

## CAN HIGH MORAL PURPOSES UNDERMINE SCIENTIFIC INTEGRITY?

Lee Jussim, Jarret T. Crawford, Sean T. Stevens,  
Stephanie M. Anglin, and Jose L. Duarte

*In this chapter, we review basic processes by which moral purposes can sometimes motivate immoral behavior, and then suggest how moral agendas can sometimes lead social psychology astray through an array of questionable interpretive practices (QIPs). These practices can be used to advance a moral agenda by permitting researchers to interpret the data as supporting that agenda even when it does not. The QIPs reviewed here include: blind spots (overlooking or ignoring data inconsistent with one's moral agenda), selective preference (accepting research supporting one's agenda at face value, but subjecting opposing research of comparable or greater quality to withering criticism), and phantom facts (making declarations or drawing implications without evidence). Four major areas of social psychological research—sex differences, stereotype threat, attitudes toward climate science, and the ideology-prejudice relationship—are reviewed and shown to be characterized by unjustified conclusions plausibly reflecting high moral purposes. The chapter concludes with a discussion of how to reduce QIPs in research that has moral undertones.*

Social psychological research and historical events have repeatedly shown that moral righteousness can and has led to immoral behavior in many contexts outside of science. This chapter suggests that social psychological research itself, when driven by high moral concerns, can and often does produce distorted and invalid claims. First we discuss historical and social psychological evidence that high moral purposes can lead to immoral behavior. Next we briefly review recently identified statistical and methodological threats to the validity of social psychological research. We then introduce a set of heretofore unrecognized threats involving distortions to the *claims* made on the basis of scientific research—what we term

questionable interpretive practices (QIPs). When researchers have high moral purposes, they may be enticed into using QIPs to advance their moral goals at the expense of making scientifically true claims. This chapter reviews the psychological processes by which this can happen, applies them to scientists' own behavior, shows how such processes have likely led to distortions in several areas of social psychological research, and identifies ways to limit such distortions.

### The Moral Irony of High Moral Purposes: Can High Moral Purposes "Justify" Immoral Behavior?

The answer from history is a clear "yes." History is littered with harm committed in the name of high moral purposes. Destruction or subjugation of indigenous peoples in the Americas, Australia, and Asia was believed to constitute the "advancement" of civilization. The Spanish Inquisition was conducted by the Catholic Church in the name of God. The abuses of McCarthyism were conducted to fight communist tyranny. Terrorism, suicide bombings, and the mass murder of civilians are typically conducted by those who deeply believe their cause is righteous (Baumeister, 2012). Certainly, not all who act with moral justifications commit atrocities. Nonetheless, the historical evidence indicates that believing one's cause is righteous can and has been used to justify discrimination, oppression, and violence.

Recent social psychological research on the nature of morality and the psychology of ideology is consistent with these historical observations, and may help explain how moral justifications can and have been used to perpetrate moral transgressions (see, e.g., Graham & Haidt, 2012; Skitka & Morgan, 2014). Consider the notion of "moral licensing": people who have committed a good deed feel license to behave immorally afterward (e.g., Conway & Peetz, 2012). Morality includes codes of conduct that guides people's actions (Forgas, Jussim, & Van Lange, this volume). The purpose of such codes is to produce "good" behavior, but good can be so subjective that one person's good behavior is another's bad behavior (e.g., protecting the sanctity of marriage by forbidding gay couples to marry). Sacredness refers to "the human tendency to invest people, places, times, and ideas with importance far beyond the utility they possess" (Graham & Haidt, 2012, p. 14). Holding something sacred can provide justification for immoral behavior.

For instance, Graham and Haidt (2012) presented a qualitative interpretation of *The Tinner Diaries*, a fictional work that is considered a "bible of the racist right" by the Southern Poverty Law Center (Jackson, 2004). It depicts the overthrow of the US government by an Aryan movement to restore White supremacy, which later leads to a race war. This analysis revealed that loyalty to and self-sacrifice for the Aryan rebellion, as well as the purity of the White race, were treated as moral ideals. Graham and Haidt (2012) also analyzed press releases from the Weather Underground, a militant left-wing group that engaged in terrorist attacks against

the US government. They concluded that the Weather Underground sacralized non-White populations, the poor, and the oppressed as victims of evil White capitalist America, which required extermination. This narrative drew primarily on the moral concerns associated with the provision of care and the prevention of harm. Thus, the laudable moral ideals of providing care for and preventing harm to the disadvantaged was used to justify actions that harmed others (e.g., bombing police stations and other government-affiliated buildings).

Political ideology is often anchored in moral intuitions that can help "bind" people into moral communities possessing shared values, worldviews (Graham & Haidt, 2012; Haidt, 2012) and moral social identities (Parker & Janoff-Bulman, 2013). Moral convictions, which reflect a strong and absolute belief that something is right or wrong without the need for proof or evidence, have a dark side (see Skitka, 2010; Skitka & Morgan, 2014). When people strongly identify with such groups, they are routinely intolerant of those with different values and worldviews (e.g., Chambers, Schlenker, & Collisson, 2013; Crawford & Pilanski, 2014; Wetherell, Brandt, & Reyna, 2013; for reviews, see Brandt, Wetherell, & Crawford, this volume; Brandt, Reyna, Chambers, Crawford, & Wetherell, 2014). For instance, self-identified American liberals and conservatives display equal willingness to deprive each other of constitutionally protected rights (Crawford, 2014; Crawford & Pilanski, 2014).

The fact that laypeople can and at times do use moral imperatives to justify immoral acts raises the possibility that they may also, at least sometimes, obstruct scientists' focus on "getting it right." How might this manifest? One possibility is that researchers may sometimes allow their moral purposes to influence how they conduct and interpret their research. The most extreme possibilities are that high moral purposes may distort or even invalidate their research. To understand how, it is necessary to understand (1) what we mean by scientific integrity and (2) common threats to scientific integrity.

### Can High Moral Purposes Undermine the Integrity of Social Psychology?

#### What Is Scientific Integrity?

"Scientific integrity" refers to two related but separate ideas: (1) the personal honesty of individual scientists in the conduct and reporting of their research; and (2) developing robust bodies of conclusions that are valid and unimpaired by errors and biases. Even when researchers suffer no lack of personal integrity, conventional practices common in their field may produce findings that are misleading or invalid. Nonetheless, "getting it right" is the sine qua non of science (Funder et al., 2013). Science can tolerate individual mistakes and flawed theories, but only if it has reliable mechanisms for efficient correction. This is perhaps most

### Confirmation Bias Among Social Scientists

Scientists are not immune to confirmation bias (Ioannidis, 2012; Lilienfeld, 2010). Reviewers' theoretical (Epstein, 2004; Mahoney, 1977) and ideological (Abramowitz, Gomes, & Abramowitz, 1975) views influence their evaluation of research reports. Values influence each phase of the research process, including how people interpret research findings (Duarte et al., 2015). Although scientists strive to be objective, confirmation biases have considerable potential to influence study design and the evaluation and interpretation of evidence.

