

Good Morals Gone Bad:  
Can High Moral Purposes Undermine Scientific Integrity?  
(draft chapter to appear in:  
J. Forgas, P. van Lange & L. Jussim, *Sydney Symposium on Social Psychology of Morality*)

Lee Jussim  
Rutgers University, New Brunswick and The Center for Advanced Study in the Behavioral Sciences,  
Stanford

Jarret T. Crawford  
The College of New Jersey

Sean T. Stevens  
Rutgers University, New Brunswick

Stephanie M. Anglin  
Rutgers University, New Brunswick

Jose L. Duarte  
Arizona State University

*Abstract*

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 towards 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 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 afterwards (e.g., Conway & Peetz, 2012). Similarly, 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 Turner 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 United States 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 United States 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, 2010; 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, & Collison, 2013; Crawford & Pilanski, 2014; Wetherell, Brandt, & Renya, 2013; for a review, see Brandt, Reyna, Chambers, Crawford, & Wetherell, 2014). For instance, self-identified American liberals and conservatives each 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 scientists may be vulnerable to similar processes. 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, Levine, Mackie, Morf, Vazire, & West, 2013). Science can tolerate individual mistakes and flawed theories, but only if it has reliable mechanisms for correction.

#### **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 pre-existing 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 counter-attitudinal 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., Munro & Ditto, 1997; Jacks & Devine, 2000; Zuwerink & Devine, 1996), people generate convincing arguments to justify their automatic evaluations, producing an illusion of objectivity (Haidt, 2001; Nickerson, 1998).

**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, Crawford, Stern, Haidt, Jussim, & Tetlock, 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, in press for a discussion of the related phenomenon of *QTIPs* – Questionable Theoretical and Interpretive Practices), this paper 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 longstanding efforts of social psychologists towards understanding these enduring social problems.

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 model (Winegard, Winegard, & Geary, in press):

*“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. 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 United States. Many psychologists studied sex differences from a feminist perspective, seeking to demonstrate that gender differences are small to non-existent 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 Men and Women*,” 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 above, 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 on many cognitive/academic characteristics among children (e.g., Briton & Hall, 1995; Halpern, Straight, & Stephenson, 2011; Swim, 1994).

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; Jussim et al., in press b, for others) – the erroneous and misleading interpretations of stereotype threat research.

Differences between White and African-American academic achievement is one of the great social problems plaguing the United States. 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 (and sometimes still do) 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 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.

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, Lustina, Good, Keough, Steele and Brown (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 (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) puts it on its current (as of January 2015) 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."

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 1?

**The confusing (misre)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 1 this manuscript). 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 pre-existing differences are unaffected by the manipulation, not because the means are equal. The equal *adjusted* means indicate that pre-existing 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 (p. 48) that, "... in fact, without this [covariate] adjustment, they would be shown to perform still worse than Whites..." They claimed that ANCOVA was conducted in order to reduce error variance. Although this is a valid use of ANCOVA, why they never acknowledged that removing stereotype threat simply maintained the prior difference until Sackett et al. (2004) pointed it out 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 -- pre-existing differences were maintained under non-threatening 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 had *no effect whatsoever* on African-American achievement in Steele & 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” that 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). 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., in press; Jussim et al., in press a, b). 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 (Lewandowsky, Oberauer, & Gignac, 2013)**

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 Lewandowski 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 1145 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 silly things like the faking of the moon landing is another *phantom fact*. Out of over 1145 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 non-mutually exclusive possible explanations for this pattern: 1. Conservatives really do hold more prejudices and are more intolerant than are those on the left; 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 towards 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 leftwing worldviews of social psychologists (see Duarte et al. for a review) which, until recently, rarely addressed prejudice against rightwing 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, Brandt, Chambers, Inbar, Motyl, & Wance, 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.

### **Conclusions: 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; Jussim et al, in press a,b 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 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. 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**

Unfortunately, we feel compelled to end this chapter on a mixed, rather than purely optimistic note. On the positive side, we think there are many researchers who will be earnestly interested in maximizing the objectivity and validity of their major conclusions. We hope such scientists find this chapter useful, both in recognizing the distorting effects of high moral purposes among others, and, perhaps, to limiting such biases in their own work.

On the other hand, we are more skeptical that this review and others like it (Duarte et al., in press; Jussim et al, in press a,b) will have any salutary effect on the validity or objectivity of the claims of those who are driven to promote what they believe to be high moral purposes. Knowing how to avoid QIPs will advance science only among those who give top priority to accurately describing scientific evidence. The history of research on moralized topics, however, suggests that there may be some distance to go before there is a broad consensus on how to prioritize evidence over moral and other agendas.

