emily Pronin and Alex Todorov, for nurturing my interests and training me to be a researcher. Your presence and input were an invaluable part of my experience at Princeton and will continue to have a lasting impact on me. Also, thank you to the rest of the members of my committee – Uri Hasson, Joel Cooper, and Alin Coman – for supporting me in this process.

My cohort at Princeton was by chance exceptionally large, affording me the opportunity to make many great friends. First, I could not have asked for two better officemates to have for five years than Deborah Son Holoien and Mike North. It was truly wonderful to have you to bounce ideas off of and to commiserate with. Second, I would particularly like to mention Sarah Getz Lob, Brooke Macnamara, and Abby Sussman for being such amazing friends for the past five years. Others not in my cohort, Sara Verosky, Andreana Kenrick, Lily Jampol, Mary Steffel, and Meg Gerbasi – you all undoubtedly improved my experience not only in the department, but in my overall life. I am very thankful for the collection of friends I will take away from Princeton.

Finally, I would like to thank my family and my friends outside of Princeton. You have always been so supportive of my decision to pursue a PhD and I love you for that – some of you even managed to make the trip to New Jersey!
Preface

This dissertation represents work from across my career at Princeton. Chapter 1 is adapted from an empirical article currently under review that was co-authored with Emily Pronin, Margaret Gerbasi, and Alex Todorov, “Unaffected by bias: the effects of knowingly utilizing a biased strategy.” Chapter 2 contains new, unpublished research that presents and tests a manipulation to increase the objectivity of people’s decisions taking into consideration the extent of the bias blind spot presented in Chapter 1.

Much of this research has been presented at academic conferences. Experiment 1.1 was presented at the annual meeting of the Society for Personality and Social Psychology (SPSP) in Las Vegas, NV in 2010. Experiments 1.1 and 1.2 were presented at SPSP in San Antonio, TX in 2011 and at the annual meeting of the Association for Psychological Science in Washington, DC in 2011. Experiments 1.2, 1.4, 2.1, and 2.4 were presented at SPSP in New Orleans, LA in 2013.
Introduction

Despite the fact that we each view the world through our own subjective lens, we put great value in objectivity. To be called “biased” is neither a complement nor something to strive for. Even with this valuation of objectivity, research shows that people frequently and systematically fall prey to a large host of biases.

People have a pervasive tendency to see themselves in an overly-rosy light, giving way to a number of self-enhancement biases. For example, despite the statistical impossibility, people reliably rate themselves as friendlier, smarter, and more honest than the average person – a bias known as the “better than average effect” (Alicke, 1985). When making such ratings, people interpret the meanings of traits in idiosyncratic and self-serving ways (Dunning, Meyerowitz, & Holzberg, 1989). Not even academics are impervious to this bias: one study found that over 90% of college professors rated themselves as above average teachers (Cross, 1977). In fact, people are often so inaccurate when assessing their own skills that they fail to notice their own incompetence (Dunning, Johnson, Ehrlinger, & Kruger, 2003).

Individuals are similarly self-serving when evaluating their behavior in comparison to others. People are biased toward giving themselves more credit than they are due (Ross & Sicoly, 1979). When evaluating their contributions to a joint effort, for instance, people see themselves as having contributed more than others think they did. Each person within a romantic couple is biased to think that he/she is responsible for more in the relationship, from planning activities to carrying on conversations (Thompson & Kelley, 1981).

Research has shown that people’s behavior is biased by their social situation. The mere presence of other people can lead to failure to intervene as a bystander when witnessing an
emergency (Darley & Latane, 1968). People will also conform to the behavior of those around them, giving an incorrect answer on a clear perceptual task simply because others present responded with that wrong answer (Asch, 1951). Interestingly, people are biased to believe that they would be less biased by the social situation than others; for example, claiming that they would be more likely than others to rebel in the Milgram obedience studies (Bierbrauer, 1976).

Additionally, people are often biased when it comes to making predictions about their own lives or future actions. When predicting their reactions to future events (known as affective forecasting), people predict that the events will impact them for longer than they actually do (Wilson & Gilbert, 2003). For instance, people have a tendency to overestimate the duration of their reaction to negative events, because they underestimate the operating of their psychological immune system, a tendency called immune neglect (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998).

People display overoptimism about the future course of their romantic relationships (Buehler, Griffin, & Ross, 1995) and their likelihood of attaining positive life events (Pronin, 2009). This overoptimism even extends to chance events – people assume that a desired outcome was brought about by their own behavior, even in cases where that was not the case, an effect known as the illusion of control (Langer, 1975).

People similarly believe that their future behavior will be more selfless than it actually is (Epley & Dunning, 2000) – a tendency tied to a general failure to realize that they often do not follow through with their good intentions (Kruger & Gilovich, 2004). Research shows that people similarly fail at predicting how long it will take them to complete a task, reliably underestimating their completion times, a bias known as the planning fallacy (Buehler, Griffin,
& Ross, 1994). People generally fail to take into consideration possibilities that are not in line with their perceptions at a given moment (Lord, Lepper, & Preston, 1984).

These biases in behavior can have serious social consequences. For example, while people may espouse egalitarian and unbiased beliefs, their implicit attitudes can display bias against outgroups (e.g., Devine, 1989; Greenwald & Banaji, 1995). In one study, white participants erroneously thought that they had just had a successful, friendly interaction with a black participant (Dovidio, Kawakami, & Gaertner, 2002). In reality, their implicit bias was revealed through their nonverbal behavior, leading to an uncomfortable interaction for the black participant.

While a whole host of biases have been empirically documented, research suggests that removing these biases is no simple task. Previous theory has said that in order to overcome bias, a person must be aware of the bias, motivated to correct it, and capable of controlling it (Wilson & Brekke, 1994). Problems are manifold with each of these steps, making it unlikely that people will successfully triumph over bias. Empirical attempts to de-bias people are often unsuccessful, or met with only limited, domain-specific, or short-lived success (e.g., Wegener, Clark, & Petty, 2006).

The Bias Blind Spot

How is it that as a society we can place great value in objectivity, yet we are continually biased in our own behavior? Extensive research has been conducted examining people’s perceptions of bias and objectivity in themselves and in others. While people are quick to point out bias in those around them, they are not inclined to acknowledge those same biases in themselves (Pronin, Lin, & Ross, 2002). That is, people see more bias in others than they do in themselves. For example, when given a list of biases, people will report that they personally are
far less likely to fall prey to them compared to the average person, or even compared to the
people in the same room.

This reluctance to see bias in oneself is referred to as the “bias blind spot” (for a review, see Pronin, 2007). The bias blind spot is so strong, that even when people are informed about how they could have been affected by a bias, they still maintain faith in their objectivity (Pronin, Lin, & Ross, 2002).

This bias blind spot can be traced back to differences in the information that people use when assessing bias in themselves compared to others (Ehrlinger et al., 2005; Pronin et al., 2004; Van Boven et al., 1999). When evaluating others, people focus outward on their behavior, but when evaluating themselves, they look inward, relying on introspection. This difference in focus is driven by an introspection illusion, or tendency to disproportionately weight introspective information (over behavioral information) when assessing oneself, but not others (e.g. Pronin, 2008, 2009; Pronin, Berger, & Molouki, 2007). In other words, there is a self-other asymmetry in the valuation of introspective versus behavioral information when evaluating bias.

In fact, people value internal information more for themselves than for others even when given access to others’ internal information (Pronin & Kugler, 2006). Researchers had actors list their thoughts as they made a series of trait ratings about themselves. They then gave observers access to that information and asked them to indicate the degree to which the observer was guilty of displaying the better-than-average effect. While the actors themselves did not believe they were biased, the observers did – despite having access to the actors’ ongoing thoughts. Observers consistently rated the actors as more biased than the actors rated themselves, and this tendency remained the same regardless of whether or not the observer was given access to the actor’s introspections.
In a similar vein, additional research shows that people feel that there is a stronger tie between unobservable intentions and external actions for themselves than for others (Kruger & Gilovich, 2004). This self-other asymmetry, therefore, is not simply guided by a difference in what kind of information is accessible, but in how people value internal information for themselves versus others.

This reliance on introspection as a source of knowledge about bias for the self is misplaced because bias typically operates outside of conscious awareness and is therefore inaccessible via introspection. Nisbett and Wilson (1977) famously argued that people lack access to their mental processes and only have access to their mental contents (which are the outputs of those processes). Since bias is often an unconscious mental process, there will be no introspective evidence of its occurrence (e.g., Dawson, Gilovich, & Regan, 2002; Wilson, Centerbar, & Brekke, 2002). By focusing inward on their mental contents (thoughts, feelings, intentions, etc.), people thus fail to recognize the influence of biased mental processes on their behavior (i.e., their decisions). People are relying on introspection to gain information about themselves without realizing its limitations (Wilson, Hodges, & LaFleur, 1995).

Further evidence of the introspection illusion as the source of the bias blind spot comes from a study by Pronin and Kugler (2006). They provided participants with an article about the role of unconscious processes in guiding judgment. After this education that shed light on the unreliability of introspection, participants no longer exhibited the bias blind spot.

Chapter 1 of this dissertation suggests that the bias blind spot runs deeper than previously realized; even when individuals acknowledge that they are doing something in a biased manner, they are still inclined to see themselves (and the decisions that they make) as objective. A series of experiments investigated the psychological effects of knowingly using a biased decision-
making strategy. Despite recognizing bias in their strategy, participants still saw their decisions as objective (even though they had actually been biased). The current research suggests that this is due to bias operating unconsciously.

Chapter 2 introduces a manipulation aimed at increasing the objectivity of decisions that can be implemented across a variety of domains despite the existence of the bias blind spot. Past research shows that others are often more accurate about us than we are. For example, while people display the planning fallacy (exhibiting overoptimism) when estimating their own completion times, they do not underestimate the completion times of others (Buehler, Griffin, & Ross, 1994). People were more likely to use diagnostic past behavioral information when forecasting the completion times of others than themselves, leading to less overoptimism. Similarly, while people are overoptimistic about their own likelihood of future charitable giving, they are far more accurate at estimating the future charitable giving of others (Epley & Dunning, 2000). When evaluating others (but not the self), people take into consideration behavioral base rates.

This greater accuracy is due to the fact that when assessing others, we actually take into account the diagnostic behavioral information that we ignore when assessing ourselves (Pronin & Kugler, 2006). The current research takes advantage of this by having participants consider the perspective of an observer when making their own decision; that is, asking them to make a decision that would appear objective to another. At first glance, this method emphasizes the objectivity of the participant’s actual decision, regardless of their personal beliefs. In many real-life situations – such as hiring – what is most important is that an individual makes an objective decision, regardless of what biases he or she may bring to the situation.
On the surface, this research resembles manipulations of accountability. Accountability is the “implicit or explicit expectation that one may be called on to justify one’s beliefs, feelings, and actions to others” (for a review, see Lerner & Tetlock, 1999; pg. 255). While both manipulations lead people to consider the perspective of an outsider, the current research will be differentiated from work on accountability in terms of the amount of effort people are putting into their decision-making. Accountability can be beneficial to decision-making by leading people to put more effort into their decision-making process. Though accountability is a complex construct, one of its primary goals is to change cognitive processing by inducing self-critical and effortful thinking in decision-makers.

Lerner and Tetlock’s (1999) review of accountability suggests that it is a multifaceted phenomenon that is only sometimes beneficial. At times accountability attenuates cognitive biases, other times it has no effect, and sometimes it even amplifies biases.

While being held accountable may be effortful, making a decision that appears objective is actually quite simple: merely disregard one’s personal belief and make a decision that matches a normative criterion for objectivity. I hypothesize that even though people instructed to appear objective will not be increasing the effort they put into their decision-making, there will still be positive outcomes for their decisions.

**Overview of Studies**

This dissertation consists of two chapters of empirical research. Chapter 1 presents four studies looking at the psychological effects of knowingly using a biased decision-making strategy. Studies 1.1 and 1.2 will demonstrate these effects across two bias domains: biases in test evaluation and artistic evaluation. In each of these studies, despite knowing about bias in their decision-making strategy, participants still see their decisions as objective (despite having
been biased). Studies 1.3 and 1.4 will explore the mechanism underlying these mistaken claims of objectivity, first investigating an alternative mechanism (the possibility that participants are attempting to correct for their biases) and then suggesting that this finding is due to the unconscious nature of bias.

In Chapter 2, I aim to establish the efficacy of a manipulation to increase the objectivity of people’s decisions despite the existence of the bias blind spot by asking people to appear objective. Given that people can easily point out bias and objectivity in others, I will ask them to make decisions that will appear objective to an outside observer. I will compare the effectiveness of asking people to appear objective to a control condition and to a condition where people are simply asked to be objective, which, as mentioned earlier, is generally unsuccessful. First, I will look at the effectiveness of this manipulation in the domain of sexism (Studies 2.1a-2.1c), followed by a self-serving bias, the better-than-average effect (Study 2.2), and then prediction bias in achievement (Study 2.3). Additionally, I will examine the limits of this manipulation, establishing circumstances under which it is most successful.

Studies 2.1c and 2.2 will distinguish the present work from research on accountability, which leads people to increase the amount of effort they put into their decisions-making (Lerner & Tetlock, 1999). I hypothesize that instructions to appear objective do not lead to increases in effort. In fact, given the generally straight-forward nature of making a decision that would appear objective to an observer, I hypothesize that people will put in equal or less effort compared to those asked to be objective or in a control.

After establishing the success of the appear-objective manipulation, I examine its potential downstream consequences in Study 2.4. What effect does asking someone to appear objective on an outward decision have on their private beliefs? Will it influence their
interpersonal behavior? It is important to understand the potential consequences (positive or negative) of this manipulation to fully understand its effectiveness. This study will also bring the appear-objective manipulation into a new bias domain: political partisan bias.

Finally, I will examine the possible role of this manipulation in situations lacking a clearly objective response by looking at whether asking people to appear objective increases their willingness to blind themselves to potentially biasing information. Not every decision has an answer that clearly appears objective; however, if appearing objective leads people to remove the source of bias, they will still make more objective decisions. These studies will be in the bias domains of racial discrimination in criminal sentencing (Study 2.5) and bias in artistic evaluation (Study 2.6).

Together, this research sheds light on the extent to which people are inclined to wrongly perceive their decisions as objective. Despite this pervasive tendency, a practical manipulation to increase the objectivity of decisions is introduced. This manipulation is explored across a number of bias domains, and its limitations and consequences are investigated. With a new perspective, people can overcome bias in their decisions. This manipulation has practical significance for real decision-makers, from human resources officials to college admission officers.
Chapter 1

Knowingly Using a Biased Decision-Making Strategy

Following rejection after a job interview, one might seek consolation from friends. Suppose a sympathetic friend points out that although you had been excited about the job, you also had some reservations. She prompts you to focus on those concerns, while putting aside the positives. You realize that this is a biased exercise, but you go along. Afterwards, you come to feel that things worked out for the best. Moreover, even though you recognize that your exercise was biased, you think it led to the objectively right conclusion.

Past research has made clear that people fail to see the bias in their judgments when those judgments result from unknowingly engaging in a biased thought process. But what happens when one knowingly engages in such a process, as in the above scenario? The scenario differs from one in which a person “automatically” engages in a biased process—for example, automatically focusing on the downsides of a job after bombing the interview. When people’s judgments result from biased strategies that are recruited automatically, they are blind to the bias that those strategies produce (e.g., Ehrlinger, Gilovich, & Ross, 2005; Pronin, Gilovich, & Ross, 2004; Wilson, Centerbar, & Brekke, 2002). In this chapter, it is hypothesized that even when people are aware that their decision-making strategy preceding a particular judgment is biased, they still tend to view the judgment they reach as objective (even if that judgment is, in fact, biased).

The particular sort of biases that this chapter concerns are those that actors themselves would label as biases (regardless of whether some normative criterion for objectivity is violated; see, e.g., Pronin, Gilovich, & Ross, 2004). An important question, then, is: Why might an actor label a decision-making strategy as biased and then claim to have made an unbiased judgment after using that strategy?
Although people may be able to acknowledge the bias inherent in the explicit procedure or strategy they use to make a judgment or decision, they do not have access to the biased internal mental processes that accompany that procedure (Nisbett and Wilson, 1977). For example, while one may recognize that concentrating on the negatives of an unattained job is biased, one is unlikely to feel that bias influencing one’s judgment. The research in this chapter predicts that when people recognize bias in their decision-making or judgment strategies, they still may sincerely (but often incorrectly) claim personal objectivity.

There is another possibility, though, for why people might deny bias in such circumstances. By virtue of being aware that their decision-making strategy is biased, they might attempt to correct for any bias in their ultimate judgment. Those correction efforts might make them feel that they have been objective. This mechanism could be responsible even if participants fail to show increased objectivity, since conscious efforts at bias correction are notoriously unreliable (e.g., Wilson & Brekke, 1994; Wegener, Clark, & Petty, 2006).