### Questionable Interpretive Practices as an "Under the Radar" Threat to the Integrity of Social Psychological Science

Even if methods and statistics are impeccable, it is *still* possible to reach distorted conclusions through questionable interpretive practices (QIPs)—conceptual and narrative tools for reaching desired conclusions, even when data are inconsistent with those conclusions (Jussim et al., in press). QIPs are a mechanism by which researcher confirmation biases (e.g., Lilienfeld, 2010) can distort conclusions, even with untainted data.

Although there are many QIPs (Jussim et al., in press; see Brandt & Proulx, 2015, for a discussion of the related phenomenon of QTIPs, or questionable theoretical and interpretive practices), this chapter focuses on:

- *blind spots*: overlooking data and studies that conflict with one's preferred conclusions.
- *selective preference*: highlighting studies consistent with one's preferred conclusions and downplaying, criticizing, or dismissing equally high quality studies inconsistent with those conclusions.
- *phantom facts*: declaring something true, or making an implicit assumption that something is true, without providing empirical evidence.

We next review four areas of research in which a plausible case can be made that high moral purposes have led to scientific distortions.

### High Moral Purpose I: Combating "Isms"

Social psychology has a rich history of advocacy research designed to combat inaccurate stereotypes, unjustified prejudices, and many types of discrimination. These are undoubtedly high moral aims—and many social psychologists may take "making a difference" as a high moral purpose (Unger, 2011). Furthermore, many valuable social scientific insights have emerged from the long-standing efforts of social psychologists toward understanding these enduring social problems.

obvious in engineering (where, for example, a poorly designed product may not work or, in some cases, be dangerous) and medical interventions (some of which make people less rather than more healthy). However, even within psychology, large amounts of time, effort, and resources can and have been wasted chasing phenomena that only appeared to be real (e.g., fewer options increases satisfaction with one's choice; thinking about Florida leads people to walk slowly) through a combination of p-hacking and publication biases (e.g., Lakens, 2015; Simonsohn, Nelson, & Simmons, 2014).

### Known Threats to the Integrity of Social Psychology Statistical and Methodological Threats

Questionable research practices, failures to replicate, faulty statistical practices, lack of transparency, publication biases, and political biases all potentially threaten scientific integrity (Begley & Ellis, 2012; Cummings, 2013; Ioannidis, 2005; Jussim, Crawford, Stevens, & Anglin, in press; Simmons, Nelson, & Simonsohn, 2011). Individuals and organizations have begun addressing scientific integrity failures through reforms primarily targeting transparency, statistics, and methods.

Nonetheless, there is a set of practices unrelated to methods or statistics that has not received much attention in ongoing scientific integrity discussions, but which can similarly threaten the validity of scientific conclusions: questionable interpretive practices. Because these practices involve subjective decisions about which there may often be no clearly right or wrong methodological or statistical guidelines, they are easily triggered by morally righteous motivations. The rest of this chapter addresses how these subjective judgments can threaten scientific integrity.

### Confirmation Bias Among Laypeople

People's motivations can influence their reasoning. Motivated reasoning refers to biased information processing that is driven by goals unrelated to accurate belief formation (Kahan, 2011; Kunda, 1990). A specific type of motivated reasoning, confirmation bias, occurs when people seek out and evaluate information in ways that confirm their preexisting views while downplaying, ignoring, or discrediting information opposing their views (Nickerson, 1998; also referred to as "myside bias," see Stanovich, West, & Toplak, 2013). People intensely scrutinize counterattitudinal evidence while easily accepting information supporting their views (e.g., Ditto & Lopez, 1992; Lord, Ross, & Lepper, 1979). Although these processes are affectively driven (e.g., Jaks & Devine, 2000; Munro & Ditto, 1997; Zuwerink & Devine, 1996), people generate convincing arguments to justify their automatic evaluations, producing an illusion of objectivity (Haidt, 2001; Nickerson, 1998).

Conducting research to reduce oppression, however, is exactly the type of worthy moral purpose that risks compromising the scientific integrity of social psychology. Morally motivated scientists may suborn their subtle and complex theories to generate false or misleading claims that advance their agenda. This vulnerability was aptly captured by the recently proposed Paranoid Egalitarian Meliorism (PEM) model (Winegard, Winegard, & Geary, 2015):

We do not mean paranoid pejoratively; rather we mean it as a form of error-management (Haselton & Nettle, 2006). In this view, paranoid refers to a heightened sensitivity to perceive injustices and/or threats to equality. Because of this, many social psychologists: (1) study topics that are related to perceived injustices (stereotyping, prejudice, hierarchies, immorality of the wealthy, obedience); (2) ignore topics that are perceived to threaten egalitarianism (heritability, stereotype accuracy, possible benefits of conformity/hierarchy); and (3) become hostile/biased against research suggesting that some outcome differences among individuals and/or groups are at least partially caused by differences in personal traits rather than by discrimination or social oppression (e.g., that sex differences in STEM field representation are partially caused by cognitive differences and the different occupational preferences of men and women). At its most extreme, PEM can lead to the creation of “victim groups” who become quarantined from objective scientific analysis. Protection of such perceived victim groups becomes a sacred value (Tetlock, 2003), and those who are perceived as violating this sacred value are assailed. Biased reviews, funding, and hiring decisions are justified because they are means to protecting a sacred cause.

Although the term “paranoid” may be too strong, the phenomena they described strongly implicate high moral purposes as threats to scientific integrity. Indeed, it is even possible that some scientists view research that exposes and combats the evils of oppression, discrimination, and injustice as a “moral opportunity” (Miller & Monin, this volume) to advance the impression that they are good, decent egalitarians. The following examples strongly suggest that the PEM can give some insight into how social psychologists sometimes get their science wrong and, therefore, inform current and future perspectives about how to get it right.

### *The Science and Politics of Minimizing Gender Differences*

The systematic study of gender similarities and differences became a prominent area of research in psychology in the 1970s, coinciding with the modern feminist movement in the US. Many psychologists studied sex differences from a feminist perspective, seeking to demonstrate that gender differences are small to

nonexistent in order to promote gender equality and advance women's status in society (Eagly, 1995). Although the goal of improving women's access to equal opportunity was undeniably moral, the evidence for trivial or nonexistent gender differences was decidedly mixed. This constituted a classic case of a series of QIPs, whereby research that contested claims of gender equivalence was consistently overlooked, dismissed, or downplayed.

