

Examining the asymmetry in judgments of racism in self and others

Angela C. Bell, Melissa Burkley & Jarrod Bock

To cite this article: Angela C. Bell, Melissa Burkley & Jarrod Bock (2018): Examining the asymmetry in judgments of racism in self and others, The Journal of Social Psychology, DOI: [10.1080/00224545.2018.1538930](https://doi.org/10.1080/00224545.2018.1538930)

To link to this article: <https://doi.org/10.1080/00224545.2018.1538930>



Published online: 15 Nov 2018.



Submit your article to this journal [↗](#)




Article views: 185



View Crossmark data [↗](#)



Examining the asymmetry in judgments of racism in self and others

Angela C. Bell ^a, Melissa Burkley^b, and Jarrod Bock^b

^aLafayette College; ^bOklahoma State University

ABSTRACT

Across three experiments, participants were provided with a list of racist behaviors that purportedly were enacted from a fellow student but in fact were based on the participants' own behaviors. People consistently evaluated themselves as less racist than this comparison other, even though this other's racist behaviors were identical to their own. Studies 2a and 2b demonstrate this effect is quite robust and even occurs under social pressure and social consensus conditions in which participants were free to express their racial biases. Thus, it appears that people are less likely to base their racist trait ratings on behavioral evidence when evaluating themselves compared to when they are evaluating another. Taken together, this work provides evidence for the consistency and robustness of self-enhanced social comparisons as applied to the trait domain of racism. Further, this work sheds insight into why people deny they are racist when they act racist.

ARTICLE HISTORY

Received 14 July 2018
Revised 8 October 2018
Accepted 11 October 2018

KEYWORDS

Better than average effect; racism; self-enhancement; self-perception; self-other differences

People who exhibit behaviors labeled as racist often deny their behaviors indicate that they are a racist person. For example, comedian/actor Michael Richards called a Black audience member a racial slur during a standup performance and later apologized, stating, "I am not a racist" (Faber, 2010). Similarly, celebrity chef Paula Deen admitted under oath during a trial that she used racial slurs in her past. When asked if she was a racist, her reply was "No, I am not" (Moraski, 2013). More recently, self-identified "white nationalist" Peter Cvjetanovic was photographed at the Unite The Right rally in Charlottesville, Virginia, and following public criticism, stated, "I'm not the angry racist they see in that photo" (Sinclair, 2017). These individuals' assessments of their behavior were at odds with public opinion that viewed such behaviors as a straightforward indication of being racist. Thus, these examples appear to highlight an asymmetry between the ways people view racism in themselves versus how others view their racism. What could be the reason for this asymmetry? It could be that they perceive racism differently than their observers and would judge anyone who engages in similar behaviors as non-racist. Alternatively, it could be that they perceive racism similarly to their observers but are unable or unwilling to see racism in themselves.

The present work sought to investigate how social comparisons influence evaluations of one's own prejudiced attitudes and behaviors. Specifically, we posed the question of whether people view the exact same behaviors as indicative of being racist when enacted by oneself versus another person. To do so, people viewed a description of another person's behavior that, unbeknownst to them, was in fact their own behavior recorded earlier. Consistent with the notion that people are often reluctant to see flaws in themselves (Critcher, Helzer, & Dunning, 2011; Sedikides & Alicke, 2012; Taylor & Brown, 1988), we predicted people would label themselves as less racist than this other person, even when this other person was in fact themselves.

The link between behavioral and trait judgments

It is often assumed that people form judgments about their own traits by examining their behaviors and comparing to that of relevant comparison targets (Festinger, 1954; Gilbert, Giesler, & Morris, 1995). For example, a woman will likely label herself as extraverted if she determines that she engages in more sociable behaviors than similar others around her. This basic assumption that behaviors are used to form trait judgments has been asserted by a number of trait theorists (Allport, 1937; Buss & Craik, 1983; Funder, 1991) and also serves as the foundation for several classic theories in social psychology, including social comparison theory (Festinger, 1954; Suls, Martin, & Wheeler, 2002) and self-perception theory (Bem, 1967, 1972). Given this assumption, those who engage in racist behaviors would clearly and consistently be labeled as racist people.