This chapter presents four studies investigating the psychological effects of making a judgment while knowing that the strategy behind it is biased. In Study 1.1, participants received false feedback on an alleged test of their social intelligence. They then were instructed to evaluate the test using either an explicitly biased strategy (e.g., evaluating the test solely based on its weaknesses), an explicitly objective strategy, or an unrestricted strategy. We predicted that participants in the explicitly biased condition would rate the evaluative procedure as biased, engage in it, and then claim to have been objective. In Study 1.2, participants rated various paintings using either a strategy that they viewed as explicitly biased (looking at the identity of the painter before assessing a painting’s inherent quality) or a strategy that they viewed as explicitly objective (shielding themselves from the painters’ identities). We again predicted that
participants in the explicitly biased condition would rate the evaluative procedure as biased, engage in it, and then claim to have been objective (a claim that we predicted would be undermined by their ratings of the paintings).

Studies 1.3 and 1.4 investigated underlining mechanism. Study 1.3 sought evidence for the competing hypothesis that participants denied their bias because they were initially conscious of it, and thus believed they had corrected for it. Study 1.4 more directly examined our proposed mechanism involving reliance on introspective cues to bias, by testing whether the effect observed in our previous experiments would disappear when conscious cues to bias were present.

**Study 1.1**

**Method**

**Participants.** A total of 101 Princeton undergraduates (38% male; 56% Caucasian) participated for course credit. The median age was 19 ($SD = 1.20$). Seven participants were excluded prior to data analysis because they strongly suspected that their test score was phony.

**Materials and Procedure.** Participants arrived at the laboratory and took a purported newly-developed test of “Social Intelligence.” The test involved 18 photographs of people’s faces and 18 statements allegedly written by those individuals about their hobbies, careers, and preferences (See Appendix A); participants were asked to match each face with the correct self-description. Afterward, they received false feedback on their performance. They then were asked to evaluate the quality of the test. Participants in three of the four experimental conditions received failure feedback (a score of 6 correct matches, ranking in the 27th percentile). They then were randomly assigned to evaluate the test using one of three strategies: In the explicitly objective condition, they were asked to list three strengths and three weaknesses of the test; in
the unrestricted condition, they were asked to list strengths and weaknesses at their discretion; in
the explicitly biased (weakness-focus) condition, they were asked to list only weaknesses.

Previous work using this alleged social-intelligence test has demonstrated that
individuals’ evaluations of it are self-servingly biased such that individuals who believe that they
have performed poorly evaluate the test more negatively than individuals who believe they have
performed well (Pronin et al., 2002; Pronin & Kugler, 2007). In order to be sure that this bias
emerged in the current experiment, we included a fourth experimental condition in which
participants were led to believe that they performed well. Those participants were given success
feedback (a score of 14 correct matches, 80th percentile) and asked to evaluate the test in an
explicitly positively-biased way—i.e., by listing only strengths of the test (explicitly biased
strength-focus condition).

Before engaging in their assigned strategy for evaluating the test, all participants rated the
objectivity of their assigned strategy on a scale of 1 (completely biased) – 9 (completely
objective). After engaging in that strategy, they provided a scale rating of the quality of the test
from 1 (very bad test) – 9 (very good test). Lastly, they rated the objectivity of the rating that
they had provided on a scale of 1 (completely biased) – 9 (completely objective).

Results and Discussion

Manipulation check: Adherence to strategy. The number of strengths and weaknesses
that participants listed when evaluating the test was coded as a manipulation check of their
adherence to their assigned evaluative strategy. Participants complied with the strategy to which
they were assigned, with the exception of those in the explicitly objective condition. In the
explicitly biased weakness-focus condition, participants listed more weaknesses of the test
($M = 2.12, SD = 1.09$) than strengths ($M = 0.04, SD = 0.20$), $t(24) = 9.66, p < .0001$. In the
explicitly biased strength-focus condition, participants listed more strengths ($M = 3.38, SD = 1.38$) than weaknesses ($M = 0.13, SD = 0.45$), $t(23) = 9.66, p < .0001$. Participants in the unrestricted condition responded in a manner consistent with the presence of an ego-protective bias: after receiving failure feedback, they listed more weaknesses than strengths ($M = 2.77, SD = 1.54$, vs. $M = 1.45, SD = 0.96$), $t(21) = 3.70, p = .001$. In the explicitly objective condition, participants were asked to write an equal number of strengths and weaknesses (three of each); however, they wrote slightly more weaknesses than strengths ($M = 3.22, SD = 0.42$, vs. $M = 3.04, SD = 0.37$), $t(22) = 2.15, p = .04$.

**Assessments of bias.** Our primary prediction concerned participants in the two explicitly biased conditions. Those participants either received failure feedback and then were asked to evaluate the test by focusing on its weaknesses or they received success feedback and then were asked to evaluate the test by focusing on its strengths. We predicted that both sets of participants would: 1) Rate their assigned strategy as biased compared to participants in the other conditions (i.e., the unrestricted condition and the explicitly objective condition), and 2) Fail to rate their evaluation of the test as biased compared to participants in those other two conditions. Thus, we predicted a 4x2 interaction effect (Evaluation Strategy: *Explicitly Biased–Weakness Focus, Explicitly Biased–Strength Focus, Explicitly Objective, Unrestricted* x Bias Assessment: *Self-Assessment, Assessment of Strategy*).

The predicted interaction effect was tested using a Mixed-Model Analysis of Variance, with evaluation strategy as a between-subjects variable, and type of bias assessment as a within-subjects variable. The interaction was significant, $F(3, 89) = 13.02, p < .0001$ (see Figure 1.1). To further explore this interaction effect, we next analyzed participants’ assessments of bias in their strategy separately from their assessments of bias in their own evaluations.
We predicted that participants in the explicitly biased conditions would see their assigned strategies as biased compared to participants in the other conditions. An omnibus test of all four conditions was significant, \( F(3, 90) = 15.76, p < .0001 \). As predicted, participants in the explicitly biased weakness-focus condition rated their assigned strategy as less objective \((M = 4.28, SD = 1.62)\) than did participants in the unrestricted condition \((M = 6.50, SD = 1.60)\), \( p < .0001 \), and in the explicitly objective condition \((M = 5.78, SD = 1.48)\), \( p = .002 \). Likewise, those in the explicitly biased strength-focus condition \((M = 3.63, SD = 1.72)\) saw their strategy as less objective than did participants in the unrestricted condition, \( p < .0001 \), and in the explicitly objective condition, \( p < .0001 \).

While participants in the explicitly biased conditions perceived greater bias in their assigned strategies than did their peers, we predicted that they would rate their test evaluations as no more biased. As predicted, an omnibus test of the effect of participants’ experimental condition on their assessments of the objectivity of their test evaluations revealed no differences, \( F(3, 89) = 1.21, p = .31 \), with participants in all conditions rating their choices as similarly objective \((M_{\text{explicitly-biased, weakness-focus}} = 4.76, SD = 1.99, M_{\text{explicitly-biased, strength-focus}} = 5.75, SD = 1.85, M_{\text{explicitly-objective}} = 4.91, SD = 2.02, M_{\text{unrestricted}} = 5.14, SD = 1.94)\), \( ps > .08 \).
Figure 1.1. Participants’ ratings of bias in their judgmental strategy and in their subsequent self-judgment. Error bars indicate 1 SE above and below the mean.

Bias in explicitly-biased conditions. The present study concerns individuals’ tendency to deny bias in their judgments even when they have acknowledged that the processes leading up to those judgments are biased. Although such a tendency to deny bias seems consistent with a bias blind spot, one could argue that such denials were accurate—if participants were able to be objective even though their judgmental processes were biased. In order to address this possibility, we next examined whether participants in the explicitly-biased conditions had been biased.

As predicted, participants in our explicitly-biased conditions showed a “self-serving” bias whereby they saw the test as less valid after allegedly performing poorly on it rather than well. Those who failed the test and then were instructed to list test flaws rated the test as less valid than did those who succeeded on the test and then were instructed to list test strengths ($M = 4.72$,
\[ SD = 1.24, \text{ vs. } M = 5.58, \ SD = 1.53), \ t(47) = 2.17, \ p = .04 \text{ (See Figure 1.2).} \] It also may be worth noting that participants in the three failure conditions (those who were instructed to be biased, to be objective, or to use their own discretion) all evaluated the test similarly to each other, all ts < 1.05, n.s., and more negatively than those in the one success condition, all ts > 2.17, ps < .04. It is not surprising that participants in the explicitly objective condition were as negative about the test as their peers. As noted earlier, they did not fully comply with the objectivity instruction and instead listed more weaknesses than strengths.

**Figure 1.2.** Mean ratings of test quality by explicitly biased (positive feedback v. negative feedback) condition. Error bars indicate 1 SE above and below the mean.

In this experiment, participants in the explicitly biased conditions rated their judgmental strategies as biased, made biased judgments using those strategies, and then claimed that their
judgments were objective. In this experiment, we were able to ascertain the presence of bias among participants in the explicitly biased conditions via a comparative analysis—that is, participants’ ratings of the test were more negative when they had been told they did poorly rather than well.

**Study 1.2**

This study aims to augment the results of Study 1.1 by using an external criterion to assess bias rather than doing so comparatively. It also introduces a new domain of bias commission in order to test our hypotheses outside of the realm of self-serving biases and test performance.

**Method**

**Participants.** Seventy-four Princeton undergraduates (36% male; 55% Caucasian), all of whom reported no knowledge of art history beyond “visiting museums,” participated for course credit. The median age was 19 ($SD = 1.23$).

**Materials and Procedure.** Participants arrived at the laboratory and rated a series of 80 paintings in terms of their “artistic merit” (See Appendix B). They were told that “by artistic merit, we mean the quality of the painting.” Participants were then told that they would be asked to make those ratings after either choosing to see, or choosing not to see, the name of the artist who made the painting. In the explicitly biased condition, participants then were asked to always choose to look at the name of the artist prior to judging the paintings\(^1\). In the explicitly objective condition, participants instead were asked to always choose not to see the name of the artist.

Before seeing the paintings, participants rated the objectivity of their assigned strategy from 1 (completely biased) – 9 (completely objective). Then, for each painting, they were asked to indicate via button press whether or not they would like to see the name of the artist prior to
seeing the painting. After making their choice, and therefore seeing the artist’s name or not, they saw the painting for five seconds. Then, they evaluated its artistic merit from 1 (*no artistic merit*) – 9 (*very high artistic merit*). This process continued for all 80 paintings. Finally, participants assessed how biased versus objective they had been in evaluating the paintings from 1 (*completely biased*) – 9 (*completely objective*).

The 80 paintings were retrieved from the websites of internationally-recognized art museums. Each painting was selected from a larger set (of 120 paintings) that were rated for their perceived quality by a separate sample of undergraduates (*N* = 17) who were untrained in art history and uninformed of the painters’ identities. These ratings were used to produce two sets of 40 paintings of comparable perceived quality, so that one set could be imputed to famous artists, and the other to non-famous artists. For the paintings allegedly made by non-famous artists, we consulted a phone book to obtain names to assign to those paintings. For the paintings allegedly made by famous artists, we used the true artist’s name except in cases where that name was not recognized by a separate sample of undergraduates (*N* = 143) untrained in art history. In those cases, we substituted a more recognizable artist from the same era and style (e.g., Picasso substituted for Braque).

**Results and Discussion**

**Manipulation check: Adherence to strategy.** As a manipulation check on whether participants used the evaluative strategy to which they were assigned, analyses were conducted on participants’ decisions to look versus not look at the artists’ names prior to rating each painting. Participants complied with their assigned conditions: In the explicitly biased condition, they chose to look at the name of the artist 96% of the time, and in the explicitly objective condition, they chose NOT to look at the name of the artist 99% of the time.
**Assessments of bias.** As in Study 1.1, participants rated the objectivity/bias of their assigned evaluative strategy prior to engaging in it. We predicted that participants in the explicitly biased condition would rate the strategy they used as significantly less objective than participants in the explicitly objective condition, but that they would nonetheless rate their resulting evaluations as no less objective. This predicted 2x2 interaction effect (Evaluation Strategy: *Explicitly Biased, Explicitly Objective* x Bias Assessment: *Self-assessment, Assessment of Strategy*) emerged, $F(1, 72) = 19.13, p < .0001$ (See Figure 1.3).

We further explored this interaction effect by looking separately at participants’ assessments of bias in their strategy versus in themselves. A simple effects test of the effect of experimental condition on participants’ assessment of objectivity/bias in their evaluation strategy was significant, $F(1, 72) = 34.21, p < .0001$. That is, participants in the explicitly biased condition viewed their strategy as less objective than participants in the explicitly objective condition rated *their* strategy ($M = 4.00, SD = 2.12$, vs. $M = 6.78, SD = 1.96$). Moreover, and also as predicted, participants in the explicitly biased condition did not view their own evaluations as any less objective than did participants in the explicitly objective condition ($M = 5.89, SD = 1.67$, vs. $M = 5.97, SD = 2.04$), $F(1, 72) = 0.03, ns$. 
Figure 1.3. Participants’ ratings of bias in their judgmental strategy and in their subsequent self-judgment. Error bars indicate 1 SE above and below the mean.

**Bias in explicitly biased condition.** Participants in the explicitly biased condition acknowledged the bias inherent in their assigned evaluative strategy yet failed to acknowledge the presence of that bias in their own evaluations. Did they *actually* show bias in their evaluations? In order to answer that question, we examined whether participants in the explicitly biased condition saw greater “artistic merit” in the paintings attributed to famous artists than the paintings attributed to non-famous artists. Participants in the explicitly biased condition rated the artistic merit of paintings attributed to famous artists ($M = 6.19, SD = 0.87$) significantly higher than paintings attributed to non-famous artists ($M = 5.91, SD = 0.83$), $t(37) = 4.12, p < .001$. Participants in the explicitly objective condition (who did not see the alleged names of the artists) rated the artistic merit of the two groups of paintings the same ($M = 5.82, SD = 0.96$, vs.
$M = 5.80, SD = 1.05), t(35) = 0.36, \ ns$. The relevant interaction was significant, $F(1, 72) = 10.10, p < .01$ (See Figure 1.4).

![Figure 1.4](image.png)

**Figure 1.4.** Mean ratings of artistic merit for paintings attributed to famous and non-famous painters by condition. Error bars indicate 1 SE above and below the mean.

In this experiment, participants who were instructed to use a biased strategy (for evaluating the artistic merit of paintings) rated that strategy as biased, then engaged in it and showed the relevant bias, and, finally, denied having been biased. This experiment extends the results of the previous study by using an external method for assessing bias (i.e., a standard not reliant on between-condition comparisons) and by exploring our hypotheses in the context of a non-motivational bias.
Study 1.3

Why might our participants repeatedly fail to see bias in their decisions despite acknowledging the bias in their decision-making strategy? We have suggested that the answer lies in the fact that bias operates unconsciously and yet, people look to internal conscious cues to assess their own bias. Study 1.3 begins to explore the question of mechanism by investigating a possible alternative. That is, upon recognizing the bias in their explicitly-biased decision-making strategy, participants might try to correct for that bias and, on the basis of that effort, they might conclude that they had been objective.

Method

Participants. Forty-three Princeton undergraduates (56% male; 51% Caucasian) with limited art-history knowledge participated for course credit. The median age was 19 ($SD = 1.21$).

Materials and Procedure. This study used the same paradigm as Study 1.2, in which participants evaluated paintings for “artistic merit” (See Appendix B). As in Study 1.2, some participants were asked to evaluate the paintings using an explicitly biased strategy (choosing to look at the name of the artist prior to judging the painting), and other participants were asked to use an explicitly objective strategy (choosing not to look at the name of the artist). Another group of participants was in a new, implicitly biased condition. They were not asked to choose any strategy for evaluating the paintings and were, instead, automatically shown the names of the painters prior to judging each painting. Participants in this condition were asked no questions about objectivity prior to evaluating the paintings and their attention was not drawn to the possibility of bias; instead, they simply rated the paintings, seeing the name of each artist prior to each painting. Given that these participants’ attention was not drawn to the possibility that a bias
might be influencing their assessments, they should have had little reason to try to correct for that potentially biasing influence.

Participants in the explicit conditions rated the objectivity of their assigned strategies [1 (completely biased) – 9 (completely objective)] prior to evaluating the paintings. They then rated the paintings, after choosing whether or not to see the name of each painter. Participants in the implicitly biased condition were not asked to evaluate a strategy or to even use one – they immediately began the rating task with the names of the artists shown by default. After rating the paintings, participants in all three conditions assessed their own bias or objectivity in providing those ratings on a scale of 1 (completely biased) – 9 (completely objective).

Next, participants in all conditions rated the degree to which they tried to rid their minds of the “contaminating effects of bias” during the rating task on a scale ranging from 1 (not at all) – 9 (very much). Finally, all participants rated each artist (alleged and famous) in terms of how much emotional intensity he/she felt toward the artist from 1 (excited in a negative way) – 5 (neutral/no reaction) – 9 (excited in a positive way).

Results and Discussion

Below, we first present analyses concerning the explicitly biased and explicitly objective conditions, in order to test whether these replicate our findings from Study 1.2. Next, we examine our implicitly biased condition in order to examine additional mechanism questions regarding the possibility of bias correction.