Starting with Maccoby and Jacklin's (1974) classic review downplaying the size and importance of many gender differences, many feminist and social psychological scholars, with the manifest and worthy agenda of breaking down barriers to women, followed suit. In “The Science and Politics of Comparing Women and Men,” an early classic article showing how worthy moral goals (in this case, gender equality) can obstruct scientific progress, Eagly (1995, p. 149) described this state of affairs as follows:

To the extent that the “gender-neutral” strategy of making no distinctions between women and men leads to gender equality (see Bem, 1993), scientific research showing that women are not different from men should help ensure women equal access to a variety of roles from which they were excluded. In contrast, evidence of differences might be seen as disqualifying women in relation to certain roles and opportunities and as justifying unequal treatment under the law.

As a result, research showing substantial gender differences was intensely scrutinized and challenged. As Eagly (1995, p. 149) continues:

Feminist writing of the 1970s generally portrayed research purportedly demonstrating sex differences as largely prescientific and obviously faulty, or, if the research was more modern, as riddled with artifacts, methodological deficiencies, and unexplained inconsistencies in findings (e.g., Sherif, 1979).

One possibility is that these criticisms were scientifically well justified and that subsequent research would eliminate the artifacts and deficiencies and then demonstrate the validity of perspectives downplaying gender differences. However, another possibility is that this reflects the QIP of *selective preference*: perhaps these scholars engaged in confirmation biases, aggressively seeking flaws in research they disliked (demonstrating gender differences) and not applying the same critical standards to research they liked (demonstrating little or no gender differences). As noted earlier, people (including scientists; Abramowitz et al., 1975; Ceci, Peters, & Plotkin, 1985) often work harder to discredit disliked than liked findings. However, in a manner entirely consistent with Haidt's (2001) intuitionist model of morality, they deploy their logical reasoning skills to make it appear as

if their reasoning explains their evaluation, when in fact it was their evaluation that triggered their reasoning. This raises the possibility that decades of declaring research demonstrating gender differences as trivial and fatally flawed reflected the motivated biases and QIPs of the researchers, not the quality of the research itself.

So, which was it? By the late 1990s, the empirical answer was vividly clear: The size of gender differences varies a great deal across different characteristics. Some are trivially small. But many are at least moderate, and some are quite large. Several meta-analyses show at least moderate gender differences for characteristics such as restlessness, math test scores, helping, leadership, a slew of nonverbal skills and characteristics, and many cognitive/academic characteristics among children (e.g., Briton & Hall, 1995; Halpern, Straight, & Stephenson, 2011; Swim, 1994). To be sure, men and women have many characteristics on which they, in fact, do not greatly differ (Hyde, 2005); however, even Hyde acknowledges that differences are quite large in certain physical skills, sexuality, and aggression (and see Eagly, 2013, for an analysis that suggests Hyde understates the size and power of gender differences).

Unfortunately, the outdated consensus on the ostensibly trivial size of sex differences continued to appear in textbooks and reviews, independent of the growing evidence to the contrary, and sometimes could be found even post-Eagly's (1995) exposé of these dysfunctions in the scientific literature (see Eagly, 2013, for a review). As Eagly (1995, 2013) has pointed out, achieving equal opportunities for women is an undeniably moral goal. However, the clear record of so many scientists allowing their benevolent moral intentions to distort their scientific conclusions constitutes one of the clearest records of good morals going bad in ways that obstructed the advancement of valid scientific conclusions.

### **But for Stereotype Threat, White and African-American Standardized Test Scores Would (Not) Be Equal**

Alongside fighting sexism, fighting racism has long been one of many social psychologists' high moral purposes. The fight for civil rights and equality of opportunity is undoubtedly a high moral good. The question addressed by the present chapter, however, is whether this high moral purpose has sometimes led to distorted scientific conclusions. There are many reasons to believe that it has, and here we focus on one (see Jussim, 2012, and Jussim et al., in press, for others): the erroneous and misleading interpretations of stereotype threat research.

The difference between White and African-American academic achievement is one of the great social problems plaguing the US. One manifestation of this problem is the very large gap in standardized achievement test scores, averaging about a full standard deviation, or the equivalent of 100 SAT points on each test (e.g., Neisser et al., 1996). A simple situational tweak that could eliminate these very large differences would have dramatic implications. It would dispel biological

explanations, suggest that race differences have social causes, and reduce obstacles to racial equality.

Psychologists once claimed that they had found just such a magic bullet: stereotype threat. In a classic early study, Steele and Aronson (1995) examined the performance of African-American and White students under various conditions designed to lead African-Americans to become concerned about confirming stereotypes of African-American inferiority (the stereotype threat conditions) or designed to eliminate such threats. The results, which are reproduced here in Figure 11.1, appeared to be striking. Under threat, there seemed to appear the typical, dramatic race differences in performance. But under no threat conditions, there seemed to be no race difference.

The Figure 11.1 caption statement is technically incorrect (they are covariate adjusted means, not "mean test performance" scores, thereby rendering the figure deeply misleading). The nearly equal covariate adjusted means in the nondiagnostic condition *do not mean* that Blacks and Whites had equal scores. Instead, they mean that the *preexisting differences* (of about 40 points) were maintained in the nondiagnostic condition. Stereotype threat increased achievement test differences; removing it did not eliminate the mean differences between African-Americans and Whites.

It is clear that stereotype threat researchers once routinely presented their findings in this manner (see Sackett, Hardison, & Cullen, 2004, for a review). Referring to Steele and Aronson (1995), Aronson et al. (1999, p. 30) claimed that African-American students performed "about as well as Whites when the same test was presented as a nonevaluative problem solving task." Wolfe and Spencer

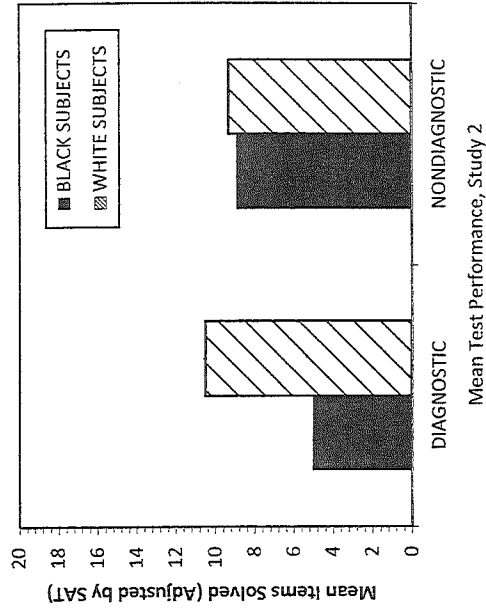


FIGURE 11.1 Based on Figure 2 from Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811, page 802.