However, not all researchers agree with this assumption. Work by Klein, Loftus, and colleagues challenges the notion that trait judgments always rely on the examination of relevant behaviors. Instead, their research suggests that people may also rely on abstract trait summaries (e.g., “I am an extraverted person), rather than specific behaviors, when forming trait judgments (Klein, Loftus, & Sherman, 1993; Klein, Loftus, Trafton, & Fuhrman, 1992). In such a case, a woman would not need to retrieve specific sociable behaviors from her memory to label herself extraverted. She would simply need to retrieve the abstract trait of being extraverted. In support of this assertion, several studies found evidence that information about traits and behaviors are stored in separate memory systems (Klein et al., 1993, 1992). Such work lends support to the assertion that behavior and trait information often exist independently and that people do not necessarily access behavior information when forming trait judgments. Given this assumption, just because an individual engages in racist behaviors does not mean they will in turn label themselves a racist.

It appears then that sometimes people rely on specific behaviors when forming traits and sometimes they do not. What can account for when one process will occur over the other? Although several factors have shown to influence which type of process will be enacted (Klein et al., 1993, 1992; Klein, Sherman, & Loftus, 1996), one such factor has to do with whether the trait judgments are in regards to oneself or another person. For instance, one study examined an individual who had a developmental disorder that impaired his ability to retrieve episodic memories, thus preventing him from being able to recall specific behaviors when forming trait judgments (Klein, Cosmides, Murray, & Tooby, 2004). This case study revealed that this individual was unable to form accurate trait judgments of others but was able to form an accurate trait judgment of himself. This outcome suggests that people are likely to rely on specific behaviors when forming trait judgments of others but rely on abstract trait concepts when forming trait judgments of themselves (Klein et al., 1992, 1993, 1996). As such, there appears to be an asymmetry in the behavior-trait link when it comes to judging others versus one’s own traits (Malle, 2005). Because of this asymmetry, people are free to form trait judgments of themselves that may run in direct contrast to the behaviors they exhibit (Vazire, 2010).

Overly positive self-judgments: the better-than-average effect

Because people are less likely to rely on their own behaviors when evaluating their own traits, they are allowed considerable flexibility in the types of self-judgments they can create. As a result, people often form self-judgments that are overly positive and driven by a need for self-enhancement (Sedikides & Alicke, 2012; Taylor & Brown, 1988). A prominent example of this positive bias can be seen in the better-than-average effect (BTAE), which refers to the tendency for people to evaluate themselves more favorably than they evaluate the average other (Alicke, 1985; Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; Brown, 2012; Dunning, Heath, & Suls, 2004). For example, people consistently rate themselves as more honest, intelligent, talented, dependable, and sympathetic than their average peer (for review, see Alicke & Govorun, 2005; Brown, 2012). Similarly, stock market investors believe they are better than their peers are at identifying profitable investment opportunities (Odean, 1998), managers believe they are more skilled than their

colleagues (Malmendier & Tate, 2005), and college professors believe they are better teachers than the average college instructor (Cross, 1977). Ironically, people even think they are less likely to fall prey to psychological biases such as the BTAE than their average peer (Pronin, Gilovich, & Lee, 2004). A study by Saucier (2002) first established the link between the BTAE and evaluations of racial prejudice and found that participants rated themselves as (1) having less negative views about African Americans and (2) finding racist arguments as less convincing than the average person. Similarly, Howell and Ratliff (2017) found that people evaluated themselves as less biased than the average person across multiple domains of prejudice measured via the IAT. Although the BTAE reflects a biased view of oneself, it serves a beneficial function by allowing people to maintain high self-esteem and avoid depression (Critcher et al., 2011; Taylor & Brown, 1988).