The basic effect (explicitly biased vs. explicitly objective conditions). Participants adhered to their assigned evaluation strategies regarding whether or not to look at the artists’ names prior to rating the paintings. In the explicitly biased condition, they chose to look at the
name of the artist 99.9% of the time, and in the explicitly objective condition, they chose NOT to
look at the name of the artist 100% of the time.

As in Study 1.2, we expected that participants in the explicitly biased condition would
rate their strategy as significantly less objective than participants in the explicitly objective
condition would rate theirs, despite claiming equal objectivity in their actual decisions. This 2x2
interaction effect (Evaluation Strategy: Explicitly Biased, Explicitly Objective x Bias
Assessment: Self-assessment, Assessment of Strategy) was significant, $F(1, 28) = 6.15, p < .05$.
Participants in the explicitly biased condition viewed their strategy as significantly less objective
($M = 3.80, SD = 2.27$) than participants in the explicitly objective condition viewed their strategy
($M = 7.40, SD = 1.84$), $F(1,28) = 22.68$, $p < 0.001$. Once again, though, participants’ ratings of
their own evaluations did not reflect this discrepancy. Participants in the explicitly biased
condition did not see themselves as significantly less objective ($M = 5.53, SD = 2.07$) than those
in the explicitly objective condition ($M = 6.60, SD = 1.60$), $F(1, 28) = 2.51$, $p = .13$.

**Bias correction analyses.** In order to examine the possibility of bias correction (or at
least **efforts** at bias correction) on the part of participants in the explicitly biased condition, we
next examined our participants’ responses in the context of the implicitly biased condition, where
participants were subjected to the same biased decision making process (i.e., seeing the artists’
names prior to rating their paintings) but without the potential for bias being brought to their
attention. Participants in the implicitly biased condition resembled those in the explicitly biased
condition in their denials of personal bias ($Ms = 5.15$ vs. $5.53$, $SDs = 2.08$ and $2.07$), $F(1,26) = 0.23$, ns. But, were participants in the explicitly biased condition actually less biased by the fame
of the artists due to their being made aware of the potential for bias? We predicted that
participants in the explicitly biased condition were **not** correcting, and that their degree of bias
would be comparable to the implicitly biased condition, with paintings attributed to famous artists being rated as higher in artistic merit.

As predicted, there was not a 2x2 interaction effect for the biased conditions (Evaluative Strategy: Explicitly Biased, Implicitly Biased x Painting Fame: Famous Paintings, Non-famous Paintings), $F(1, 26) = 0.55, p = .47$ (See Figure 1.5). While there was an overall effect of painter fame, $F(1, 26) = 11.49, p < .01$, there was no effect of implicit vs. explicit bias condition, $F(1, 26) = 0.55, p = .47$. Participants in the implicitly biased condition rated paintings attributed to famous artists as having significantly higher artistic merit than those attributed to non-famous artists ($M = 6.04, SD = 0.78$ vs. $M = 5.82, SD = 0.78$), $t(12) = 2.45, p < .05$, and participants in the explicitly biased condition also did that ($M = 6.27, SD = 0.50$ vs. $M = 5.93, SD = 0.59$), $t(14) = 2.57, p < .05$, suggesting that participants in the latter condition were not correcting for the bias even though they recognized its presence in their evaluative strategy. Participants in the explicitly objective condition rated the artistic merit of the two groups of paintings the same ($M = 6.02, SD = 0.96$ vs. $M = 6.05, SD = 0.67$), $t(14) = -0.45, p = .66$. 
Figure 1.5. Mean ratings of artistic merit for paintings attributed to famous and non-famous painters by bias condition. Error bars indicate 1 SE above and below the mean.

Additional tests of correction. While the ratings of the paintings indicate that participants in the explicitly biased condition were not, upon recognizing the bias in their assigned evaluative strategy, engaging in correction during the painting-rating task, we looked at the possibility of correction in two other ways as well. First, we looked at participants’ self-reported attempts to rid their minds of the contaminating effects of bias during the painting rating task. Greater self-reported attempts at ridding the mind of bias would be supportive of attempted correction as a mechanism. We, however, predicted that participants across conditions would not vary in their experience of ridding their minds of bias during the task (since we hypothesized that the bias was operating unconsciously and therefore not prompting participants to try to overcome it). As predicted, an omnibus test of the effect of participants’ experimental condition on their
self-reported attempts to rid their minds of the contaminating effects of bias revealed no significant differences, \( F(2, 40) = 1.72, p = .19 \), with participants in all conditions reporting similar attempts (\( M_{\text{explicitly-biased}} = 5.20, SD = 2.15, M_{\text{implicitly-biased}} = 6.00, SD = 1.96, M_{\text{explicitly-objective}} = 4.47, SD = 2.39 \)).

The final way we looked at the possibility of correction as a mechanism was to look at the relationship between participants’ ratings of the paintings and their attitudes toward the painters. Since participants in the biased conditions were biased by the fame of the painters, their attitudes toward the painters should be predictive of their ratings of the paintings’ artistic merit. If participants in the explicitly biased condition were engaging in correction, their correction efforts should disrupt this relationship. That is, their attitudes toward the painters should be less predictive of their ratings of the paintings compared to participants in the implicitly biased condition. The possibility of correction here was investigated using multilevel modeling, allowing us to look at the effect of condition on the relationship between attitudes toward painters and ratings of paintings. We conducted a random intercept model with painter ratings as the Level 1 predictor, condition (explicitly biased or implicitly biased) as a Level 2 predictor, the cross-level interaction between painter ratings and condition as a Level 2 predictor, and painting ratings as the dependent variable. The ICC of painter ratings is 0.14. This model is shown in the equations below.

Level 1: \( y (\text{rating of paintings}) = b_0 + b_1 \times \text{painter rating} + e \)

Level 2: \( b_0 = c_0 + c_{01} \times \text{condition} + u_0 \)

\[ b_1 = c_{10} + c_{11} \times \text{condition} \]

Combined:

\[ y = c_0 + c_{01} \times \text{condition} + c_{10} \times \text{painter rating} + c_{11} \times \text{condition} \times \text{painter rating} + u_0 + e \]
As predicted, the model indicated a significant effect of painter rating on painting ratings ($p < .001$), but no effect of condition ($p = .90$), nor a cross-level interaction between painter rating and condition ($p = .49$), see Table 1.1. This lack of difference between the two biased conditions is further indication that participants in the explicitly biased condition were not engaging in correction when rating the paintings.

**Table 1.1. Effect of painter ratings and condition on painting ratings.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Painter Ratings</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Est.</td>
</tr>
<tr>
<td>Painter Rating</td>
<td>0.33*</td>
<td>0.04</td>
</tr>
<tr>
<td>Condition</td>
<td>0.05</td>
<td>0.41</td>
</tr>
<tr>
<td>Painter Rating*Condition</td>
<td>-0.04</td>
<td>.06</td>
</tr>
</tbody>
</table>

* $p < .001$

In this experiment, participants who were asked to evaluate paintings using a biased strategy rated that strategy as biased, engaged in it, showed the relevant bias, and yet denied having been biased. This experiment added a comparison implicitly biased condition, in which participants' attention was not drawn to the possibility of bias, thus making efforts at bias correction unlikely. Bias correction (or attempted correction) as an alternative mechanism was investigated in three different ways, none of which yielded evidence that correction or correction efforts were especially pronounced in the explicitly biased condition.
Study 1.4

This study investigates our proposed mechanism: participants do not recognize the bias present in their decisions because the bias in their decision-making strategy fails to leave conscious cues. This study gives participants a conscious, emotional cue to the possibility of bias in their judgments in order to test the hypothesis that this will diminish their tendency to deny their bias.

Method

Participants. Twenty-eight Princeton undergraduates (54% male, 46% Caucasian) with limited art-history knowledge participated for course credit. The median age was 20 (SD = 1.14).

Materials and Procedure. This experiment used the same basic paradigm as Studies 1.2 and 1.3, but with an important twist. Again, participants were asked to evaluate a series of paintings for “artistic merit”. In this experiment, however, the non-famous painters were swapped out for alleged convicted murderers.

Participants were told that they would see some paintings by famous artists, and some by convicted murderers who took part in an art-based rehabilitation program. Half of participants were asked to evaluate the paintings with an explicitly biased strategy (choosing to look at the identity of the artist prior to judging each painting), and the other half were asked to evaluate the paintings with an explicitly objective strategy (choosing not to look at the identity). As in prior experiments, participants rated the objectivity of their assigned strategy for evaluating the paintings, and then evaluated each one. Finally, they rated the level of bias or objectivity in their ratings.

The same 80 paintings from Studies 1.2 and 1.3 were used. While the famous painter names remained the same, the non-famous painters’ names were replaced with alleged criminal
inmate numbers and crimes (e.g., #79923-974, Stabbed Victim). The crimes were chosen to reflect the prevalence of various forms of murder in the United States. Names of the famous artists were accompanied by a portrait of the artist, and inmate numbers were accompanied by a mugshot (all mugshots were of Caucasian males; See Appendix B). Mugshots and descriptions of the crimes were included with the intention of eliciting an emotional reaction in the participants (a visceral feeling of dislike) that would serve as a conscious cue to the bias affecting their evaluations of the paintings.

**Results and Discussion**

**Manipulation check: Adherence to strategy.** Participants complied with their assigned conditions regarding whether or not to look at the identities of the artists: In the explicitly biased condition, they chose to look at the artists’ identities 99.7% of the time, and in the explicitly objective condition, they chose NOT to look 94.3% of the time.

**Assessments of bias.** As in the previous experiments, it was expected that participants in the explicitly biased condition (who were asked to look at the identities of the painters) would rate their strategy as less objective than participants in the explicitly objective condition (who were asked not to look). Unlike in those experiments, however, it was predicted that participants in the explicitly biased condition would also rate their own evaluations as less objective than would those in the explicitly objective condition (since we expected that participants’ reactions to the artists’ violent crimes and mugshots would act as a cue to bias). The results revealed a 2x2 interaction effect (Evaluation Strategy: *Explicitly Biased, Explicitly Objective* x Bias Assessment: *Self-assessment, Assessment of Strategy*), $F(1, 26) = 30.32, p < .001$, but one that differed importantly from Studies 1.2 and 1.3 (see Figure 1.6). Participants in the explicitly biased condition viewed their strategy as far less objective than participants in the explicitly
objective condition ($M = 2.93, SD = 1.64$, vs. $M = 8.50, SD = 0.65$), $F(1, 26) = 139.74, p < .001$.

As predicted, participants in the explicitly biased condition also viewed their own judgments as less objective than did participants in the explicitly objective condition, although this difference was less pronounced ($M = 6.07, SD = 1.94$ vs. $M = 8.21, SD = 0.80$), $F(1, 26) = 32.14, p < .01$.

![Figure 1.6. Participants’ ratings of bias in their judgmental strategy and in their own evaluations.](image)

Error bars indicate 1 SE above and below the mean.

**Presence of bias.** Participants in the explicitly biased condition rated both their strategy and their own evaluations as being significantly less objective than participants in the explicitly objective condition rated their strategy and evaluations. This begs the question: Were participants in the explicitly biased condition actually influenced by the identity of the artists? The predicted interaction effect emerged, $F(1, 26) = 4.36, p < .05$. In the explicitly biased condition, participants rated the artistic merit of paintings attributed to famous artists ($M = 6.89, SD = 0.64$) higher than that of paintings attributed to murderers ($M = 6.49, SD = 0.82$),
Participants in the explicitly objective condition (who did not see the artists’ identities) rated the artistic merit of the two groups of paintings similarly ($M = 5.73$, $SD = 0.71$, vs. $M = 5.70$, $SD = 0.83$, respectively), $t(13) = 0.47$, $p = .65$.

In this experiment, participants who were asked to use an explicitly biased strategy when evaluating paintings rated their decision-making strategy as biased, were biased by it, and then proceeded to rate their own evaluations as biased (relative to ratings of participants in the explicitly objective condition). Changing the non-famous artists of Studies 1.2 and 1.3 to murderous criminals led to participants perceiving bias in their own judgments where they previously had not. This suggests that their emotional reaction to seeing murderers acted as a conscious cue to bias.

**Discussion**

In the context of judgments and decisions in the real world of work and relationships, the costs of bias can be high. As a consequence, the ability to recognize one’s biases—and, thereby, to work towards preventing, correcting, and overcoming them—is valuable. When individuals’ judgments are influenced by strategies that they do not see as biased, it seems inevitable that they will fail to see the bias in their judgments. This research suggests that the problem is deeper and more persistent than that. Even when individuals knowingly use biased judgmental strategies, they still may insist on their objectivity.

In this series of experiments, we found that when individuals were induced to use blatantly biased judgmental strategies, they acknowledged the bias in those strategies but nonetheless claimed that their decisions emerging from those strategies were objective. Moreover, we found that this claim of objectivity was false—i.e., that those participants were in fact biased. This research suggests the persistence of people’s blindness to their own bias.
Recognizing one’s bias is a critical first step in trying to correct for it; these experiments make clear how difficult that first step can be to reach.

Two experiments explored the mechanism underlying this phenomenon. We proposed that the reason for this failure in self-perception of bias was due to the unconscious nature of bias. People, we suggest, introspect to look for signs of bias, see none (because it affects them unconsciously), and therefore sincerely conclude that they have been objective. Study 1.3 addressed an alternative mechanism: that participants were trying to correct for the bias and concluding from their correction efforts that they had been objective. After looking for this mechanism in three ways, no signs of correction were found. Finally, Study 1.4 more directly investigated our proposed mechanism. When participants were given a conscious cue to the possibility of their bias, the tendency to see their decisions as objective diminished.

The effects in this chapter were shown in two domains of bias commission. Study 1.1 showed the effects of knowingly engaging in a biased decision-making strategy in the context of a classic self-serving bias—i.e., the tendency to criticize a test after performing poorly on it. Studies 1.2-1.4 showed the effects in the context of a bias involving the tendency for people to view paintings as having more inherent artistic merit when those paintings are attributed to famous artists.

Those participants in our experiments who were induced to use biased decision-making strategies readily saw those strategies as biased but nonetheless claimed personal objectivity. Could it be that their denials of personal bias partly reflected a desire to appear positively to the experimenter? Being biased is generally viewed as negative, and our participants may have wanted to distance themselves from this negative characteristic. A couple of features of our studies suggest that social desirability concerns are not the principle cause of our effects. First, it
is unclear that bias denials would be viewed as socially desirable in the explicitly biased conditions. Rather, it might appear hypocritical for individuals to call their judgmental strategy biased and then, minutes later, claim objectivity in their resulting judgment. Given the social undesirability of appearing hypocritical, participants may have had to put social desirability concerns aside in order to disclose their conviction that they had been objective. Second, participants in Studies 1.2-1.4 knew that the experimenter would be able to discern whether their assessments had been biased by the identities of the painters. Therefore, concerns about social desirability should have magnified their interest in responding as accurately as possible—in order to avoid the embarrassment of claiming objectivity when the experimenter could readily see their bias.

While we cannot rule out the possibility that social desirability played a role in our effects, another explanation seems more compelling to us. We have suggested that participants’ persistent denials of bias in this research reflect the nature of bias commissions: Although people may have access to the products of the judgmental strategies they use, and even to those strategies themselves, they lack access to the underlying mental processes that produce bias in their judgments. As a consequence, the presence of bias is hidden from consciousness. This notion is supported by past research revealing the unconscious nature of bias commissions (e.g., Ditto & Lopez, 1992; Gilbert et al., 1998; Williams & Gilovich, 2008; Wilson & Brekke, 1994) and is additionally supported by the results of Study 1.4. The present experiments suggest the depth of the problem of unconscious bias. Even when individuals become aware of bias in the explicit processes they use to make a decision, they are still unaware of the implicit processes that produce their bias and thus their judgments still “feel” unbiased.
The consequences of this persistent bias blindness can be serious. Consider a group of jurors who have just been exposed to testimony that they are now told to disregard as inadmissible. Each juror might admit a priori that being exposed to inadmissible evidence biases judgment. But, having been exposed to that information, each is likely to believe that his or her own judgment is objective (even though that judgment likely took into account the “inadmissible” evidence; e.g., Sommers & Kassin, 2001). Or, consider a team of human resources officials who have a hiring process in which they see photographs of applicants before evaluating the merits of their applications. Each official might view that process as introducing a host of potential biases, but each might believe that they personally can overcome that bias (and, thereby, benefit from being able to “attach a face” to an application). This blindness may be especially likely to foster conflict between people who disagree after going through the same judgmental process. Each may not only be confident in their own objectivity but also quick to impute bias to the person who disagrees—and who was subject to a biased decision-making process.
Chapter 2

Decreasing Bias through Instructions to Appear Objective

“A judge shall avoid impropriety and the appearance of impropriety in all of the judge’s activities,” states the second canon of the U.S. judicial code of conduct. Interestingly, this rule not only limits a judge’s actions themselves, but also the appearance of his or her actions. This emphasis on appearances seems to imply that for judges, the pillars of objectivity in our society, avoiding the appearance of impropriety is a key step in the actual avoidance of impropriety; that the perspective of an outside observer should be taken into consideration prior to taking action. This suggests that there is agreement in society as to what behavior appears proper or objective and what does not – and the latter is to be avoided (even if the actor does not personally feel that impropriety has taken place). This chapter, like the US judicial code, stresses the appearance of objectivity, instructing people to appear objective in their decision-making. It does not seek to encourage people to simply hide their biases under a veil of objectivity like the above might suggest, but rather aims to take advantage of this knowledge of what objectivity looks like in order to lead people to make more objective decisions.