(1996, p. 180) declared that “one simple adjustment to the situation (changing the description of the test) eliminated the performance differences between Whites and African-Americans.”

As the American Psychological Association (2006, n.p.) puts it on its web page:

In the no stereotype-threat condition, in which the exact same test was described as a lab task that did not indicate ability, Blacks' performance rose to match that of equally skilled Whites. Additional experiments that minimized the stereotype threat endemic to standardized tests also resulted in equal performance.

And then later on the same page: “At the very least, the findings undercut the tendency to lay the blame on unsupported genetic and cultural factors, such as whether African Americans' 'value' education or girls can't do math” (American Psychological Association, 2006, n.p.).

This latter sentence demonstrates the moral purpose driving some of the promotion of stereotype threat research. The problem, however, is not the moral posturing. It is the unjustified *scientific* claim. The original stereotype threat research *never* showed that removing threat eliminates race differences in standardized test scores. This is a *phantom fact*. How can that possibly be, given the results displayed in Figure 11.1?

### The Confusing (Mis)presentation

The widespread misinterpretation probably derived from two sources: a confusing presentation of the original result, and the failure to correct this misleading presentation by subsequent stereotype threat researchers. Although Steele and Aronson's (1995) text clearly states that they performed an *analysis of covariance* (ANCOVA) and reported *adjusted means*, their Figure 2 presents the *covariate adjusted means* labeled only as “Mean test performance Study 2” (p. 802; see Figure 11.1). The figure shows that the (adjusted) mean performance of the African-American students was equal to the (adjusted) mean performance of the White students.

Misinterpreting this study is easy absent a close reading of the text and a sophisticated understanding of adjusted means in ANCOVA. Because the adjusted means for African-American and White students were nearly identical in the no threat (nondiagnostic test) condition, it is easy to come away with the false impression that removing stereotype threat eliminated racial differences. Equal adjusted means in ANCOVA occur because preexisting differences are unaffected by the manipulation, not because the means are equal. The equal adjusted means indicate that preexisting differences (52 SAT points, in their sample) were *maintained* (not eliminated) when not under stereotype threat.

### Misleading Presentations 2.0

This situation changed, slightly, after Sackett et al. (2004) pointed all this out. In a reply that primarily defended the validity of stereotype threat research, Steele and Aronson (2004) acknowledged that “in fact, without this [covariate] adjustment, they would be shown to perform still worse than Whites” (p. 48). They claimed that ANCOVA was conducted in order to reduce error variance. This is a valid use of ANCOVA, and one could view this exchange as a nice example of science (eventually) self-correcting. Why it took so long, and did not occur until after Sackett et al.'s (2004) critique, remains unclear.

One possibility is that the high moral purpose of combating racism impelled stereotype threat researchers to promote the idea that there were no “real” race differences in academic test score performance. Indeed, if the priority is to “get the science right,” one might expect stereotype threat researchers to describe Steele and Aronson's (1995) results simply and accurately subsequent to Sackett et al.'s (2004) critique. This can be done in 13 words: “Stereotype threat *increased* the achievement gap; removing threat left the prior gap intact.”

Alternatively, however, modern stereotype threat researchers might go to considerable lengths to retain the original claim of racial equivalence under non-threatening conditions in their earnest pursuit of egalitarianism. Perhaps this explains why modern stereotype threat researchers *still* promote a technically true but misleading claim. For example, Schmader, Johns, and Forbes (2008, p. 336) claimed that Steele and Aronson (1995) showed that:

African American college students performed worse than their White peers on standardized test questions when this task was described to them as being diagnostic of their verbal ability but that their performance was equivalent to that of their White peers when the same questions were simply framed as an exercise in problem solving (and after accounting for prior SAT scores).

Similarly, Walton, Spencer, and Erman (2013, p. 5) wrote:

In a classic series of studies, Black students performed worse than White students on a GRE test described as evaluative of verbal ability, an arena in which Blacks are negatively stereotyped. But when the same test was described as nonevaluative—rendering the stereotype irrelevant—Blacks performed as well as Whites (controlling for SAT scores; Steele & Aronson, 1995).

These latter statements take 50–60 words, whereas the true result of the study can be summarized in 13. These latter statements must be convoluted in order to (1) be technically true and (2) maintain the claim that “remove stereotype threat, and African-American/White standardized test score differences disappear.” To be

true, the declaration that African-American and White scores are “equivalent” in nonthreatening conditions needs to be walked back by adding the parenthetical regarding controlling for prior SAT scores. The actual result—preexisting differences were maintained under nonthreatening conditions—is never explicitly stated in these articles.

After Sackett et al.’s (2004) exposé, stereotype threat researchers have a special responsibility to insure that their statements are accurate and not likely to be misinterpreted. The persistence of misleading statements strongly suggests that insuring that readers understand that removing stereotype threat did not significantly reduce the achievement gap in Steele and Aronson’s (1995) research has taken a backseat to the high moral purpose of combating racism. Instead, despite Sackett et al.’s (2004) best attempt to bury the *phantom fact* that “remove threat, and the gap disappeared,” the undead claim (with the newer addition of the parenthetical but easily missed and misunderstood caveat, “controlling for prior scores,” which makes it deeply misleading rather than false) continues to appear in some of psychology’s most influential outlets.

### High Moral Purpose II: Deploying Science to Crush Our Political Opponents

People with strong ideological views are often intolerant of those who hold different views (Brandt et al., 2014; Brandt et al., this volume). People view opposing partisans as more extreme than they really are, as immoral, as holding illogical beliefs, and as not deserving the same rights and protections as other people (e.g., Chambers, Baron, & Inman, 2006; Chambers & Melnyk, 2006; Crawford, 2014; Crawford & Pilanski, 2014; Graham, Nosek, & Haidt, 2012).

What does this have to do with high moral purposes undermining scientific integrity? There is no reason to think social psychologists are personally immune from these attitudes, or that the scientific method offers immunity from the biases such attitudes often produce (Duarte et al., 2015; Jussim et al., 2015, in press). If scientists believe that it is their moral obligation to marginalize their ignorant and immoral ideological opponents, they put themselves at risk for purveying invalid scientific claims. Because strongly held ideological beliefs subjectively feel like objective truths (Morgan, Skitka, & Lytle, 2014), it is possible that such scientists are unaware of the biased nature of their science; squashing their ideological opponents may be subjectively experienced as a core component of advancing science.