There seems to be a division between cognitive and motivational explanations for why the BTAE occurs. Cognitive models claim the processes underlying comparative judgments between people are not unlike comparisons of other objects, such as food (Giladi & Klar, 2002). Conversely, motivational models argue the BTAE occurs due to a desire to preserve or enhance self-worth (Alicke & Sedikides, 2009; Brown, 2012; Guenther & Alicke, 2010). For example, the BTAE occurs more frequently among people high in subjective well-being (Goetz, Ehret, Jullien, & Hall, 2006) and when people evaluate qualities that are “important” compared to those rated as “unimportant” (Brown, 2012). Being labeled as a racist person is a potential threat to a positive self-image that underlies psychological well-being, and being perceived as non-racist is likely important to maintain a positive self-image in the eyes of others. Therefore, the desire to perceive oneself as less racist than others gives credence to the motivational explanation. However, if the desire to view the self as non-racist is automatic and leads to a blindspot for recalling evidence of the contrary, this self-other asymmetry within the prejudice domain is another cognitive bias.

Overly positive self-judgments: the better-than-myself effect

Although people generally rate themselves more favorably in comparison to an abstract concept, like the aggregate “average” person, research has also found that people still rate themselves more favorably in comparison to a concrete individual (Alicke et al., 1995; Brown, 2012). Interestingly, this occurs even when the comparison individual’s trait or behavioral ratings are actually made by the participants themselves, and the tendency for people to perceive themselves as better than those who are identical to them is known as the better-than-myself effect (BTME; Alicke, Vredenburg, Hiatt, & Govorun, 2001; Guenther & Alicke, 2010). In a series of studies, Alicke et al. (2001) compared participants’ self-ratings on a wide range of personality traits (e.g., cooperative, polite, intelligent, musical) to their ratings of a hypothetical other. Unbeknownst to participants was the fact that the behaviors reportedly enacted by this hypothetical other were all behaviors the participant had reported committing themselves during an earlier study session. The hypothetical other was an exact reflection of the actual participant. The point of this work was to determine if an above-average bias in trait ratings would still occur even if the trait-relevant behaviors were held constant across both comparison targets. As expected, the results showed that participants continued to rate themselves as better than the comparison other, even when this other engaged in identical behaviors as the participant. This BTME also emerged when individuals compared themselves to another concrete individual, and not just the abstract concept of the “average” person (Alicke et al., 2001).

The BTME represents an interesting extension of the well-established BTAE, yet it has not received as much empirical attention as its predecessor. One difficulty in studying the BTAE is that it often involves comparing trait ratings of two targets (i.e., self and other) that may truly differ in terms of their behaviors. This is especially true for previous studies linking the BTAE and evaluations of prejudice (Howell & Ratliff, 2017; O’Brien et al., 2010; Saucier, 2002). Statistically speaking, not everyone can be better than average in the way that is often reported in BTAE studies, but some people truly are better than average or better than their comparison target. The BTME studies addressed this issue by holding the behavior in question constant across both comparison

targets. They also provide additional support to the assertion that behaviors and traits can exist largely independently when it comes to self-evaluations. Further, these studies offer guidance for a methodology that could be used to determine if people are more willing to view racism in others than themselves, and unwilling to label their own behaviors as racist, even when these others engage in the same racist behaviors. We designed the present work to test this intriguing possibility.

Present research: a less-racist-than-myself effect?

As the opening examples suggest, people who commit racist acts often deny that they are racist. When such an event occurs, two potential explanations seem possible. On the one hand, these people may just hold a different definition of prejudice than their observers. If this were the case, then they should be reticent to label a person as racist who engages in such behaviors, regardless of whether that person is themselves or someone else. On the other hand, these people may hold the same definition but may be reluctant to perceive themselves as racist because of self-enhancement reasons. If this were the case, denial of racism would not occur if they were to observe someone else committing the same racist behaviors that they had. To address this question empirically, the present studies were designed to test if people consider themselves less racist than those who engage in the exact same questionably racist behaviors.