### 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. (1991). In the Supreme Court of the United States: Price Waterhouse v. Ann B. Hopkins (Amicus curiae brief). *American Psychologist, 46*, 1061-1070.
- American Psychological Association (2006). Stereotype threat widens achievement gap. Retrieved on 12/28/14 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.
- Brant, M. J., & Proulx. (in press). QTIPS: Questionable theoretical and interpretive practices in social psychology, Commentary on Duarte et al. (in press). *Behavioral and Brain Sciences*.
- Brandt, M. J., Reyna, C., Chambers, J. R., Crawford, J. T., & Wetherell, G. (2014). The ideological-conflict hypothesis: Intolerance among both liberals and conservatives. *Current Directions in Psychological Science, 23*, 27-34.
- Briton, 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 Nam, 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, 0956797613504966*.
- 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. (in press). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*. doi:  
<http://dx.doi.org/10.1017/S0140525X14000430>
- 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, England: 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.  
doi:10.1371/journal.pone.0050092
- 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*. Pantheon Books: New York.
- Halpern, D. F., Straight, C. A., & Stephenson, C. L. (2011). Beliefs about cognitive gender differences: Accurate for direction, underestimated for size. *Sex Roles*, *64*, 336-347.

- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47-66.
- 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 Turner Diaries, other racist novels, inspire extremist violence. *Intelligence Report*, Southern Poverty Law Center. Retrieved 2014-12-30.
- Jost, J. T., Glaser, 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. (in press). Ideological bias in social psychological research. To appear in J. Forgas, K. Fiedler, & W. Crano (Eds.), *Sydney Symposium on Social Psychology and Politics*. (cited as "Jussim et al, in press a").
- 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. (cited as "Jussim et al, in press b").
- Kahan, D. M. (2011). Neutral principles, motivated cognition, and some problems for constitutional law. *Harvard Law Review, 125*, 1.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.
- 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. (in press). Lack of political diversity and the framing of findings in personality and clinical psychology: Commentary on Duarte et al. (in press). *Behavioral and Brain Sciences*.
- 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, California: 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., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. 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.

- Redding, R. E. (2001). Sociopolitical diversity in psychology. *American Psychologist, 56*, 205-215.
- 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.
- 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, 0956797614533802*.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology. *Psychological Science, 22*, 1359–1366.
- Skitka, L. J. (2010). The psychology of moral conviction. *Personality and Social Psychology Compass, 4*, 267-281.
- Skitka, L. J., Bauman, C. W., & Mullen, E. (2008). Morality and justice: An expanded theoretical perspective and empirical review. *Advances in Group Processes, 25*, 1-27.
- 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.
- Stevens, S. T., Anglin, S. M., & Jussim, L. (in press). The political self. To appear in D. McInerney, R. Craven, H. Farsh, & F. Guay (Eds.), *Self-concept, motivation and identity: Underpinning success with research and practice*.

- 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.
- 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 on 12/28/14 from <http://www.un.org/apps/news/story.asp?NewsID=47471>
- 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. (in press). A checklist to facilitate objective hypothesis testing: Commentary on Duarte et al. (in press). *Behavioral and Brain Sciences*.
- 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. (in press). Too paranoid to see progress: Social psychology is probably liberal, but it doesn't believe in progress, Commentary on Duarte et al. (in press). *Behavioral and Brain Sciences*.
- 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.

Figure 1

Figure 2 from Steele &amp; Aronson (1995)

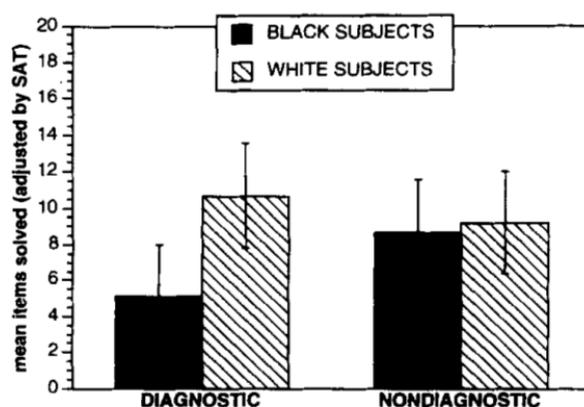


Figure 2. Mean test performance Study 2.

Copied and pasted whole from the original article, page 802. The *Figure 2 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 *pre-existing differences* (of about 40 points) were maintained in the nondiagnostic condition. Stereotype threat increased achievement test differences; removing it did not reduce the mean differences between African-Americans and Whites *at all*.