Increasing the objectivity of people’s decisions is no easy feat. While people are quick to point out bias in those around them, they are not inclined to acknowledge those same biases in themselves (Pronin, Lin, & Ross, 2002). This can be explained by people’s tendency to weight introspective information more strongly than behavioral information when evaluating themselves, but not others (Pronin & Kugler, 2007).

Chapter 1 of this dissertation explored the stubborn nature of this bias blind spot; even when individuals acknowledge that they are doing something in a biased manner, they are still inclined to see themselves (and the decisions that they make) as objective. It is perhaps
unsurprising then that research suggests that de-biasing people is fraught with difficulty (e.g., Wilson & Brekke, 1994; Wegener, Clark, & Petty).

This chapter introduces a manipulation aimed at increasing the objectivity of decisions despite the bias blind spot – a manipulation that can be implemented across a variety of domains. This manipulation takes advantage of the self-other asymmetry in bias attribution by having participants consider the perspective of an observer when making their own decision: that is, asking them to make a decision that would appear objective to another.

This chapter first aims to establish the efficacy of asking people to appear objective over asking them to be objective or leaving them to make their own decision (a control condition) on increasing the objectivity of their decisions. Study 2.1a begins by establishing the effectiveness of appearing objective at reducing sexism within the technology industry. Study 2.1b suggests a boundary condition of this manipulation by reducing the specificity of the objectivity instruction and examining its effectiveness. In Study 2.1c, the sexism paradigm is altered so that the ratio of men to women more closely reflects a real workplace environment in the technology industry.

In Study 2.2, the appear-objective manipulation is brought into a new bias domain: the better-than-average effect. This study looks at the effectiveness of this manipulation on this new bias as well as examining how it affects participants’ confidence in their decisions. One would expect that being asked to appear objective (i.e. making a decision different from what you truly believe) would reduce participants’ confidence in the accuracy of their decisions. Ironically, being asked to be objective may increase participants’ confidence in their accuracy, a feeling that will not be mirrored in the actually objectivity of their decisions.

Studies 2.1c and 2.2 examine differences in the degree of effort participants put into their decision-making. This is to differentiate the present research from work on accountability,
which aims to positively impact decision-making by increasing the amount of cognitive effort people put into their decisions (Lerner & Tetlock, 1999).

A new bias domain is explored in Study 2.3 – bias in grade prediction. This study looks at the effect of instructions to appear objective on the accuracy of students’ grade and rank predictions for an upcoming exam. This study again looked at the possibility of differences in participants’ confidence in the accuracy of their decisions across conditions.

Study 2.4 looks at political partisan bias, asking participants to rate quotes said by President Barack Obama and Mitt Romney prior to the 2012 Presidential election. This study examines a downstream effect of the appear-objective manipulation, suggesting that the manipulation positively impacts interpersonal behavior as well as decision-making. The effect of appearing objective on participants’ private beliefs was explored to see whether participants were displaying reactance to the manipulation.

The final pair of studies examine whether instructions to appear objective increase participants’ willingness to blind themselves to potentially biasing information. Since not all decisions have a clearly objective response, these studies suggest an additional way in which the appear-objective manipulation may be useful.

Together, these studies explore the efficacy of instructing people to appear objective across a wide variety of bias domains. In addition to looking at its effect on decisions themselves, these study examine the manipulation’s limitations, its downstream consequences, and its alternative uses.

**Study 2.1a**

Study 2.1a is the first attempt to demonstrate the effectiveness of appearing objective on people’s decisions. In particular, this study looks at the effectiveness of appearing objective at
reducing sexism. Sexism remains a problem for women in the workforce. For example, a study situated in the academic workplace found that when evaluating the qualifications of male and female assistant professors, reviewers were four times more likely to ask for supporting evidence for the female professors, such as a chance to see her teach, than they were for the male professors (National Academy of Sciences, 2006). The present study concentrates on sexism within the domain of the technology industry – a male-dominated field. Despite increases in gender equality, women remain underrepresented in the fields of science, technology, engineering, and mathematics (STEM; Snyder, Dillow, & Hoffman, 2009).

In the present study, participants were asked to assemble a committee from a selection of male and female employees at a large technology corporation. I hypothesized that asking participants to appear objective would lead to greater gender equality in their decisions compared to both a control condition and a condition where participants were asked to be objective.

**Method**

**Participants.** Participants consisted of 199 Amazon Mechanical Turk workers (61% male; 77% Caucasian). The median age was 27 ($SD = 10.51$). Participants were compensated at prevailing rates.

**Materials and Procedure.** Participants were asked to assemble a committee of 10 employees from a selection of 16 at a large technology corporation. The purpose of the committee was ostensibly to increase the quality of data security at the corporation. Participants were instructed to select the 10 employees that they felt were most qualified for the committee.

Participants were presented with 16 employee biographies with photographs (See Appendix C). The biographies were adapted from the actual biographies of employees at a major technology corporation (all of whom were male) and edited for length. The photos they
were paired with were all professional head shots of Caucasian adults taken with neutral backgrounds.

A separate group of raters (N = 26, 14 male) rated the photos for age and biographies for quality. Cross-gender employee pairs were created to control for effects of age and biography quality. The cross-gender pairs were created by selecting photos that matched on age and biographies that matched on quality. The photo pairs and biography pairs were then randomly matched together. The gender of the photo appearing with each biography in a pair was counterbalanced across participants.

Analyses indicate that the male and female employees did not significantly differ in perceived age or quality. The group of raters rated the photos on a scale of 1-10 for apparent age: 1 (21-25 years old), 2 (26-30), 3 (31-35), 4 (36-40), 5 (41-45), 6 (46-50), 7 (51-55), 8 (56-60), 9 (61-65), 10 (66-70). There was not a significant difference in average age between the female photos (M = 5.24, SD = 1.44) and the male photos (M = 5.49, SD = 1.48), F(1, 14) = 0.12, ns. Similarly, there was not a significant difference in quality (1 Not at all qualified – 9 Very qualified) between the two groups of biographies [Group 1 = 6.25 (1.04); Group 2 = 6.26 (0.88)], F(1, 14) = 0.00, ns. Participants did not see any names for the employees nor know what corporation the employees purportedly worked for.

There were three conditions in this study. First, there was a control condition in which participants were given no instructions beyond picking the most qualified employees for the committee and were given no instructions pertaining to objectivity. In the second condition, participants were asked to be objective. They were given special instructions such that they were told: “As you make your decision, please do not be biased by the gender of the employee.” Lastly, participants in the key condition were asked to appear objective. That is, they were told:
“As you make your decision, please do not appear to be biased by the gender of the employee. Think of how your responses would look to an outside observer, and respond in a way that would make that person believe you had not been biased by employee gender.”

I hypothesized that participants in the appear-objective condition would show the most gender equality in their decisions, while those in the other two conditions would be biased by gender. In other words, I expected that those told to appear objective would overwhelmingly be balanced in their selection of the committee (choosing equal numbers of men and women) while participants in the control and be-objective conditions would be biased in favor of male employees.

**Results and Discussion**

**Committee gender ratio.** First, the gender make-up of participants’ selected committees was analyzed for gender bias favoring male employees. The control, be-objective, and appear-objective conditions did differ in the gender ratios that participants chose for the data security committee. Participants’ choices were coded as whether their committee favored men, was balanced (contained an equal number of men and women), or favored women. A chi square analysis indicates a significantly different pattern in gender ratio across the three conditions, \( \chi^2(4) = 18.35, p < .01 \) (see Figure 2.1). More participants in the control condition favored men (44.1%) than were balanced (36.8%) or favored women (19.1%). The be-objective condition was spread similarly across favoring women (31.7%), being balanced (39.7%) and favoring men (28.6%), and did not differ from the control, \( \chi^2(2) = 4.3, p = .12 \). While participants asked to be objective did not favor men like the control condition, only about one third of those participants were balanced in their committee selection. It appears that when asked to be objective, a number
of participants over-corrected and favored women over being balanced in their committee selection, though this pattern was not significantly different than the control.

Finally, as predicted, participants in the appear-objective condition overwhelmingly were balanced in their selection (61.8%) rather than favoring women (7.4%) or favoring men (30.9%), significantly differing from both the control, $\chi^2(4) = 9.46, p < .01$, and be-objective conditions, $\chi^2(2) = 13.37, p < .01$. Given that I put forth an equal gender ratio as the most objective committee arrangement, this indicates a reduction in sexism.

![Figure 2.1](image.png)

*Figure 2.1.* Percentage of participants in each condition favoring women, being balanced (selecting an equal number of men and women), and favoring men.

**Committee quality.** Are participants instructed to appear objective able to avoid sexism while *simultaneously* picking the most qualified applicants? To answer this question, the average quality of employees selected for participants’ committees was analyzed. Using the quality ratings of the bios without photos [1 (*Not at all qualified*) – 9 (*Very qualified*)], there was
no difference in quality across the three conditions, with the control condition ($M = 6.41, SD = 0.22$), be-objective condition ($M = 6.43, SD = 0.21$), and appear-objective condition ($M = 6.45, SD = 0.21$) exhibiting similar levels of quality, $F(2,196) = 0.45, ns$. This indicates that participants’ reduced sexism in the appear-objective conditions was not at the cost of overall committee quality.

However, there was a significant effect of gender ratio on quality, $F(2, 196) = 6.31, p < .01$, such that across conditions, participants that were balanced chose committees of higher quality ($M = 6.48, SD = 0.19$) than those that favored men ($M = 6.37, SD = 0.19$), $p < .01$, and those that favored women ($M = 6.40, SD = 0.25$), $p = .05$. The committees of individuals that favored men and favored women did not differ, $p = .38$. There was no interaction between condition and committee gender ratio, $F(2, 190) = 0.76, p = .55$. This finding is unsurprising, as the male and female employees were matched on quality in this study – the highest quality committee would therefore have to consist of equal numbers of men and women.

In this study, participants asked to appear objective overwhelmingly chose committees with an even gender ratio, selecting an equal number of male and female employees. This reduced sexism in the appear-objective condition did not come at the cost of the quality of the employees they selected, as committees with an even gender ratio were actually of the highest quality. These participants were able to simultaneously balance the needs of appearing objective and selecting qualified employees. These findings have important implications for overcoming gender inequality in fields in STEM, providing a practical way to reduce sexism in hiring.

**Study 2.1b**

In the previous study, participants were asked not to be or appear biased by the employee’s gender. This study aimed to see whether asking participants to appear objective
would still be effective even when a precise bias was not specified; in this study, participants were asked simply to not be or appear biased (in general). Can participants balance the demands of multiple potential sources of bias (e.g., gender, age, physical attractiveness) while reducing bias in the domain under examination (here: sexism)?

**Method**

**Participants.** One hundred and fifty-two Mechanical Turk workers participated (43% male; 71% Caucasian) in this study. The median age was 29 ($SD = 11.72$). Participants were compensated at prevailing rates.

**Materials and Procedure.**

The stimuli and task remained the same as those in Study 2.1a. Participants were asked to pick a committee from a selection of 16 employees, half of which were women. Male and female employees were matched for apparent age and biography quality then counterbalanced across participants (See Appendix C).

This study diverges from Study 2.1a with changes in manipulation wording. Participants in the control condition were given no additional instructions beyond the basic task that they were to select 10 out of the 16 employees from this large technology corporation for a committee aimed at increasing data security. Instructions in the be-objective condition were altered such that they were told: “As you make your decision, please do not be biased.” Participants in the appear-objective condition were told: “As you make your decision, please do not appear to be biased. Think of how your responses would look to an outside observer, and respond in a way that would make that person believe you had not been biased.”

After engaging in the selection task, participants were asked if they had tried to avoid gender bias in the study. The answer options were 1 (*I tried to avoid being biased against male*...
employees), 2 (I tried to avoid being biased against female employees), and 3 (I did not try to do either of the above). This measure will allow us to see if participants were specifically trying to correct for sexism. Participants were also asked what bias related to gender the experimenters would have wanted them to avoid: 1(People picking more male employees than female employees), 2, (People picking more female employees than male employees).

Results and Discussion

Committee gender ratio. First, the gender make-up of participants’ selected committees was examined for differences in gender bias. Whether participants favored men, favored women, or were balanced (picking an equal number of men and women) was analyzed. There was not a significant effect of condition on gender ratio, $\chi^2(4) = 1.83, p = .77$. The degree to which participants in the control condition favored men (33%), favored women (33%), and were balanced (35%), did not differ from the be-objective condition (26% favored men, 28 % favored women, and 47 % were balanced), nor the appear-objective condition (28% favored men, 34% favored women, and 38% were balanced).

Attempts to avoid sexism. Since this study did not indicate what bias participants should avoid, they were asked to self-report any attempts to avoid sexism. Participants were asked at the end of the survey to indicate if they had attempted to avoid bias against men, against women, or neither. There was not a significant effect of condition on this variable, $\chi^2(4) = 4.13, p = .39$. In fact, self-reports in the control (18% attempted to avoid bias against men, 33% attempted to avoid bias against women, 49% attempted neither), appear-objective (9% attempted to avoid bias against men, 45% attempted to avoid bias against women, 47% attempted neither), and be-objective conditions (22% attempted to avoid bias against men, 34% attempted to avoid bias
against women, 44% attempted neither), all had the highest number of participants reporting that they did not attempt to avoid bias based on gender at all.

Unsurprisingly, there was not an effect of condition on what gender bias participants thought the experimenters would want them to avoid, $\chi^2(2) = 1.27, p = .53$. In fact, the vast majority of participants indicated that the experimenters would want them to avoid picking more male employees than female employees (control = 91%, be-objective = 85%, and appear-objective = 84%), indicating that participants agreed on the direction of potential gender bias in the study.

Unfortunately, in this study, when no longer specifically told not to appear biased by gender, but rather simply not to appear biased, the appear-objective manipulation no longer successfully increased gender equality in participants’ decisions. Despite the majority of participants recognizing that gender bias in this study would be against women (and not men), most participants said they did not attempt to correct for gender bias at all.

With multiple facets of a person being a potential source of bias (e.g., age, gender, race), this study suggests that it is necessary to specify which source of bias participants need to appear to be uninfluenced by. This means that in practice, the goal of this manipulation would have to be to reduce a specific bias. Future research should investigate whether people can balance appearing objective across more than one bias when both biases are specified.

**Study 2.1c**

The previous studies were very controlled in the employees available for participants to pick from: there were the exact same number of men and women and the two genders were matched to biographies of equal quality. Unfortunately, this is not representative of actual gender representation in STEM fields such as the technology industry where women remain
underrepresented (Snyder, Dillow, & Hoffman, 2009). Study 2.1c changes the ratio of men to women available in order to better approximate real-world conditions. This study will attempt to answer the question: Can participants balance the competing needs of appearing objective while still selecting the most qualified candidates for the committee?

In particular, this study will look at the effect of the conditions (appear-objective, be-objective, and a control) on how many ambiguously qualified female employees are selected for the committee. Since an uneven gender ratio makes the selection decision more difficult, it is hypothesized that sexism would most likely emerge in ambiguous cases. Past research on racial discrimination has shown that participants did not discriminate against black candidates when their qualifications were strong or weak, only when they were ambiguous (Dovidio & Gaertner, 2000).

The instructional manipulations for participants asked to be objective and to appear objective were equated for approximate length in this study. This was to rule out the possibility that the longer instructions for the appear-objective condition in previous studies were responsible for some decrease in bias.

Additionally, this study begins to differentiate the current research from research on accountability, which can operate through increases in cognitive effort on part of the decision-maker (Lerner & Tetlock, 1999). Study 2.1c will ask participants to rate how much effort they put into selecting their committee. Despite the complexity of this task given the uneven gender ratio of male and female employees available, I do not believe that the appear-objective manipulation will necessitate increased effort in decision-making. Making a decision that conforms to an outside observer’s criterion for objectivity should be equally, if not less difficult to make than under control conditions.
Method

Participants. Two hundred participants (76% male, 72% Caucasian) participated on Mechanical Turk. The median age was 25 (SD = 9.13). Participants were compensated at prevailing rates.

Materials and Procedure. Like Studies 2.1a and 2.1b, participants in this study were again asked to select employees from a large technology company to be a part of a committee aimed at increasing data security. However, in this study, the number of employees (out of 16) that were female was reduced from eight to six. The two previously female employee photos were exchanged for male photos of a similar age.