When there is abundant objective evidence that some widely held belief is false, scientists are justified for challenging such beliefs. However, we are not confident that scientists with high moral purposes can always distinguish between, on one hand, overwhelming objective evidence and, on the other, promoting their own personal moral and political agendas, even in the absence of overwhelming

objective evidence. Next, we next present two examples of how attempts to vilify one’s ideological opponents has distorted science.

### The Curious Case of Condemning Climate Skeptics as Conspiracy Theorists

Global warming may be one of the greatest social and scientific problems of our era. The potential disruption produced by melting polar ice, rising seas, expanding deserts, and increased extreme weather outbreaks is vast, and the evidence is overwhelming that humans have either created or exacerbated the pace with which warming has occurred (United Nations, 2014). Nonetheless, it is very difficult to get people, organizations, and especially governments to do anything to address the problem.

Further compounding the problem are active efforts to thwart major policy changes by challenging the scientific basis for evidence of human-caused global warming. Thus, to some, fighting the “deniers” of global warming may have taken on a high moral purpose.

Into this mix stepped Lewandowsky et al. (2013) with a paper titled, “NASA Faked the Moon Landing—Therefore (Climate) Science Is a Hoax”—which strongly implies that people who doubt global warming believe bizarre conspiracy theories. As Lewandowsky et al. (2013, p. 622) put it, “conspiratorial thinking contributes to the rejection of science.”

One possibility is that this was true—that a disproportionately high number of people who disbelieve climate science also believe in something as silly as the faking of the moon landing. Another, however, was that this was essentially trumped up in order to cast those who are most skeptical of the climate science as fools. Fortunately, and to their credit, Lewandowsky et al. (2013) publicly posted their data, so we can evaluate these two alternative explanations for the claim in the title.

Their evidence for these conclusions was drawn from 1,145 readers of environmentalist blogs who completed a web survey asking about their belief in conspiracies and acceptance of scientific conclusions (HIV causes AIDS, burning fossil fuels increases atmospheric temperatures, etc.). Lewandowsky et al. (2013) subjected responses to latent variable modeling and did indeed find that “conspiracist ideation” negatively predicted (–.21, standardized regression coefficient) acceptance of climate science. So, where is the problem?

The implication that climate skeptics believe in the faking of the moon landing is another phantom fact. Out of over 1,145 respondents, there was a grand total of 10 who believed the moon landing was faked. Among the 134 of participants who “rejected climate science,” only *three* people (2%) endorsed the moon-landing hoax. The link asserted in the title of the paper did not exist in the sample. Correlations primarily resulted from covariance in levels of agreement among



reasonable positions (i.e., people varied in how much they *disbelieved* hoaxes and in how strongly they *accepted* science). It would be fair to characterize their results as indicating “the more strongly people disbelieved hoaxes, the more strongly they believed in climate science”—people varied in *how strongly* they rejected hoaxes and accepted science, but almost no one believed the moon hoax.

Understanding when people are and are not persuaded by science is an interesting and important area of research. But this curious case highlights the threat to scientific integrity that can stem from high moral missions. The notion that skeptics believed something so silly as the faking of the moon landing is yet another myth essentially concocted by the researchers. No matter how worthy the efforts to advance policy changes to combat human sources of global warming, the goal of “getting it right” is jeopardized when scientists claim their data shows their ideological opponents hold silly beliefs when they, in fact, do not. As such, this constitutes another example of high moral purposes undermining scientific integrity.

### ***The Unjustified Claim That Prejudice Is a Province of the Right***

Several theoretical perspectives disproportionately blame conservatives for prejudice, discrimination, and political intolerance (e.g., Jost, Glaser, Kruglanski, & Sulloway, 2003; Lindner & Nosek, 2009; Sibley & Duckitt, 2008). We next consider two nonmutually exclusive possible explanations for this pattern: (1) conservatives really do hold more prejudices and are more intolerant than are those on the left; and (2) this pattern reflects the biases and moral agendas of the researchers studying the relationship between ideology and prejudice.

Conservatives are indeed prejudiced against many groups, including racial and ethnic minorities, women, and homosexuals (e.g., Altemeyer, 1996; Jost et al., 2003). However, is prejudice restricted to oppressed or disadvantaged target groups? One might reasonably consider prejudice against corporate executives or wealthy White men to be irrelevant to advancing egalitarian *political and moral* goals, but does that render such prejudice impossible, uninteresting, nonexistent, or psychologically inconsequential? The answer can be yes only if one limits the scientific purpose of social psychology to the political/moral action agenda of advancing the status of disadvantaged groups.

If prejudice simply refers to negative attitudes toward groups and is not restricted to oppressed demographic groups, it becomes clear that *any* group can be the target of prejudice. Considering only prejudices against the subset of demographic groups who are disadvantaged does not fully answer the question, “Is prejudice the particular province of the right?”

There is another problem with limiting prejudice to the study of such demographic groups. These are the groups that are of highest concern to people on the left, and in fact, most such groups are themselves left-wing or left-aligned groups

(e.g., African-Americans; gay men and lesbians; see Brandt et al., 2014, for a discussion). Restricting prejudice to such groups constituted a large *blind spot* in the “psychology of prejudice” induced by the overwhelmingly left-wing worldviews of social psychologists (see Duarte et al., 2015, for a review), which until recently rarely addressed prejudice against right-wing or right-aligned groups.

If conservatives are *generally* more prejudiced than liberals, then they should also show more prejudice against right-wing groups. This, however, has not happened. Research supporting the *ideological conflict hypothesis* (ICH; see Brandt et al., 2014, for a review) shows that conservatives are prejudiced against left-aligned groups (e.g., atheists, welfare recipients) to about the same degree as liberals are prejudiced against right-aligned groups (e.g., evangelical Christians, businesspeople). Such findings have replicated across convenience and nationally representative samples, with a variety of social target groups, and various operationalizations of intergroup antipathy, including prejudice (Chambers et al., 2013; Crawford, 2014; Wetherell et al., 2013), political intolerance (Crawford, 2014; Crawford & Pilanski, 2014), and outright discrimination (i.e., dictator game; Crawford et al., 2014). Thus, despite extant efforts to explain prejudice as a symptom of conservatism, this recent ICH research shows that antipathy toward ideologically dissimilar others characterizes the left and right in approximately equal measure.