Of the six remaining female photos, one was paired with a biography of very high quality (part of the cross-gender pair with the top-ranking biographies; i.e., was definitely qualified to be a part of the committee), four were paired with biographies of moderate quality (from pairs with mid-ranking biographies; i.e., were ambiguously qualified for the committee), and one was paired with a biography of low quality (from the pair with the bottom-ranking biographies; i.e., was not qualified to be a part of the committee). The photos and biographies were matched cross-gender for age and quality and counterbalanced as in previous studies (See Appendix C).

Additionally, the size of the committee was reduced from 10 to 8 in the present study in order to decrease the number of positions available to ambiguously qualified employees. Given the smaller committee and the quality ratings of the remaining female employees, this altered paradigm will allow the gender makeup of a participant’s selected committee to have a greater impact on the committee’s overall quality than in previous studies.

This study consisted of three conditions. Participants were given no additional instructions in the control condition beyond selecting a committee. Participants in the be-
**objective condition** were told: “As you make your decision, please do not be biased by the gender of the employee. Try to respond so that you personally do not believe you have been biased by gender.” Participants in the **appear-objective condition** were told: “As you make your decision, please do not appear to be biased by the gender of the employee. Think of how your responses would look to an outside observer, and respond in a way that would make that person believe you had not been biased by gender.”

Additionally, participants were asked at the end of the survey to indicate how much effort they had put into making their selection on a scale of 1 (*None*) – 9 (*A lot*). Participants were then asked to indicate (via multiple choice) what option matched their instruction prior to the task as a manipulation check.

I hypothesize that in this study, participants instructed to appear objective will display less sexism in selecting their committees. Additionally, given the alterations to the paradigm, this decrease in sexism should be reflected in an increase in the overall quality of their committees.

**Results and Discussion**

**Selection of female candidates.** I first looked for differences in sexism by looking at the total number of female employees selected and the number of ambiguously qualified female employees selected. The total number of female employees selected (out of six) did not differ across the control ($M = 2.60, SD = 1.03$), be-objective ($M = 2.86, SD = 1.06$), and appear-objective ($M = 2.76, SD = 0.96$) conditions, $F(2, 197) = 1.10, p = .34$. They also did not differ in the number of ambiguously qualified female employees (out of four) that they selected, $F(2, 197) = 1.40, p = .25$. 
Given the quality ratings of the biographies, the optimal committee would have to include at least two of the ambiguously qualified female employees. Any less than two would negatively impact the overall quality of the committee and indicate sexism against the female employees. When we compare the number of ambiguously qualified female employees selected compared to this number (two), the control condition \((M = 1.70, SD = 0.89)\) is the only condition selecting significantly fewer than two, \(t(69) = -2.82, p < .01\). Neither the appear-objective \((M = 1.91, SD = 0.94)\), nor the be-objective \((M = 1.95, SD = 0.99)\) condition selected significantly fewer than two ambiguously qualified female employees on average, \(t's < -0.38, p's > .44\). This suggests a reduction in sexism in both the be-objective and appear-objective conditions compared to the control condition.

As one would expect, there was not a difference across conditions in the percentage of participants selecting the highly qualified female employee, \(\chi^2(2) = 1.34, p = .51\) (control = 56%, be-objective = 56%, and appear-objective = 64%). While there was likewise not an overall effect on the percentage of participants selecting the unqualified female employee across conditions, \(\chi^2(2) = 3.98, p = .14\), participants in the appear-objective condition were marginally less likely to pick the unqualified female employee (21%) than both the control condition (34%), \(\chi^2(1) = 3.06, p = .08\), and the be-objective condition (35%), \(\chi^2(1) = 3.19, p = .07\). The control and be-objective conditions did not differ, \(\chi^2(1) = .006, p = .94\).

**Committee Quality.** This tendency for the appear-objective condition to be less likely to pick the unqualified female employee is reflected in a significant difference in the average quality ratings of the committees across conditions, \(F(2, 197) = 3.53, p = .03\) (See Figure 2.2). The committees of participants in the appear-objective condition \((M = 6.52, SD = 0.27)\) were of significantly higher quality than those of participants in the control \((M = 6.40, SD = 0.28)\),
$p < .01$. The quality of committees selected by participants in the be-objective condition ($M = 6.45, SD = 0.28$) did not differ from either the control or the appear-objective condition, $p's > .12$.

![Average committee quality by condition. Error bars indicate 1 SE above and below the mean.](image)

**Figure 2.2.** Average committee quality by condition. Error bars indicate 1 SE above and below the mean.

**Effort.** This study examined participants’ self-reported degree of effort in making their decision in order to compare the appear-objective manipulation to research on accountability (driven, in part, by an increase in effort; Lerner & Tetlock, 1999). There was no effect of condition on self-reported effort in making the committee decision, $F(2, 196) = 0.14, ns$. Participants in the control ($M = 7.27, SD = 1.50$), be-objective ($M = 7.38, SD = 1.22$), and appear-objective ($M = 7.39, SD = 1.55$) conditions all reported comparable degrees of effort. Despite participants in the appear-objective more successfully balancing their competing goals,
they did not self-report that this took greater effort. This provides initial support that the appear-objective manipulation does not operate similarly to accountability manipulations.

While experimental condition did not have an effect directly on the number of ambiguously qualified female employees chosen, the control condition was the only condition that displayed any sexism against women. Participants asked to appear objective selected significantly higher quality committees than did the control condition. This is reflected in their decreased likelihood to select the unqualified female employee than either the control or be-objective conditions. This data suggests that participants asked to appear objective were better at balancing the dual goals of avoiding sexism and selecting the most qualified employees for the committee. Additional research is needed to clarify the degree to which participants asked to appear objective can successfully do so when there is not a clearly objective decision to be made (i.e., having an equal number of men and women to choose from as in Study 1.1a).

**Study 2.2**

This next study takes the appear-objective manipulation into a new bias domain: self-serving bias. In particular, this study looks at the better-than-average effect (BAE) in regard to traits. The better-than-average effect is people’s tendency to see themselves in a self-aggrandizing light, for example, claiming to be friendlier and more intelligent than average, despite the statistical impossibility (Alicke, 1985). I predict that instructing participants to appear objective will decrease the degree to which they display the better-than-average effect.

Additionally, this study will examine the effects of being asked to appear objective, be objective, and a control on participants’ confidence in the accuracy of their ratings. I hypothesize that the appear-objective manipulation will lead to more objective decisions than asking participants to be objective, despite the fact that this will not be reflected in their
confidence in the accuracy of their ratings. I hypothesize that participants asked to be objective will ironically report greater levels of confidence in the accuracy of their self-ratings. This would underscore the importance of instructing people to appear objective rather than to be objective – when asked to be objective, people will not only fail in the pursuit of that goal, but will also be falsely confident in their decisions.

Finally, this study will look at how much effort participants put into making their decisions. No differences were found in Study 1.1c across conditions for how much effort participants put into their decisions – despite participants who were asked to appear objective doing the best job at balancing employee gender and quality. In the current study, the difficulty of appearing objective will be less than in Study 1.1c, and I hypothesize, will actually take less effort than being objective or being left to make one’s own decision (control condition), continuing to distinguish the current work from research on accountability (e.g., Lerner & Tetlock, 1999).

**Method**

**Participants.** One hundred and three Princeton undergraduates (21% male, 49% Caucasian) participated in exchange for entry into a lottery. The median age was 19 (SD = 1.67). Four participants were excluded from analysis for indicating that they had participated in a pilot version of this study. Another four participants were excluded for failing to follow directions (i.e., did not type out the manipulation as instructed).

**Materials and Procedure.** Participants were asked to rate themselves compared to the average Princeton student on a series of traits. The trait ratings used in this study were adapted from Kugler and Pronin (2007) and Williams and Gilovich (2008) to demonstrate the better-than-average effect and consisted of the following: considerate, mature, and able to get along
with others. Trait ratings were on a five-point scale: 1 (Much less than average Princeton student), 2 (Somewhat less than average Princeton student), 3 (Same as average Princeton student), 4 (Somewhat more than average Princeton student), 5 (Much more than average Princeton student).

In the control condition, participants were given no additional instructions. In the be-objective condition, participants were told: “People sometimes have a self-enhancement bias – i.e., they are motivated to see themselves in a positive light. As you make your ratings, please avoid any influence of bias. Try to respond so that you personally do not believe you have been biased in this way.” In the appear-objective condition, participants were told: “People sometimes have a self-enhancement bias – i.e., they are motivated to see themselves in a positive light. As you make your ratings, please avoid any appearance of bias. Try to respond so that an outside observer would not believe you have been biased in this way.” In the two experimental conditions, participants were asked to type out their instructions word-for-word in order to increase retention of the manipulation.

After making the series of trait ratings, participants were asked to indicate how confident they were in the accuracy of their ratings on a scale ranging from 1 (Not at all) – 5 (Extremely). Participants then indicated how much effort they put into making their ratings on a scale ranging from 1 (None) – 5 (A lot). Participants indicated via multiple choice what option matched their instruction at the end of the survey as a manipulation check.

Results and Discussion

Better than average effect. First the effect of condition on trait ratings was explored, followed by an analysis of the better-than-average effect. There was a significant effect of condition on participants’ trait ratings, \( F(2, 92) = 4.93, p < .01 \) (see Figure 2.3). Participants
instructed to appear objective ($M = 3.20, SD = 0.50$) rated themselves significantly lower than control participants ($M = 3.60, SD = 0.47$), $p < .01$, and participants instructed to be objective ($M = 3.52, SD = 0.59$), $p = .02$. The control and be-objective conditions did not differ, $p > .50$.

The average ratings for the control condition were significantly higher than “average” (a scale rating of three), $t(35) = 7.69$, $p < .001$, as were those for the be-objective condition, $t(30) = 4.83$, $p < .001$, indicating that both conditions displayed the better-than-average effect. While participants’ ratings in the appear-objective condition were significantly lower than the other conditions, they were also above three, $t(27) = 2.14$, $p = .04$.

![Average Trait Ratings by Condition](image)

**Figure 2.3.** Average trait ratings by condition. Error bars indicate 1 SE above and below the mean.

**Confidence.** Participants’ self-reported confidence in the accuracy of their ratings was explored next. There was no effect of condition on participants’ confidence in the accuracy of
their ratings, $F(2, 92) = 0.53, p = .59$. While participants in the be-objective condition did self-report the highest level of confidence ($M = 3.55, SD = 0.72$), this was not significantly higher than ratings for the control ($M = 3.36, SD = 0.87$) or appear-objective ($M = 3.46, SD = 0.58$) conditions, $p$’s > .31.

While an ironic effect on participants’ confidence in the accuracy of their ratings was hypothesized (such that participants instructed to be objective would report greater confidence despite displaying greater bias), this was not the case as there was no effect of condition on confidence. Interestingly, asking participants to make a decision that would appear objective did not reduce their confidence. One would expect that being instructed to make a decision that differs from what you really believe would lead to a decrease in the confidence in that decision’s accuracy. This is particularly surprising given that their ratings were significantly lower than participants’ ratings in the other two conditions.

**Effort.** Participants rated how much effort they put into making their trait ratings. There a main effect of experimental condition on effort, $F(2, 92) = 3.22, p = .04$ (see Figure 2.4) Participants in the appear-objective condition self-reported expending less effort ($M = 2.82, SD = 0.72$) when making their ratings than did participants in the control ($M = 3.22, SD = 0.76$) and be-objective conditions ($M = 3.23, SD = 0.62$), $p$’s < .03. The control and be-objective conditions did not differ, *ns.*
This study successfully shows the effectiveness of the appear-objective manipulation at reducing the better-than-average effect in trait ratings. When participants were asked to appear objective, they rated themselves as significantly closer to average than participants asked to be objective or participants in a control. Like Study 2.1c, this study did not find that participants asked to appear objective were putting greater effort into their decision-making. In fact, with this simpler decision, the effect is in the opposite direction. This further suggests that the appear-objective manipulation affects decision-making in a way that is distinctive to that of accountability manipulations which aim to improve decisions by increasing the amount of cognitive effort that people put into those decisions (Lerner & Tetlock, 1999).
Study 2.3

This next study aimed to show the effectiveness of appearing objective in a new bias domain: prediction bias. This study will look at prediction error and the better-than-average effect (BAE). Research shows that people are consistently overoptimistic when predicting their own life events and future success (e.g., Bueller, Griffin, & Ross, 2002; Langer, 1975). This study examined students’ predicted performance on an upcoming midterm exam: both their predicted score (for prediction bias) and predicted rank (for BAE) compared to their peers. I hypothesize that participants asked to appear objective will show significantly reduced prediction bias and BAE compared to participants asked to be objective and a control. This study again looked at participants’ level of confidence in their prediction, to see if any ironic effect of confidence takes place (i.e., that despite participants asked to be objective failing to correct for their biases, they report greater levels of confidence in the accuracy of their decisions).

Method

Participants. One hundred eleven undergraduate students (33% male; 57% Caucasian) in an Introduction to Psychology course volunteered to participate in this study. The median age was 19 (SD = 0.93).

Materials and Procedure. Participants were told that the study was interested in grade prediction and performance. They were asked to predict their score (% out of 100) and their percentile rank on an upcoming midterm exam.

There were three conditions in this study. Participants in the control condition were given no additional instructions. Those in the be-objective condition were told: “Research shows that people have an over-optimism bias when making predictions about their futures. As you make your prediction, please avoid being biased in this way. Try to respond so that you
personally do not believe you have been biased.” Participants in the appear-objective condition were told: “Research shows that people have an over-optimism bias when making predictions about their futures. As you make your prediction, please avoid appearing biased in this way. Try to respond so that an outside observer would not believe you had been biased.” Students were then asked to indicate their confidence in the accuracy of their ratings on a scale of 1 (Not at all) – 9 (Very). They were then asked to give permission to the researcher to obtain their actual midterm grades.

A follow-up survey was emailed to participants after the midterm took place. Respondents ($N = 55$; 31% male) were asked to answer questions about how they would expect a peer to respond to the questions they had previously answered themselves. They were asked to indicate what score (% out of 100) they would expect someone who was objective to respond with. They were likewise asked what percentile rank a person who was objective would give.

**Results and Discussion**

**Prediction bias.** Participants’ predicted scores were compared to their actual midterm scores to assess degree of prediction bias. In no condition was this difference significant, $t’s < 1.72, p’s > .09$, indicating that participants did not show a prediction bias. The effect of condition on the difference score between predicted score and actual score was not significant, with the control ($M = 0.15, SD = 9.81$), be-objective ($M = 0.69, SD = 10.60$), and appear-objective ($M = 2.32, SD = 8.23$) conditions all showing comparable differences, $F(2, 108) = 0.52, p = .59$.

**Better-than-average effect.** Participants’ projected ranks were compared to 50%, or average, to assess the better-than-average effect. The predicted ranks in the control ($M = 67.16, SD = 19.00$), be-objective ($M = 70.17, SD = 15.97$), and appear-objective ($M = 68.64$, $SD = 8.23$) conditions all showing comparable differences, $F(2, 108) = 0.52, p = .59$. 

60
conditions were all higher than average, $t's > 5.57$, $p's < .001$, indicating that participants in all conditions displayed the BAE. The effect of condition on the degree of BAE displayed was not significant, $F(2,107) = 0.29$, $ns$.

When asked in a follow-up survey to indicate what rank would actually be objective as an estimate for someone, participants reported a mean rank of 64.71 ($SD = 17.50$). When compared to this standard, the appear-objective condition’s self-predictions are no longer significantly different, $t(36) = 1.51$, $p = .14$. Likewise, the difference for the control condition was not significantly different from this new standard, $t(37) = 0.79$, $p = .43$. However, the difference was marginal for the be-objective condition, $t(34) = 2.01$, $p = .05$, indicating that participants in this condition were still showing the BAE with this new standard.

The students in this study failed to show a prediction bias in terms of their predicted score, and were actually quite well calibrated in that regard. They did, however, show the BAE in terms of their predicted ranks. At first it would appear that those aiming to appear objective failed, as their mean predicted rank was significantly higher than 50%, or “average”. When compared to what participants themselves thought was an objective response to the question, however, this difference disappeared. According to the students’ criterion for objectivity, they had in fact succeeded. Interestingly, the be-objective condition was the only condition trending toward showing the BAE giving the students’ criterion.

This result underscores the importance of understanding what the criterion for objectivity is in the population under study. What researchers view as an objective response (here: 50%) may differ from the perspective of the population (here: 65%). In such instances, it may be necessary to clarify what the researchers’ criterion is, so that participants are correctly calibrated.
This seems like an interesting boundary condition where appearing objective does not work when what people think is objective does not approximate actual objectivity.

**Study 2.4**

This study brings the appear-objective manipulation into the domain of political partisan bias. Past research has shown people can have diametrically opposing perceptions of the same objective stimulus given their stance in a situation. For example, when Hastorf and Cantril (1954) asked Princeton and Dartmouth students to watch a controversial football game between their respective universities, the two groups of students seemed to have watched two entirely different “games”; each group claimed that the other team was guilty of more infractions than their own.