Reducing intergroup antipathy such as prejudice and discrimination is a justified moral concern. However, adopting this moral concern has made some social and political psychologists blind to contradictory evidence, and has narrowed the very definitions of prejudice and related constructs in ways that, intentionally or not, unjustifiably advance the notion that conservatives are morally inferior to liberals. The subsequent conclusions typify, perhaps unintentionally, Lilienfeld’s (2015) conclusion that conservatism is often characterized as deficient and in need of explanation, relative to “normal” liberalism. These decisions demonstrate yet again how a particular moral agenda—combating the demographic prejudices of conservatives—has distorted scientific practice and constituted an obstacle to advancing our understanding of prejudice.

### ***Recommendations for Limiting the Distorting Effects of High Moral Missions***

In this chapter, we have identified how moral purposes can lead to QIPs that undermine the validity of social psychology. Although it may be difficult, we believe that social psychologists, in their roles as authors, reviewers, and editors, can engage in practices that avoid, reduce, and defuse QIPs. A few such practices are discussed here (see Duarte et al., 2015 and Jussim, Crawford, Anglin, & Stevens, 2015, and Jussim, Crawford, Stevens, & Anglin, in press, for additional recommendations).



### Acknowledging Competing Perspectives

Clark and Hatfield (1989) provide a useful template for how to interpret results in a balanced manner that might otherwise have been exploited to advance a particular moral agenda. Instead of relying on a pet theory to interpret their results (showing huge differences between men and women in willingness to accept an offer from a stranger to have casual sex), they offered several potential explanations (some of which were ideologically contradictory) rather than excluding disfavored over favored options. Such an approach is not only more scientifically honest, but helps advance scholarship by providing authors multiple alternative testable hypotheses for future research.

### Adversarial Collaboration and Popperian Falsification

We echo the Popperian imperative that scientific practices are often at their best when researchers, in their own work and in reviewing that of others, attempt to falsify rather than confirm their pet hypotheses. This approach can help researchers identify how their own biases can be scrutinized, which can reduce morally motivated blind spots and selective preferences.

One way to maximize falsification attempts is through adversarial collaborations (Kahneman, 2003; Van Lange, 2013). There are challenges to such an approach, and projects may break down over disagreements prior to data collection. Nonetheless, if researchers can get past their personal (including moral) commitments to particular outcomes, there are both personal and scientific advantages to such collaborations. Personally, such collaborations are likely to advance the careers of all involved by yielding highly publishable findings. Scientifically, adversaries will likely be highly motivated to disconfirm one another's theories, thereby stacking the scientific deck in favor of Popperian falsification. Consequently, such collaborations are likely to (and have already) constructively advance(d) the field by resolving scientific controversies (Crawford, Collins, & Brandt, 2015; Silberzahn, Simonsohn, & Uhlmann, 2014).

### Meta-Analytic Thinking

A single study cannot resolve a question of human thought or behavior. By focusing on single studies, we have considerable freedom to allow our moral concerns to bias just what evidence we see as informative. It is very easy to cherry-pick studies and results in narrative reviews in such a manner as to create the impression that there is widespread support for our preferred claims, even when there is not (as indicated by the research one has, in classic *blind-spot* manner, intentionally or unintentionally overlooked). Instead, whether or not we conduct actual meta-analyses, we should get into the habit of compiling, citing, including, and

considering broad swaths of evidence (whether individually or collectively; see Tsuji, Bergmann, & Cristia, 2014) pertinent to our research question, regardless of whether it supports or contests our pet hypotheses. Efforts to create checklists that can encourage a balanced and meta-analytic approach (see Washburn, Morgan, & Skitka, 2015, for such an example) can reduce morally motivated biases at multiple stages of the research process.

### Ways Forward

We generally oppose adding onerous bureaucratic requirements to the already difficult research process (see also Fiedler, this volume). However, one method social psychologists can use to limit their potential for motivated reasoning and confirmation biases is to ask themselves a few pointed questions. In that spirit, we conclude our chapter with this *Personal Use Checklist*, which we envision as something for personal use, not to be required by journals or other organizations. Furthermore, just because this version of the checklist works for us does not mean it will necessarily work for others. We encourage researchers to adapt this as they see appropriate, or develop their own. This is intended to assist well-meaning researchers to become more aware of their own potential for biases, in order to be more able to limit or eliminate them, *without mandates from authorities, editors, or organizations*.

The following *Personal Use Checklist* is for increasing confidence that our empirical research is relatively free of motivated biases:

1. What do I want to happen and why? \* An honest and explicit self-assessment is a good first step toward recognizing our own tendencies toward bias, and is, therefore, a first step to building in checks and balances in our research to reduce them.
2. Do I have a long track record of research that does not systematically validate a particular political or social justice narrative or agenda? Note this is not about one's *intentions*. It is about one's *results*. If one's results consistently validate a particular set of beliefs, values, or ideology, one has failed this check, and this suggests attempts at falsification may be in order.
3. Have I generated theoretical arguments for *competing and alternative hypotheses* and designed studies to incorporate and test them? \* Honest tests of alternatives can go a long way to reducing personal bias.
4. Have I carefully edited my manuscript to eliminate the use of pejorative terms to refer to the psychological characteristics of types of people I dislike or disagree with?
5. Have I read some of the literature highlighting the invidious ways our motivated biases, morals, and politics can creep into our scientific scholarship? Doing so can alert one to ways in which our politics and values might distort

our science. After having done so, have I made a good faith attempt to eliminate such biases from my scholarship?

6. Have I sought feedback from colleagues with very different theoretical and political views than mine or with track records of scholarship that often contests my preferred narratives?

Starred (\*) items are from the checklist first developed by Washburn, Morgan, and Skitka (2015).

It may not always be possible for researchers to meet all six of these checks. However, as a starting heuristic, meeting five of the six probably justifies confidence that the research has kept bias mostly in check. What to do if one cannot meet at least five of the six (or, alternatively, one fails too many of one's own such questions)? Although that, too, is a matter of judgment, one possibility will be to *start over*. The first check may appear to be an open-ended question that one can neither pass nor fail. However, if one has strong preferences for how a study "should" come out, then one's ego may be invested in the outcome and one has failed this check. Checks 2 and 3 are easy enough to conduct, though *implementing* Check 3 after one has realized one has not met it may require new research. Check 4 is no more difficult than eliminating sexist language from manuscripts. Check 5 requires a little reading and probably can do double duty as a required assignment in advanced undergraduate and graduate courses on methodology, social cognition (confirmation bias among scientists!), and scientific practices. The hardest part about Check 6 is finding enough people so that the ones from whom one seeks feedback are not overburdened.

## Conclusion

In this chapter, we have reviewed evidence that people sometimes engage in immoral actions when they are motivated by high moral principles. One possible manifestation of this sort of moral irony is that when researchers are motivated by high moral principles, such as combating global warming or advancing egalitarianism, such motivations may lead to practices that threaten scientific integrity. This can occur because scientists are subject to the same sorts of motivated reasoning and tendencies toward confirmation bias as are laypeople. When they are highly motivated by moral, political, or social action agendas, therefore, they may be at heightened risk of engaging in questionable interpretive practices—discursive techniques for reaching desired conclusions even when the data do not support those conclusions. In this chapter, we have reviewed how blind spots (overlooking data and studies that conflict with one's preferred conclusions), selective preference (highlighting studies consistent with one's preferred conclusions and downplaying, criticizing, or dismissing equally high-quality studies inconsistent with those conclusions), and phantom facts (declaring something true, or making an

implicit assumption that something is true, without providing empirical evidence) have led to underestimating the size of gender differences, claims that removing stereotype threat eliminated racial differences in standardized test scores when it did not do so, and claims that climate skeptics believe bizarre conspiracy theories that, in fact, they do not believe.

Such researcher biases are undoubtedly well intentioned (or, at least, we are sympathetic to the moral values of egalitarianism and to the importance of combating global warming). However well intentioned they are, such biases are also generally subtle and invidious, and, as such, many researchers may be unaware of the potentially unintentional ways in which they can lead to unjustified conclusions. Thus we ended our chapter by reviewing specific steps researchers can take to limit their vulnerability to such biases. Our hope is that by making such processes salient, researchers will be more willing and able to prevent them, and thereby lead to a more credible and valid psychological science.

## References

- Abramowitz, S.I., Gomes, B., & Abramowitz, C.V. (1975). Publish or perish: Referee bias in manuscript review. *Journal of Applied Social Psychology, 5*, 187–200.
- Altemeyer, R.A. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.
- American Psychological Association. (2006). Stereotype threat widens achievement gap. Retrieved December 28, 2014, from <http://www.apa.org/research/action/stereotype.aspx>
- Aronson, J., Lustina, M.J., Good, C., Keough, K., Steele, C.M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29–46.
- Baumeister, R.F. (2012). Human evil: The myth of pure evil and the true causes of violence. In M. Mikulincer & P.R. Shaver (Eds.), *The social psychology of morality: Exploring causes of good and evil* (pp. 367–380). Washington, DC: American Psychological Association.
- Begley, C.G., & Ellis, L.M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature, 483*, 531–533.
- Brandt, M.J., & Proutlx, T. (2015). QTIPS: Questionable theoretical and interpretive practices in social psychology: Commentary on Duarte et al. *Behavioral and Brain Sciences, 38*, 19–20.
- Brandt, M.J., Reyna, C., Chambers, J.R., Crawford, J.T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Citizen Directions in Psychological Science, 23*, 27–34.
- Britton, N.J., & Hall, J.A. (1995). Beliefs about female and male nonverbal communication. *Sex Roles, 32*, 79–90.
- Ceci, S.J., Peters, D., & Plotkin, J. (1985). Human subjects review, personal values, and the regulation of social science research. *American Psychologist, 40*, 994–1002.
- Chambers, J.R., Baron, R.S., & Inman, M.L. (2006). Misperceptions in intergroup conflict: Disagreeing about what we disagree about. *Psychological Science, 17*, 38–45.
- Chambers, J.R., & Melnyk, D. (2006). Why do I hate thee? Conflict misperceptions and intergroup mistrust. *Personality and Social Psychology Bulletin, 32*, 1295–1311.

- Chambers, J.R., Schlenker, B.R., Collisson, B. (2013). Ideology and prejudice: The role of value conflicts. *Psychological Science*, *24*, 140–149.
- Clark, R.D., & Hatfield, E. (1989). Gender differences in receptivity to sexual offers. *Journal of Psychology & Human Sexuality*, *2*, 39–55.
- Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, *38*, 907–919.
- Crawford, J.T. (2014). Ideological symmetries and asymmetries in political intolerance and prejudice toward political activist groups. *Journal of Experimental Social Psychology*, *55*, 284–298.
- Crawford, J.T., Brandt, M.J., Chambers, J.R., Inbar, Y., Motyl, M., & Wance, N.M. (2014). *A multi-dimensional approach to political prejudice: Social and economic ideologies differentially predict prejudice across the political spectrum*. Manuscript in preparation.
- Crawford, J.T., Collins, T.P., & Brandt, M.J. (2015). *Ideological symmetry in people's avoidance of dissonance-arousing situations: A failure to closely or conceptually replicate Niimi, Jost and Van Bavel (2013)*. Manuscript in preparation.
- Crawford, J.T., & Pilanski, J.M. (2014). Political intolerance, right and left. *Political Psychology*, *35*, 841–851.
- Cummings, G. (2013). The new statistics why and how. *Psychological Science*, *25*, 7–29.
- Ditto, P.H., & Lopez, D.F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 569–584.
- Duarte, J.L., Crawford, J.T., Stern, C., Haidt, J., Jussim, L., & Tetlock, P. (2015). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*, *38*, 1–13.
- Eagly, A. (1995). The science and politics of comparing women and men. *American Psychologist*, *50*, 145–158.
- Eagly, A.H. (2013). The science and politics of comparing men and women: A reconsideration. In M.K. Ryan & N.R. Branscombe (Eds.), *Sage handbook of gender and psychology* (pp. 11–28). London: Sage.
- Epstein, W.M. (2004). Informational response bias and the quality of the editorial processes among American social work journals. *Research on Social Work Practice*, *14*, 450–458.
- Funder, D.C., Levine, J.M., Mackie, D.M., Morf, C.C., Vazire, S., & West, S.G. (2013). Improving the dependability of research in personality and social psychology: Recommendations for research and educational practice. *Personality and Social Psychology Review*, *18*, 3–12.
- Graham, J., & Haidt, J. (2012). Sacred values and evil adversaries: A moral foundations approach. In M. Mikulincer, & P.R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 11–31). Washington, DC: American Psychological Association.
- Graham, J., Nosek, B.A., & Haidt, J. (2012). The moral stereotypes of liberals and conservatives: Exaggeration of differences across the political spectrum. *PLOS ONE*, *7*, e50092.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by religion and politics*. New York: Pantheon Books.
- Halpern, D.E., Straight, C.A., & Stephenson, C.L. (2011). Beliefs about cognitive gender differences: Accurate for direction, underestimated for size. *Sex Roles*, *64*, 336–347.
- Haseleton, M.G., & Neute, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, *10*, 47–66.
- Hyde, J.S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*, 581–592.
- Ioannidis, J.P.A. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, e124.
- Ioannidis, J.P.A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, *7*, 645–654.
- Jacks, J.Z., & Devine, P.G. (2000). Attitude importance, forewarning of message content, and resistance to persuasion. *Basic and Applied Social Psychology*, *22*, 19–29.
- Jackson, C. (2004). *The Tinner Diaries*, other racist novels, inspire extremist violence. *Intelligence Report*, Southern Poverty Law Center. Retrieved December 30, 2014, from <https://www.splcenter.org/fighting-hate/intelligence-report/2004/turner-diaries-other-racist-novels-inspire-extremist-violence>
- Jost, J.T., Ghaser, J., Kruglanski, A.W., & Sulloway, F.J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*, 339–375.
- Jussim, L. (2012). *Social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy*. New York: Oxford University Press.
- Jussim, L., Crawford, J.T., Anglin, S.M., & Stevens, S.T. (2015). Ideological bias in social psychological research. In J. Forgas, K. Fredler, & W. Crano (Eds.), *Sydney Symposium on Social Psychology and Politics* (pp. 91–109). New York: Taylor & Francis.
- Jussim, L., Crawford, J.T., Stevens, S.T., & Anglin, S.M. (in press). Political distortions in the social psychology of intergroup relations. To appear in P. Valdesolo & J. Graham (Eds.), *Bridging ideological divides: Claremont Symposium on Applied Social Psychology*. Wiley-Blackwell.
- Kahan, D.M. (2011). Neutral principles, motivated cognition, and some problems for constitutional law. *Harvard Law Review*, *125*, 1–77.
- Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist*, *58*, 723–730.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480–498.
- Lakens, D. (2015). *Professors are not elderly: Evaluating the evidential value of two social priming effects through p-value analysis*. Unpublished manuscript.
- Lewandowsky, S., Oberauer, K., & Gignac, G.E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science*, *24*, 622–633.
- Lilienfeld, S.O. (2010). Can psychology become a science? *Personality and Individual Differences*, *49*, 281–288.
- Lilienfeld, S.O. (2015). Lack of political diversity and the framing of findings in personality and clinical psychology: Commentary on Duarte et al. *Behavioral and Brain Sciences*, *38*, 31–32.
- Lindner, N.M., & Nosek, B.A. (2009). Alienable speech: Ideological variations in the application of free-speech principles. *Political Psychology*, *30*, 67–92.
- Lord, C.G., Ross, L., & Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.
- Maccoby, E.E., & Jacklin, C.N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- Mahoney, M.J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, *1*, 161–175.