Additionally, Vallone, Ross, and Lepper (1985) found that when pro-Arab and pro-Israeli students were shown the same news coverage of the Beirut massacre, both sides saw the clips as demonstrating media bias in favor of the other side, an effect known as the hostile media bias (see also, Giner-Sorolla & Chaiken, 1994). Research suggests that an individual’s perception of a political message is shaped equally by their political views as by the objective content of that message (Dalton, Beck, & Huckfeldt, 1998). It follows then, that partisan participants should have differential interpretations of the stimulus in the current study: candidates’ comments in a presidential debate. I predict that this tendency to be swayed in the direction of one’s political party will be mitigated when participants are asked to appear objective.

This study not only examines the effectiveness of appearing objective in a new bias domain – political partisan bias – it also examines potential downstream consequences of this manipulation. In particular, this study explores whether instructions to appear objective induce positive interpersonal behavior. Work on naïve realism indicates that people think they see that
world as it “really is” and that people who disagree with them are therefore biased (Ross & Ward, 1995). People think that members on an out-group (e.g., an opposing political party) are biased by their group membership when making evaluations (Vivian & Berkowitz, 1992) and perceive them as simply following the party line (Cohen, 2003). This predisposition can have negative consequences, as people have a tendency to respond conflictually toward people who disagree with them (and that they therefore perceive as biased), leading to a bias-perception conflict spiral (Kennedy & Pronin, 2008). I hypothesize that the appear-objective manipulation will have positive effects on participants’ behavior toward someone of the opposing political party, leading them to behave in a way that comes across as more objective and more positive. Such a change in behavior could prevent negative interpersonal encounters, such as the bias-perception conflict spiral.

Additionally, this study looks at whether there are consequences for people’s private beliefs after making a decision that appears objective. Does being forced to appear objective lead to reactance in private beliefs? That is, would having to make an outward decision that goes against what one really thinks (in order to meet the appearance criteria) lead people’s private beliefs to become even more biased in response? It would be unfortunate if that were the case, as it would indicate that the appear-objective manipulation has negative long term effects that may outweigh any short-term benefit to a single decision.

Finally, this study also manipulated the target for which participants would be appearing objective. Previous studies have left the target general and unknown. In addition to this original appear-objective condition, this study specifies the target in two new conditions. In one, participants are asked to appear objective to a member of their own political party. I hypothesize that this will not lead to a decrease in political partisan bias, as the target audience would share
the participant’s bias. In fact, it may even lead to an increase in bias. The other new condition asks participants to appear objective to a member of the opposing political party, which should lead to a decrease in political partisan bias similar to the original condition.

**Method**

**Participants.** Two hundred and fifty-one participants (54% male; 71% Caucasian) on Mechanical Turk took an online survey prior to the 2012 Presidential election debates. The median age was 29 (SD = 11.07) and participants’ political self-identification was 66% Democrat, 34% Republican. Participants were compensated at prevailing rates.

**Materials and Procedure.** Participants were asked to indicate whether they identified more strongly as a Republican or a Democrat. They were then told that they would be evaluating a series of quotations said by Barack Obama and Mitt Romney in public speeches on the campaign trail.

The quotations used in this study were collected from transcripts from President Barack Obama and Mitt Romney’s respective presidential primary debates. Quotations were divided into six topics: the Middle East, immigration, the economy, energy, education, and international trade (See Appendix D). Quotations were selected and edited such that they did not contain obvious references to the candidate or candidate’s political party and therefore were not obviously partisan. Across participants, the quotations were counterbalanced such that half of the quotations said by one candidate were attributed to the opposing candidate and vice versa. This equated the quality of the quotations across the two candidates.

Each quotation was paired with a professional portrait of the presidential candidate it was attributed to (See Appendix D). Within each topic, the order of the quotations was
counterbalanced across subjects such that for half of the quotations, Obama’s quotation was shown first and for the other half, Romney’s quotation was shown first.

Participants were asked to read the six pairs of quotations, ordered by topic. They were then asked to categorically select who had made more intelligent comments, either Obama or Romney, for each of the six topics.

There were five conditions in this study. The control condition was given no additional instructions for the task. Participants asked to be objective were told, “People often thoughtlessly follow their political party. As you make your decision about each quote, please avoid being biased by your political affiliation. Try to respond so that you personally do not believe you have been biased by your political affiliation.” Participants asked to appear objective were told, “People often thoughtlessly follow their political party. As you make your decision about each quote, please avoid appearing biased by your political affiliation. Try to respond so that an outside observer would not believe you had been biased by your political affiliation.” The two new conditions manipulating whom participants were appearing objective to resembled the appear-objective condition except for the last sentence. Participants asked to appear objective to a member of their own political party were told, “Try to respond so that a member of your own political party would not believe you had been biased by your political affiliation,” while participants asked to appear objective to a member of the opposing political party were told, “Try to respond so that a member of the opposing political party would not believe you had been biased by your political affiliation.” In all conditions except for the control, participants were asked to type out their instruction word-for-word in order to increase retention of the manipulation. Responses were summed and recoded to represent the number of times (up to
six) participants picked their own party’s candidate as having made more intelligent comments on a topic.

After answering the questions about each topic, participants gave a three sentence response to an open-ended question about who would have won had these quotations been a real debate and why that candidate would have won. They were told that their response would be anonymously shown to someone from the opposing political party and were reminded of their manipulation. These responses were then coded by two independent coders (average $\alpha = .81$). The coding items were adapted from Kennedy and Pronin (2008): how aggressive the participant was toward the other (reverse coded), how amicable the participant was toward other, and the degree of objectivity (vs. bias) in the response. Aggressiveness and amicability were combined into a composite measure of interpersonal positivity, $\alpha = .87$.

Participants were then asked to give their honest opinion on a number of questions including: which candidate won the debate overall, which candidate was more articulate, which was a better candidate overall, and which made more intelligent comments. The scale ranged from 1 (Definitely Romney) to 9 (Definitely Obama) for each question. Responses were recoded to represent 1 (Definitely the opposing candidate) to 9 (Definitely my party’s candidate) and collapsed into a composite variable of honest opinions ($\alpha = .91$) for analysis.

At the end of the survey participants were asked to identify via multiple choice the instructions they had received (i.e. the manipulation) prior to the first task as a manipulation check.

**Results and Discussion**

**Political partisan bias.** It was hypothesized that participants asked to appear objective would display significantly less political partisan bias than participants asked to be objective or
in a control. Additionally, it was hypothesized that asking participants to appear objective to a member of their own political party would not positively impact their degree of bias, but asking them to appear objective to a member of the opposing political party would.

The main effect of condition on how many times out of six participants picked their own party’s candidate as having made more intelligent comments, was not significant across all conditions, $F(4, 244) = 1.08, p = .37$. However, when looking at the planned comparison of the control condition ($M = 4.07, SD = 1.72$) compared to the appear-objective condition ($M = 3.59, SD = 1.25$), the appear-objective condition is picking their own party’s candidate less often than the control, $p = .08$. When taking into account only the participants that passed the manipulation check (83% of participants across all conditions), this difference becomes significant, $p = .04$ (See Figure 2.5). The other three conditions, be-objective ($M = 3.73, SD = 1.54$), appear-objective-to-your-own-political-party ($M = 3.86, SD = 1.12$), and appear-objective-to-the-opposing-political-party ($M = 4.02, SD = 1.16$), did not differ from the control, $p’s > .30$. The appear-objective and be-objective conditions did not differ, $p = .38$.

Since the quotations were counterbalanced between Romney and Obama, the intelligence of the quotations was held constant. Participants in all conditions picked their party’s candidate more than half of the time (i.e., significantly more than three out of six times), $t’s > 3.34$, $p’s < .01$, indicating that all conditions showed political partisan bias to some degree.
Figure 2.5. Mean number of times participants picked their own party’s candidate (out of six) as making more intelligent comments. Error bars indicate 1 SE above and below the mean.

**Interpersonal behavior.** The coding of participants’ open-ended responses for interpersonal positivity, a composite of amicability and aggressiveness (reversed), and degree of bias/objectivity was analyzed. There was a significant effect of condition on interpersonal positivity, $F(4, 244) = 3.39, p = .01$ (See Figure 2.6). The be-objective $M = 3.39, SD = 0.78$, appear-objective $M = 3.27, SD = 0.99$, appear-objective-to-own-political-party $M = 3.35, SD = 0.56$, and appear-objective-to-opposing-political-party $M = 3.27, SD = 0.57$ conditions all displayed greater interpersonal positivity than the control condition $M = 2.89, SD = 0.91$, $p$’s < .02.
Likewise, there was a significant effect of condition on coded ratings of bias/objectivity (higher numbers indicate a greater impression of objectivity), $F(4, 244) = 4.15, p < .01$ (See Figure 2.7). The be-objective ($M = 3.30, SD = 0.77$) and appear-objective ($M = 3.16, SD = 0.85$) conditions both displayed greater objectivity than did the control condition ($M = 2.83, SD = 0.87$), $p$’s $< .03$. The be-objective condition and appear-objective condition did not differ, $p = .34$. The appear-objective-to-own-political-party ($M = 2.90, SD = 0.53$) and opposing-political-party ($M = 2.84, SD = 0.62$) conditions, meanwhile, did not significantly differ from the control, $p$’s $> .63$. 

*Figure 2.6.* Mean coder ratings of open-ended responses for interpersonal positivity by condition. Error bars indicate 1 SE above and below the mean.
Private beliefs. To assess the possibility that participants could display reactance in response to the appear-objective manipulation, participants’ true, private beliefs were analyzed. There was no effect of condition on the composite measure of participants’ private beliefs, $F(4, 246) = 1.38, p = .24$. The be-objective ($M = 5.38, SD = 1.74$), appear-objective ($M = 5.80, SD = 1.60$), appear-objective-to-own-political-party ($M = 6.08, SD = 1.37$), and appear-objective-to-opposing-political-party ($M = 5.96, SD = 1.52$) conditions did not significantly differ from the control condition ($M = 5.93, SD = 1.79$), $p$’s > .08. This indicates no reactance in private beliefs as a result of any of the manipulations.

In this study, the appear-objective manipulation led to a successful decrease in political partisan bias compared to a control. The other three conditions (be-objective, appear-objective-to-own-political-party, and appear-objective-to-opposing-political-party) did not show a
reduction of political partisan bias compared to the control. As in previous studies, it is unsurprising that participants asked to be objective failed to reduce their bias. Without a behavior goal, these participants would have erroneously relied on the content of their introspections to be objective, failing in the effort.

As predicted, participants asked to appear objective to their own political party condition did not show a reduction in bias, as their political parity would have the same bias as themselves. Interestingly, the manipulation did not lead to an increase in political partisan bias. Past research indicates that people view those that agree with them as being objective (Kennedy & Pronin, 2008), a possible explanation for why participants did not have more extreme answers.

Surprisingly, asking participants to appear objective to the opposing political party did not reduce political partisan bias. It is possible that participants displayed reactance in what they may have seen as a request to pander to the opposing political party, leading them to remain firm in their beliefs. Overall, this study suggests that it is most efficacious to ask participants to appear objective without delineating a target.

It is interesting that in any condition that mentioned objectivity, participants displayed significantly greater interpersonal positivity toward a member of the opposing political party in their open-ended response when compared to the control. Regardless of changes in the objectivity of actual decisions, the word “objectivity” may have given participants the goal to be level-headed interpersonally, leading to less aggressiveness and more amicability.

However, the mention of objectivity did not have a similar effect on how biased or objective the person seemed in their open-ended response. It makes sense that participants asked to appear objective came across as more objective than the control in their open-ended responses, as their decisions actually were significantly more objective. Strangely, however, the open-
ended responses of the be-objective condition also came across as more objective. Despite their actual responses not being more objective than the control, asking participants to be objective led their behavior to come across as significantly more objective. This may be because participants asked to be objective did have the goal of objectivity in their decision-making process; they were just unable to attain it in their actual decisions (yet those efforts came through in their interpersonal behavior).

These positive behavioral consequences of being asked to appear objective are important because they are staving off the behaviors that led to the bias perception conflict spiral (Kennedy & Pronin, 2008). In this spiral, people see others who disagree with them as biased and behave conflictually toward them. The target of this behavior then sees the actor as increasingly biased, and escalates his/her conflictual behavior. This pattern continues to escalate and spiral out of control. Instructing people to appear objective has led to significant increases in the positivity and perceived objectivity of participants’ behavior that is directed at a disagreeing other. This result may keep the bias perception conflict spiral from starting in the first place, increasing the likelihood of people of opposing perspectives to successfully work together.

One could worry that asking participants to give an answer that appears objective (and thus may not correspond to their true beliefs) could have a negative impact on their honest opinion. Analyses of participants’ honest opinions, however, indicated that this was not the case. Honest opinions across all conditions were similar, indicating no reactance to the manipulation.

Only when participants were asked to appear objective did both their decisions and their interpersonal behavior become more objective. The manipulation had other beneficial effects as well: participants interacted more positively with a disagreeing other and did not display reactance in their personal beliefs.
Study 2.5

While this chapter has thus far shown the efficacy of asking people to appear objective when making decisions, not all situations have an answer that is clearly objective. How can this manipulation still be used to increase the objectivity of decisions under such circumstances? I hypothesize that asking participants to appear objective will increase their willingness to blind themselves to potentially biasing information.

Externally instituting procedures to prevent people from encountering potentially biasing information is a practice known as exposure control (Gilbert, 1993). In general, people are unlikely to implement blinding procedures for themselves due to the bias blind spot. Research shows that when asked about blinding procedures, people feel that they are more appropriate for others to use than for themselves to use (Mueller & Pronin, 2012). When asked about inputting blinding procedures in a number of domains (e.g. blind grading), participants were the least supportive when the blinding affected the domain of their own decision-making. Blinding may lead to more objective decisions, but if people fail to choose it for themselves, its benefits cannot be secured. I hypothesize that asking participants to appear objective will significantly increase their willingness to blind themselves to potentially biasing information.

Study 2.5 was conducted in the domain of racial discrimination in criminal sentencing. Meta-analyses of sentencing decisions indicate that minorities (i.e., African Americans and Latinos) are more likely to be sentenced to prison than their majority peers and to receive longer sentences (Barnes & Kingsnorth, 1996; Spohn, 2000). Additionally, a meta-analysis of experimental studies showed that cross-racial bias consistently affects mock jurors’ sentencing decisions of minority defendants (Sweeney & Haney, 1992).
In addition to looking at the effect of appearing objective on racial discrimination in this context, this study looked at the effect of appearing objective on participants’ willingness to blind themselves to potentially biasing information (mugshots). I predict that participants asked to appear objective in their decision-making will be more likely to choose to blind themselves to the potentially biasing information than participants in a control condition or participants asked to be objective.

**Method**

**Participants.** One hundred twenty-one participants (53% male; 75% Caucasian) on Mechanical Turk took part in this study. The median age was 28 (SD = 11.44). Participants were compensated at prevailing rates.

**Materials and Procedure.** Participants were asked to rank order 24 crimes in order of sentence length (i.e., position one goes to the crime deserving the longest sentence). They were told that each crime would be accompanied by the mugshot of the person who had committed the crime.

The stimuli were piloted by a separate set of raters (N = 48; 33% male). Mugshots of African American and White men were piloted for age and crimes were piloted for deserved sentence length. Twelve cross-race pairs were created such that the mugshots matched for age and the crimes matched for deserved sentence length (See Appendix E). The crimes were counterbalanced across race within each pair across subjects. Crimes consisted of the name of the crime (e.g., “Embezzlement) and a description of the crime (e.g., “At work, he regularly pocketed the cash from a sale instead of ringing it up and putting the money in the cash register.”).
The raters rated the crimes for how much time the person should serve in prison on a scale of 1 [Not much time (months)] – 4 [A moderate amount of time (years)] – 9 [A long time (decades)]. There was not a difference in severity for the first \( (M = 4.16, SD = 1.09) \) and second \( (M = 4.14, SD = 1.07) \) groups of crimes, \( F(1, 22) = .001, ns. \) Raters rated the mugshots for apparent age using a sliding scale from 20 to 80 years old. There was not a difference in age between the groups of African American \( (M = 41.58, SD = 12.34) \) and White \( (M = 41.58, SD = 12.63) \) mugshots, \( F(1, 22) = .00, ns. \)

Participants in the control condition did not receive additional instructions. Participants in the be-objective condition were told: “Research shows that the race of a defendant affects criminal sentencing. As you make your decision about each criminal, please avoid being biased by race. Try to respond so that you personally do not believe you have been biased by race.” Participants in the appear-objective condition were told: “Research shows that the race of a defendant affects criminal sentencing. As you make your decision about each criminal, please avoid appearing biased by race. Try to respond so that an outside observer would not believe you had been biased by race.” Participants in the be- and appear-objective conditions were asked to type out the above instructions word-for-word in order to increase retention of the manipulation.

Before ranking the crimes, but after receiving the manipulation, participants were told that a limited number of participants could choose to remove the mugshots paired with the crimes and that doing so would remove all appearance information about the criminal except gender. They were then given the option to leave or remove the mugshots.

Regardless of their decision, all participants were shown the mugshots during the ranking task. Participants then rank-ordered the 24 crimes/criminals. At the end of the survey,
participants were asked to indicate their instructional manipulation via multiple choice as a manipulation check.