- Morgan, G.S., Skitka, L.J., & Lytle, B.L. (2014). *Objectively and universally true: Psychological foundations of moral conviction*. Paper presented at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.
- Munro, G.D., & Ditto, P.H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, *23*, 636–653.
- Neisser, U., Boodoo, G., Bouchard, T.J., Jr., Boykin, A.W., Brody, N., Ceci, S.J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.
- Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.
- Parker, M.T., & Janoff-Bulman, R. (2013). Lessons from morality-based social identity: The power of outgroup “hate,” not just ingroup “love.” *Social Justice Research*, *26*, 81–96.
- Sackett, P.R., Hardison, C.M., & Cullen, M.J. (2004). On interpreting stereotype threat as accounting for African American–White differences on cognitive tests. *American Psychologist*, *59*, 7–13.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, *115*, 336–356.
- Sibley, C.G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, *12*, 248–279.
- Silberzahn, R., Simonsohn, U., & Uhlmann, E.L. (2014). Matched-names analysis reveals no evidence of name-meaning effects: A collaborative commentary on Silberzahn and Uhlmann (2013). *Psychological Science*, *25*(7), 1504–1505.
- Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology. *Psychological Science*, *22*, 1359–1366.
- Simonsohn, U., Nelson, L.D., & Simmons, J.P. (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, *9*, 666–681.
- Skitka, L.J. (2010). The psychology of moral conviction. *Personality and Social Psychology Compass*, *4*, 267–281.
- Skitka, L.J., & Morgan, G.S. (2014). The social and political implications of moral convictions. *Advances in Political Psychology*, *35*, 95–110.
- Stanovich, K.E., West, R.F., & Toplak, M.E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science*, *22*, 259–264.
- Steele, C.M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.
- Steele, C.M., & Aronson, J.A. (2004). Stereotype threat does not live by Steele & Aronson (1995) alone. *American Psychologist*, *59*, 47–48.
- Swim, J.K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21–36.
- Tetlock, P.E. (2003). Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences*, *7*, 320–324.
- Tsuji, S., Bergmann, C., & Cristia, A. (2014). Community-augmented meta-analyses toward cumulative data assessment. *Perspectives on Psychological Science*, *9*, 661–665.
- Unger, R. (2011). SPSSI leaders: Collective biography and the dilemma of value-laden action and value-neutral research. *Journal of Social Issues*, *67*, 73–91.
- United Nations. (2014). Climate change impacting entire planet, raising risk of hunger, floods, conflict. Retrieved December 28, 2014, from <http://www.un.org/apps/news/story.asp?NewsID=47471>
- Van Lange, P.A.M. (2013). What we should expect from theories in social psychology: truth, abstraction, progress, and applicability as standards (TAPAS). *Personality and Social Psychology Review*, *17*, 40–55.
- Walton, G.M., Spencer, S.J., & Erman, S. (2013). Affirmative meritocracy. *Social Issues and Policy Review*, *7*, 1–35.
- Washburn, A.N., Morgan, G.S., & Skitka, L.J. (2015). A checklist to facilitate objective hypothesis testing: Commentary on Duarte et al. *Behavioral and Brain Sciences*, *38*, 42–43.
- Wetherell, G., Brandt, M.J., & Reyna, C. (2013). Discrimination across the ideological divide: The role of perceptions of value violations and abstract values in discrimination by liberals and conservatives. *Social Psychology and Personality Science*, *4*, 658–667.
- Winegard, B., Winegard, B., & Geary, D.C. (2015). Too paranoid to see progress: Social psychology is probably liberal, but it doesn't believe in progress: Commentary on Duarte et al. *Behavioral and Brain Sciences*, *38*, 43.
- Wolfe, C., & Spencer, S. (1996). Stereotypes and prejudice. *American Behavioral Scientist*, *40*, 176–185.
- Zuwerink, J.R., & Devine, P.G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology*, *70*, 931–944.