**Results and Discussion.**

**Blinding.** Participants’ decisions to blind themselves (or not) to the mugshots were analyzed. Across the three conditions, there was a marginal effect on blinding preference, \( \chi^2(2) = 4.55, p = .10 \). In our key comparison of the control (61% chose to see, 39% chose to blind) and appear-objective (37% chose to see, 63% chose to blind) conditions, there was a significant effect of condition on blinding preference, \( \chi^2(1) = 4.50, p = .03 \), with participants in the appear-objective condition choosing to blind themselves significantly more often. The be-objective condition (47% chose to see, 53% chose to blind) did not significantly differ from either of the other conditions, \( p’s > .24 \) (See Figure 2.8)

![Figure 2.8](image-url)  

*Figure 2.8.* Percent of subjects in each condition choosing to either see the mugshots or hide them (i.e., blind themselves).
**Racial discrimination.** The effect of condition on the average ranking of the crimes (out of 24) when paired with an African American mugshot was analyzed (lower numbers indicate longer sentences). There was no difference between the control \((M = 12.62, \text{SD} = 1.02)\), be-objective \((M = 12.75, \text{SD} = 0.87)\), and appear-objective \((M = 12.75, \text{SD} = 0.91)\) conditions, \(F(2, 119) = 0.26, \text{ns}\). Similarly the difference in average rank by race between the control \((M = 0.25, \text{SD} = 2.04)\), be-objective \((M = 0.51, \text{SD} = 1.75)\), and appear-objective \((M = 0.50, \text{SD} = 1.82)\) conditions was not significantly different, \(F(2, 116) = 0.90, p = .78\). The difference in average ranking for African American and White mugshots was also analyzed for evidence of racial discrimination; there was not a difference in any of the conditions, \(t’s < 1.79, p’s > .08\), indicating no presence of racial discrimination in the task².

In this study, participants asked to appear objective were significantly more likely to choose blinding than participants in the control. This is evidence that the appear-objective manipulation can have a beneficial impact on decision-making beyond any effect directly on a decision. The goal to appear objective increased participants’ willingness to remove the source of bias prior to even making a decision.

Racial discrimination was not present in this study; participants rated crimes similarly when paired with African American and White mugshots. This study cannot therefore lead to any conclusions on the efficacy of appearing biased on the issue of racial discrimination in criminal sentencing. Much past research documenting the effect of racial discrimination on sentencing has been field research focused on the outcomes of actual trials (e.g., Barnes & Kingsnorth, 1996; Spohn, 2000), while this study attempted to demonstrate racial discrimination in a survey-based experiment.
Methodologically, this study may have been too overt to document racial discrimination. For example, it did not measure participants’ individual levels of prejudice. Past research shows that people high and low in prejudice vary greatly in the degree to which they endorse racial stereotypes (e.g., Devine & Elliot, 1995). Additionally, this study only compared mugshots of African-Americans to White mugshots. Research has shown that within the African-American racial category, individuals with more Afro-centric features are more likely to receive longer sentences than their peers with less pronounced Afro-centric features (Pizzi, Blair, & Judd, 2005). Modern anti-racist norms have led to a subtle form of racism that this paradigm may not have been able to tap into (e.g., Dovidio & Gaertner, 2004).

**Study 2.6**

This study again examines the effectiveness of the appear-objective manipulation on increasing people’s willingness to blind themselves to potentially biasing information. This study looks at decisions within the realm of artistic evaluation, in particular, of rating paintings for artistic merit. Studies 1.2 and 1.3 demonstrated that people are biased by the fame of paintings’ artists when rating them for artistic merit. When evaluating paintings, participants rated paintings with famous artists as having significantly higher artistic merit than those with non-famous artists. Participants did not think that the fame influenced their ratings, despite having admitted that looking at the name of the artists was a biased way to evaluate the artistic merit of the paintings.

In situations such as this, there is no clear way to make a decision that would appear objective. Objectivity may best be reached through choosing to remove the potential sources of bias: the names of the painters. I hypothesize that participants asked to appear objective will be more willing to blind themselves to the identity of the painters than will participants asked to be
objective or in a control. Participants’ self-reported reasons for blinding (or not) will be examined.

Method

Participants. One hundred and four participants (71% male, 71% Caucasian) on Mechanical Turk participated in this study. The median age was 26 (SD = 6.92). Participants were compensated at prevailing rates.

Materials and Procedure. Participants were told that they would be evaluating a series of paintings for artistic merit, or quality. They were told that each painting would be accompanied by the name of its artist and that some artists would be famous while some would not.

Participants in the control condition were not given additional instructions. Those in the be-objective condition were told: “Research shows that people are affected by the fame of an artist when evaluating art. As you make your decision about each painting, please avoid being biased by the artist. Try to respond so that you personally do not believe you have been biased by the artist.” Those in the appear-objective condition were told: “Research shows that people are affected by the fame of an artist when evaluating art. As you make your decision about each painting, please avoid appearing biased by the artist. Try to respond so that an outside observer would not believe you had been biased by the artist.” Participants in the latter two conditions were asked to type their instructions word-for-word in order to increase retention of the manipulation.

Participants were told that they could remove the names of the artists and were asked to indicate if they would like to leave the names of the artists or remove them. After making this decision, participants were never asked to rate paintings. Instead, they were asked to explain
why they made their decision in their own words in an open ended response. Participants who chose to leave the names of the artists were then asked to indicate via multiple choice, what best described why they chose to leave the names of the artists: 1 (I did not think they would bias my ratings); 2 (I thought the artists would be informative); 3 (I just wanted to see them); 4 (None of the above). Participants who chose to remove the artists were similarly asked to indicate which option best described their reasoning: 1(I was worried the artists would bias my ratings); 2 (It seemed like the more objective decision); 3 (I just did not want to see them); 4 (None of the above).

At the end of the study, participants were asked to indicate via multiple choice the wording of their original instructions as a manipulation check.

**Results and Discussion**

**Blinding.** Participants’ decisions to leave or remove the names of the artists were first analyzed. It was hypothesized that participants instructed to appear objective would be more likely to remove the names of the artists – that is, to choose to blind themselves to the potentially biasing information. There was a significant effect of condition on participants’ willingness to blind themselves to the identity of the painters, \( \chi^2(2) = 8.84, p = .01 \) (See Figure 2.9).

Participants asked to appear objective chose to blind themselves 67% of the time, significantly more often than the control (35%), \( \chi^2(2) = 7.25, p < .01 \). Similarly, participants asked to be objective chose to blind themselves 63% of the time, also significantly more often than the control, \( \chi^2(2) = 5.39, p = .02 \). The be- and appear-objective conditions did not differ, \( p = .73 \).
Figure 2.9. The percentage of participants choosing to keep the names of the painters v. remove the names of the painters (i.e., blind themselves) by condition.

Reasoning. Participants self-reported reasons for their blinding (or not) decisions were analyzed next. There was not an effect of condition on participants’ reasons not to blind themselves, $\chi^2(6) = 5.42, p = .49$, with the most participants in each condition indicating that they chose to leave the artists because they thought the artists would be informative (control = 48%, be-objective = 33%, and appear-objective = 55%).

There was, however, a significant effect of condition on participants’ reasons for blinding, $\chi^2(4) = 10.02, p = .04$ (See Table 2.1). Participants’ reasons in the appear-objective condition were significantly different from the control, $\chi^2(1) = 5.84, p = .02$, and marginally different from the be-objective condition, $\chi^2(2) = 5.92, p = .05$ (see Table 2.1). The control and be-objective conditions did not differ, $\chi^2(2) = 0.97, p = .62$. The majority of participants in the appear-objective condition indicated that they chose to blind themselves because it seemed like
the more objective decision, while the majority of participants in the control and be-objective conditions indicated that their decision to blind was due to concern that the artists would bias their ratings.

Table 2.1. Percentage of participants in each condition selecting each reason for blinding (of those that chose blinding).

<table>
<thead>
<tr>
<th>Reason</th>
<th>Control</th>
<th>Be Objective</th>
<th>Appear Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was worried the artists would bias my ratings</td>
<td>86%</td>
<td>75%</td>
<td>46%</td>
</tr>
<tr>
<td>It seemed like the more objective decision</td>
<td>14%</td>
<td>20%</td>
<td>55%</td>
</tr>
<tr>
<td>I just did not want to see them</td>
<td>0%</td>
<td>5%</td>
<td>0%</td>
</tr>
</tbody>
</table>

This study further showed that asking participants to appear objective leads to a significant increase in willingness to blind the self from potentially biasing information. It is interesting that in this study, participants asked to be objective chose to blind themselves just as often as participants asked to appear objective. The nature of the paradigm may have created a demand effect for the participants in the be-objective condition to respond this way. After being told that people are influenced by the fame of the artists, it is clear that the researcher would want participants to select this option.

It is interesting that participants gave different reasons for choosing blinding across conditions. When asked to appear objective, participants were not concerned with whether or not the painters would bias their ratings; they chose blinding because doing so was the objective
thing to do. It is not surprising that the participants in the control and be-objective conditions who chose blinding indicated that they did so due to possible influence of the painters on their ratings of the paintings. The research presented in Studies 1.2-1.4 demonstrated that people recognize that evaluating paintings while knowing the name of the artists is a biased thing to do.

**Discussion**

This chapter showed the effectiveness of instructing people to appear objective across a variety of bias domains: sexism (Studies 2.1a-2.1c); the better than average effect (Study 2.2.); and political partisan bias (Study 2.4).

Studies 2.1c and 2.2 both showed that the appear-objective manipulation does not lead to an increase in the amount of effort participants put into their decision-making. Study 2.2 actually suggests that in straight-forward situations (i.e., when the answer that appears objective is very clear) appearing objective may actually take less effort. This finding sets the appear-objective manipulation apart from work on accountability, which operates by increasing the amount of cognitive effort people put into their decisions (Lerner & Tetlock, 1999).

Interestingly, this study also showed that asking participants to appear objective did not reduce their confidence in the accuracy of their decisions.

The results of Study 2.3 underscore the importance of understanding what the population under examination believes is “objective.” If the population’s interpretation of objectivity differs from the experimenter’s objective criterion, this manipulation will not be as effective. Only under circumstances where the implementers of the manipulation have an agreed-up standard for objectivity and where the decision-makers share that standard, can this manipulation have a successful, direct impact on the objectivity of decisions.
This chapter also demonstrated that instructing people to appear objective increases their willingness to blind themselves to potentially biasing information when evaluating criminals (Study 2.5) and art (Study 2.6). This has important implications for the objectivity of decisions, as people cannot be biased by information they do not have.

Study 2.6 suggests that there are differences in why participants choose to blind themselves. Participants in a control or asked to be objective more often choose blinding out of concern for the potential of bias to enter into their decision-making process. Participants asked to appear objective seem to choose blinding for a more abstract reaction: because it is a more objective thing to do. The benefit of the latter reasoning is that it could lead people who do not fear for the potential influence of bias to still choose blinding.

It should be noted, that while asking participants to appear objective did significantly increase the objectivity of participants’ decisions compared to a control across studies, it only inconsistently increased objectivity compared to participants asked to be objective. For example, in Study 2.4, looking at political partisan bias, the level of objectivity in the be-objective conditions falls between the control and appear-objective conditions, significantly differing from neither. This inconsistency in the be-objective condition may be due participants’ motivation to make an objective decision, but lack of clarity on how to do so. While they may occasionally make movements toward objectivity, their reliance on introspection prevents them from successfully realizing that goal, and never significantly improving over a control. Participants instructed to appear objective, however, have a clear path to making an objective decision which consistently allows for the successful realization of their goal.

Similarly, in the two studies examining blinding decisions (Studies 2.5 and 2.6), the be-objective condition does not significantly differ from the appear-objective condition, and in the
latter case is even significantly different from the control. While the research in Chapter 1 indicates that these participants would not believe they had been biased after *experiencing* the decision-making task, it suggests they would recognize the possibility of bias *before* engagement. It is therefore not surprising that participants in these studies showed at least some greater willingness to blind themselves to potentially biasing information prior to engaging in the task. In conclusion, asking participants to appear objective is the only manipulation that leads to *significant increases* in objectivity and *consistently* leads to a greater willingness to implement blinding procedures.

Together, these studies not only show the effectiveness of asking people to appear objective across a variety of bias domains, but also in situations that do and do not have clearly objective answers. Repeatedly, the research reported in this chapter showed that instructing people to appear objective is more effective than leaving them to make their own decision and than instructing them to be objective. By shifting participants’ focus to that of an outside observer (and therefore away from their introspections), the objectivity of their decisions was reliably increased.
Conclusions and Future Directions

This research constitutes another chapter in a continuing story about the nature of the “bias blind spot” (e.g., Ehrlinger, Gilovich, & Ross, 2005; Frantz, 2006; Pronin, Lin, & Ross, 2002; West, Meserve, & Stanovich, 2012). Past research has shown that while individuals are quick to point out bias in others, they are far less likely to admit to bias in themselves. Chapter 1 supports and strengthens the notion that people have difficulty recognizing their own biases. It shows that even when people acknowledge that what they are about to do is biased, they still are inclined to see their resulting decisions as objective. The current research suggests that this due to people’s reliance on introspection for evaluating the possible effect of biases that actually operate unconsciously (and are therefore inaccessible via introspection).

Chapter 1 explored actors’ self-assessments of objectivity following a biased judgmental strategy, but what of observers’ assessments? If a judge, or hiring official, or athletic coach openly acknowledges the potential for bias in his or her judgment process, what do onlookers conclude about his or her resulting judgment? It seems possible that actors’ admissions of bias in their judgmental processes and their subsequent conclusions of objectivity could be mirrored by onlookers. Those observers might think something like: “If he was smart enough to know this bias existed, and honest enough to acknowledge it, surely he wouldn’t fall prey to it! At least not as much as those who don’t even recognize and admit to such things.”

This research contributes to an existing literature illustrating that efforts at debiasing are fraught with difficulty (e.g., Epley & Gilovich, 2006; Pronin et al., 2002; Stapel et al., 1998; Wegener & Petty, 1995). People have difficulty correcting for biases that have already influenced their judgments, because they are unsure of how much to correct (and often deem any correction unnecessary).
The research presented in Chapter 2 of this dissertation introduces a successful manipulation to increase the objectivity of decisions while still allowing for the bias blind spot. Asking people to make a decision that would appear objective to an outside observer led to more consistently objective decisions than asking people to be objective or leaving them to make a decision on their own. This manipulation was explored in a variety of bias domains from sexism to political partisan bias. Additionally, steps were taken to distinguish this research from work on accountability (Lerner & Tetlock, 1999), demonstrating that appearing objective can actually take less effort than attempting to be objective.

Potential downstream consequences of asking people to appear objective were explored. In addition to decreasing bias in their decisions, asking people to appear objective had a positive effect on their interpersonal behavior. This is an important consequence as the appear-objective manipulation may lead to a reduction in conflict between parties that disagree. Future research should examine how asking people to appear objective affects people’s views of disagreeing others. Since it leads them to behave more positively toward another, it may lead them to view that other more positively as well. Past research shows that people view those that disagree with them to be biased, leading them to behave conflictually toward them (Kennedy & Pronin, 2008). This pattern of perception and behavior can turn into a bias-perception conflict spiral. The present research already shows positive effects of appearing objective on behavior toward a disagreeing other; it may have positive effects on interpersonal perception as well, further helping to avoid this conflict spiral. Future research should look at the effect of appearing objective on behavior within a real interaction. Are disagreeing partners who were instructed to appear objective able to work better together?
In addition, Study 2.4 suggested that appearing objective does not lead to reactance in participant’s private beliefs. Future research should explore whether repeatedly appearing objective leads private beliefs to resemble those public decisions over time. Also, the effect of appearing objective on subsequent tasks should be examined. If a person appears objective on one decision, does it positively influence a future, unrelated decision? It is possible that appearing objective could have additional positive downstream consequences on both similar and dissimilar decisions.

Studies 2.5 and 2.6 demonstrated a way in which the appear-objective manipulation remains beneficial even in situations where there may not be a clearly objective response. In two bias domains, instructions to appear objective lead to a significant increase in participants’ willingness to blind themselves to potentially biasing information compared to a control. Given past research on people’s general unwillingness to blind themselves (Mueller & Pronin, 2012), this is an important result: a few sentences can dramatically improve not just the decisions that people make, but the way in which they make them. Participants in Study 2.6 indicated that they chose to blind themselves because it was the more objective course of action. Future research should be conducted to tie this greater willingness to blind the self to a decrease in bias in actual decisions.

Future research should explore the mechanism underlying this phenomenon. I believe that asking people to appear objective increases the objectivity of their decisions by focusing their attention outward on their behavior. When asked to be objective, people look inward, focusing on their internal thoughts and feelings when assessing their own potential bias. Past research has shown that this reliance on introspection is what leads to the bias blind spot (e.g. Pronin & Kugler, 2007). The current research suggests that asking participants to appear
objective to an outside observer shifts the focus of the attention outward, increasing the objectivity of their decisions.

This dissertation has contributed to work on the bias blind spot and bias correction. Chapter 1 demonstrated the extent to which people are inclined to view themselves and their decisions as objective. Even when people recognize that their decision-making strategy is biased, they still fail to recognize its effects on their judgments. This research suggests that this error is due to the unconscious nature of bias; when people to introspect to determine if they had been influenced by bias, the fail to see any signs of bias because that bias is operating unconsciously.

Still allowing for this bias blind spot, Chapter 2 introduced a manipulation that successfully reduces bias in decisions. A series of studies demonstrated the effectiveness of asking people to make decisions that would appear objective to an observer. The bias blind spot may have a powerful effect on people’s perceptions of their own biases, but this effect can be circumvented by asking people to take the perspective of an outsider. This work suggests that greater objectivity in decision-making can be achieved by shifting a person’s focus outward to their behavior, and asking them to mold that behavior to a normatively agreed upon objective response. In the future, it will be important to document the effectiveness of the appear-objective manipulation on decisions with greater consequences both for the decision-maker and the target of the decision. Determining the logistics and effectiveness of this manipulation in situations with important real-world consequences, like hiring, would be a valuable contribution to work on decision-making.
References


Footnotes

1 In some instances, knowing the name of the artist (and, therefore, his or her degree of fame and/or talent) might not be a “biasing” influence on judgments of artistic merit, but rather a rational criterion to use in judging that merit. Because the present research concerns people’s perceptions about whether they are biased, we are less concerned with whether this influence is “objectively” biasing than with the fact that participants themselves would view it that way.

2 A follow-up study was run where question wording was changed to focus on the criminals rather than the crime and each criminal/crime was rated on a scale (rather than ranked). There was still no effect of race on the task.
Appendix A

Social intelligence test (Study 1.1)

Photos

Statements

1. I am a Navy SEAL, but I actually do have a soft side. When I’m off duty, I love to go swing dancing. I’ve always liked dogs, and I’ll probably get a lab when I have the time.

2. I’m a construction worker. My real hobby is just spending time with my wife and two daughters. I like to kick back and watch Letterman at night.

3. I am a police officer. At the end of the day, I like to go home and read the newspaper. I also like to play with my three cats.

4. I’m a software engineer. I like to paint and listen to jazz.

5. I’m an office manager. My top hobby is probably going jogging – I have a great group of friends I go jogging with every day. I also really like going to garage sales and looking for bargains!
6. I work as a stock broker, which really takes up most of my time. When I’m off work, though, I love to watch sports. I like my new sports car.

7. I’m an administrative assistant. I enjoy doing aerobics. I also like going to hear live bands.

8. I’m an environmental consultant. In my spare time, I like to write poetry. I also enjoy good food and cooking.

9. I’m a student majoring in biology. I like to go skiing and to read.

10. I am a bank teller. I don’t really have any exciting hobbies, but I do enjoy watching TV while having a snack, like chips and salsa.

11. I work at a biotech company doing pharmaceutical research. I always go rock climbing on the weekend, and I love alternative music.

12. I’m a psychotherapist. My hobby is practicing the violin. When I have time, I like to go browsing for antiques.

13. I’m a waitress. I love going to the movies. I also really like sunsets.

14. I teach English at the local high school, but my hobby is coaching the varsity soccer teams. I also really enjoy learning foreign languages, and I hope to take a trip to Europe this summer.

15. I’m an accountant. My hobby is mountain biking. I like shopping for new electronic gadgets.

16. For my job, I’m a librarian, but my passion is designing clothing. I would love to have a fancy computer so I could use it for new designs and to sell the clothes on-line.
17. I work as a private investigator. My hobby is photography, and when I’m not working, I like vacationing at the beach.

18. I’m in sales. I love to go country-music line dancing and play pool at the local bar.
Appendix B

Example painting stimuli (Studies 1.2 - 1.34)

Example for painting attributed to famous painter (Studies 1.2 & 1.3)

Example portrait/name of famous painter (Study 1.4)
Example mugshot/identifying information for criminal painter (Study 1.4)

#79923-974

Stabbed victim
Appendix C

Employee stimuli (Studies 2.1a-2.1c)

Example employee biographies with photos
Employee biographies

1. I work in the Platforms and Services Division at -- Corporation. I am currently focused on defining -- Corporation’s future storage platform, spanning from devices to desktops to enterprises to cloud storage. I have a bachelor’s degree in Natural Resources Management from Colorado State University and two Master’s degrees--Forestry and Computer Science--from the University of Wisconsin, Madison.

2. I work with various groups within the company to help define and expand efforts in the areas of parallel and high-performance computing. I am recognized internationally in high-performance computer architecture and programming languages for parallel computers. I attended the University of New Mexico, where I earned a bachelor of science in electrical engineering, and the Massachusetts Institute of Technology, where I earned a masters and a doctorate in electrical engineering.

3. I work on -- Corporation’s datacenter operating system. I am widely recognized expert in our operating system internals as well as operating system security and design. I have been a featured speaker at major industry conferences.

4. Since joining -- Corporation, I have played a key role in shaping the company's Internet strategy and the object and distributed networking architectures. Most recently, I have made contributions to -- Corporation's vision and design for unified communications and presence services. I have served as the technical editor of "AI Expert Magazine" and have written many articles on computer languages and artificial intelligence, many of which appeared in "Computer Language Magazine".
5. I am in the Online Services Division (OSD) working as the chief architect for Search, responsible for technical direction across Search. Within OSD I advise leadership and exercise thought leadership for cross-division technology matters. I also work with leaders across the corporation to align and evolve -- Corporation’s cloud architecture and strategy.

6. I am chief architect of the -- Corporation Business Intelligence (BI) offerings and I hold over 40 patents. I joined -- Corporation when the corporation acquired Panorama Software, where I served as vice president of R&D. I have held roles as the development lead, development manager and product unit manager for analysis services in SQL Server.

7. As a technical leader in the Developer Division driving -- Corporation’s browser programmability and tools efforts, I have been responsible for the technical innovations that have helped make our web browser an industry leader in platforms across desktop and devices. I received a Ph.D. from UC Berkeley where I pursued research interests in programming language design and implementation, natural language understanding, and runtime systems for multi-core processors. I have also pursued these interests as an undergraduate at Yale, a researcher at Bell Labs and a faculty member at Carnegie Mellon University.

8. I am in -- Research, working on security, privacy, and fault-tolerance, and offering suggestions in systems, networking, and other areas. I also worked on anti-piracy, security, fault-tolerance, and user interfaces. I received an AB from Harvard University and a PhD in EECS from the University of California at Berkeley.
9. I am the chief designer of the C# programming language and a key participant in the development of the -- framework. I co-authored a book on the C# programming language, and have received numerous software patents. I was the recipient of the prestigious Dr. Dobbs Excellence in Programming Award.

10. I work as the Product Unit Manager for Team Foundation Server - a server-based product designed to dramatically improve the productivity, predictability, and agility of software development teams by ensuring that all team members have easy access to the information they need to make the right decisions at the right time. I previously worked on the problem of improving the approachability of API for the developer masses.

11. At -- Corporation, I am responsible for driving technical strategy and the cloud transition for -- directory, identity, access, privacy, and information protection technologies. In this capacity I work closely with leaders across the corporation to align overall services strategy and product architectures. I earned a Master of Computer Science degree from Brown University and hold a Bachelor in Electrical Engineering from Union College.

12. My new role is creating and leading the -- Systems Lab, a new advanced development center in association with the University of Wisconsin-Madison Computer Sciences Department. I am a member of the National Academy of Engineering, a fellow of the American Academy of Arts and Sciences, and an ACM (Association for Computing Machinery) Fellow. I have authored over 120 technical publications and served on numerous program committees and NSF (National Science Foundation) Review Panels.
13. Currently responsible for the design of the -- Operating System, I am generally considered one of the top programmers worldwide. In addition to leading the team, I contributed to the architecture of all parts of the system, and even wrote the kernel myself. I hold over 20 patents and am an affiliate professor in the Computer Science Department at the University of Washington.

14. As a key technical member of the storage engine team, I implemented the SQL Server lock manager and other critical concurrency control mechanisms. Through the SQL Server 2000 and SQL Server 2005 releases, I served in a variety of roles including product-level architect and general manager of product development. I am currently serving as general manager of -- Company’s Data and Modeling Group, which oversees data modeling and data access strategies.

15. I work in the Developer Division was one of the key people who established -- Corporation’s service and consulting business on the West Coast. My work with large customers ultimately led to help turn – Corporation into a credible enterprise software supplier. In addition to my management responsibilities, I contributed hands-on code to core components of MTS, such as the context wrapper, administration, performance monitoring and clustering.

16. Before joining -- Corporation, I worked at IBM developing fundamental data mining concepts and technologies and pioneering key concepts in data privacy, including Hippocratic Database, Sovereign Information Sharing, and Privacy-Preserving Data Mining. Scientific American named me one of the 50 top scientists and technologists. I received my M.S. and Ph.D. degrees in Computer Science from the University of Wisconsin-Madison.
Appendix D

Candidate quote stimuli (Study 2.4)

Example stimuli with photos

Education

“We’ve got to restore America’s promise in this country where people know that with hard work and education, that they’re going to be secure and prosperous and that their kids will have a brighter future than they’ve had.

The principles of testing our kids, excellent curriculum, superb teachers, and school choice, those are the answers to help our schools.”

“We’ve got to improve K through 12. And that means not just talking about how great teachers are but rewarding them for their greatness by giving them higher salaries and giving them more support and professional development.

Every dollar that we spend in early childhood education, we get $10 back in reduced dropout rates, improved reading scores. That’s the kind of commitment we have to make early on.”
Quotes by topic with correct attribution

Immigration

Obama: “We have a broken legal immigration system that has to be expedited. That's part of the problem that we're seeing. It's too cumbersome. It's often too expensive and unwieldy.

As president of the United States, I will make sure that the federal government does what it's supposed to do, which is to do a better job of closing our borders and preventing hundreds of thousands of people to pour in, and have much tougher enforcement standards when it comes to employers.”

Romney: “You know, I think Latino voters, like all voters in this country, are interested in America being an opportunity nation. People come here because they believe they want to have a brighter future and that’s been the story of America.

So, the right course for us is to, once again, talk about what you described. Secure the border. Once we do that, we can start talking about the 11 million or whatever number that may be that are in the country illegally.”

The Middle East

Obama: “I think we should always cooperate with our allies and sovereign nations in making sure that we are rooting out terrorist organizations. But if they are planning attacks on Americans like what happened on 9/11, it is my job, it will be my job as president to make sure that we are hunting them down.

I believe that that includes direct talks with the Iranians, where we are laying out very clearly for them: Here are the issues that we find unacceptable, not only development of nuclear weapons, but also funding terrorist organizations and their anti-Israel rhetoric and threats toward Israel.”
**Romney:** “Let me tell you, people who join al Qaeda are not entitled to rights of due process under our normal legal code. They are entitled instead to be treated as enemy combatants.

If we can turn Syria and Lebanon away from Iran, we finally have the capacity to get Iran to pull back. And we could, at that point, with crippling sanctions and a very clear statement that military action is an action that will be taken if they pursue nuclear weaponry, that could change the course of world history.”

**The Economy**

**Obama:** “We have to stop providing tax breaks for companies that are shipping jobs overseas and give those tax breaks to companies that are investing here in the United States of America.

What I want is not oppressive taxation. I want businesses to thrive, and I want people to be rewarded for their success.

We can hire young people who are out of work and put them to work at a trade. So there are all sorts of things that we're going to have to do to make the United States economy much more competitive and those are plans that I have put forward in this campaign and I expect to pursue as president of the United States.”

**Romney:** “There are a lot of good jobs we need in manufacturing and high-tech jobs and good service jobs, technology of all kinds. America produces an economy that’s very, very broad, and that’s why our policy to get America the most attractive place in the world for investment and job growth encompasses more than just energy. It includes that, but also tax policy, regulatory policy, trade policy, education, training and balancing the federal budget.

You know, the right course for America is to have a president who understands the economy and will make that his focus and put in place a plan to get this economy going.”
Energy

Obama: “Part of the reason that Kuwait and others are able to come in and purchase, or at least bail out, some of our financial institutions is because we don't have an energy policy. And we are sending close to a billion dollars a day.

We need to put forward a realistic plan that is going to reduce our dependence on foreign oil, is going to invest in solar and wind and biodiesel.”

Romney: “We have an energy policy that doesn’t take advantage of our natural resources. That makes no sense. We’re an energy-rich nation that’s acting like an energy-poor nation.

I want to find a way to get our energy resources — and they’re all over this country — used for us. This is time to get America growing again, and that’s what this campaign ought to be about.”

Education

Obama: “We've got to improve K through 12. And that means not just talking about how great teachers are but rewarding them for their greatness by giving them higher salaries and giving them more support and professional development.

Every dollar that we spend in early childhood education, we get $10 back in reduced dropout rates, improved reading scores. That's the kind of commitment we have to make early on.”

Romney: “We've got to restore America's promise in this country where people know that with hard work and education, that they're going to be secure and prosperous and that their kids will have a brighter future than they've had.

The principles of testing our kids, excellent curriculum, superb teachers, and school choice, those are the answers to help our schools.”
International Trade

Obama: “We are the biggest market for China. They can't afford to just say, "See ya later." They're going to have to sell here. And if we tell them you have to meet certain safety standards, that you have to enforce certain labor and environmental agreements, they will meet them.

We have to be tougher negotiators with China. They are not enemies, but they are competitors of ours. And on the economic front, on trade issues, on issues in importation, we have not been the best negotiators.”

Romney: “China is playing by different rules. They are stealing intellectual property. They are manipulating their currency, and by doing so, holding down the price of Chinese goods, and making sure their products are artificially low-priced. It's predatory pricing, it's killing jobs in America.

If I'm president of the United States, I'm making it very clear, I love free trade. I want to open markets to free trade. But I will crack down on cheaters like China. They simply cannot continue to steal our jobs.”
Appendix E

Sentencing stimuli (Study 2.5)

Example stimuli with mugshots

1. Crime: Aiding and Abetting
   Description: He convinced his coworker to carry a gun to work and helped him plan how to do it.

2. Crime: Armed Robbery
   Description: He threatened a shop owner with a gun as he forced the shop owner to empty the cash register.

3. Crime: Auto Theft
   Description: He was caught driving a Honda Civic reported stolen three days ago.

4. Crime: Breaking and Entering
   Description: He entered his ex-girlfriend’s apartment with a gun while she was sleeping.
5. Crime: Child Abuse
Description: He slapped his son across the face, breaking his nose, for disobeying.

6. Crime: Child Pornography
Description: He had a hidden folder of images of nude children in sexual poses on his laptop.

7. Crime: Conspiracy
Description: He and his friend planned to smuggle drugs across the border from Mexico. After making their plan, he bought a gun to ensure things would go smoothly.

8. Crime: Drug Possession
Description: He had 2.5 grams of cocaine hidden in his shoe.

9. Crime: Drug Trafficking or Distribution
Description: He asked his friend, an undercover cop, to help him move and repack a trunk full of marijuana.

10. Crime: Embezzlement
Description: At work, he regularly pocketed the cash from a sale instead of ringing it up and putting the money in the cash register.

11. Crime: Extortion
Description: He obtained risqué photos of a local businesswoman and told her that if she gave him $50,000, he would not sell them to the media.

12. Crime: Forgery
Description: He stole a blank check and printed new checks with the same account information. He then went to check cashing agencies to cash the false checks.
13. Crime: Hate Crime
Description: He killed a coworker a few days after discovering that the coworker was homosexual.

Description: He stole a child’s social security number and used it to take out loans and get credit cards.

15. Crime: Kidnapping
Description: He flew across the country with his young daughter where he changed their names and tried to hide from his ex-wife.

Description: While working as a high school substitute, he touched a 14 year old female student on the behind and made a suggestive comment.

17. Crime: Manslaughter
Description: He caught his girlfriend in bed with a lover, and in a fit of rage grabbed a bedside lamp and hit the lover on the head, killing him.

18. Crime: Money Laundering
Description: He had $100,000 in drug money deposited in 25 U.S. banks in amounts ranging from $500-6,000. He then had the funds transferred overseas and back and used the money to purchase real estate to make the money appear legitimate.

19. Crime: Perjury
Description: He lied about his friend’s whereabouts during a trial, giving his friend a false alibi.
20. Crime: Rape
Description: He got a woman at a bar very drunk and then followed her home and forced her to have sex with him.

21. Crime: Second Degree Murder
Description: Having a bad relationship with his neighbor, one day he decided to shoot at this neighbor as the neighbor pulled out of his driveway. His shot missed the neighbor, killing the neighbor’s young daughter instead.

22. Crime: Sexual Assault/Battery
Description: He got undressed and caressed a female who was passed out at a party.

23. Crime: Statutory Rape
Description: He had sex with a minor.

24. Crime: Wire Fraud
Description: He called up senior citizens claiming he was from a charity supporting families of 9/11 victims. He then used the information to steal from the senior citizens.