Because there are few studies in regards to the BTAE as it relates to perceptions of prejudice, our goal for the present work was modest. Essentially, we sought to determine if people would consider themselves less racist than another person who reportedly engages in the same racist behaviors that they have. By doing so, it allows us to see if people base their racist trait ratings on behavioral evidence or something else, and if they do so differently depending on whether the trait rating is for themselves or another. Consistent with research on self-enhancement (i.e., BTAE and BTME; Brown, 2012; Sedikides & Alicke, 2012), and the previous work linking the BTAE and prejudice (Howell & Ratliff, 2017; Saucier, 2002), we predicted that people would label themselves as less racist than a comparison other, even when this other was in fact themselves.

Study 1 was designed to determine if people evaluate themselves as less racist using similar methodology used by prior BTME research (Alicke et al., 2001). To provide a more conservative test of the effect in question, we had participants compare themselves to a specific comparison other, rather than a generic “average peer,” because prior research suggests that a comparison with a specific-other decreases (although it does not eliminate) the BTAE (Alicke et al., 1995). We designed Studies 2a and 2b using a similar methodology to replicate this effect and provide an initial attempt at understanding the potential process by which the effect occurs. Consistent with Alicke et al. (1995), we view the BTAE and its extension in self-other evaluations of racism as driven by a basic heuristic to view the self positively rather than a specific desire to not be viewed as racist. If this effect is truly an extension of the automatic tendency to view the self positively, then it should not be influenced by manipulations known to increase the explicit motivation to control or express one’s prejudice. Alternatively, if the motivational desire to not look racist drives this effect, then such manipulations should alter its pattern. Studies 2a and 2b explored this question uniquely by incorporating two different manipulations of prejudiced motives (e.g., social pressure and social consensus information). Materials and data for all studies reported in our article can be accessed at https://osf.io/h7k9p/?view_only=2b0f6b3dbce2430fb41b2b85b27daf75.

Study 1

Method

Participants

Ninety-nine students (62 female, 37 male; mean age = 19.82 years) from a Southwestern university participated in this study for partial course credit. The study was conducted in compliance with the university’s Institutional Review Board. Racial background was as follows: 68% European American/

White, 6% African American/Black, 1% Native American/Alaskan Native, 4% Latino/Hispanic, 3% Asian American, and 18% unspecified or multiracial. Analyses indicated that gender, age, race, and the distinction between being White or a participant of color did not influence the results.

Procedure

For all studies, we used a modified version of the procedure created by Alicke et al. (2001). The first phase of the study occurred at the beginning of the semester within the context of an online prescreener questionnaire. At this university, the prescreener consisted of a number of questionnaires used by researchers to determine the studies for which individuals qualify (e.g., individuals who scored high on a depression measure qualify for a study on depression), and our questions of interest were embedded within. Participants were presented with a list of 46 behaviors and asked to indicate if they had ever engaged in each behavior by selecting “yes” or “no.” To hide the true purpose of the study, this list included a wide range of undesirable behaviors (e.g., “Have you ever lied to get out of a gathering with friends or family?”; “Have you ever worn the same underwear two days in a row?”). Within this list were also 30 questionably racist behaviors that were identified during prior pilot testing (e.g., “Have you ever...used the N word to refer to Blacks? ...laughed at another person’s joke about Asian people? ... avoided interacting with a Muslim person out of fear? ... avoided an interaction with someone because they spoke with a thick foreign accent?”). After completing the behavior checklist, participants were asked to rate themselves compared to the average student on several filler traits, as well as the racist trait of interest.

The second phase of the study took place months later during an in-lab session. Upon signing up for this seemingly unrelated study, the participants’ name provided in the prescreener was temporarily used by the lead investigator and first author to generate a de-identified student profile with their prescreener responses. The participants’ names were necessary to determine exactly what information each individual participant was presented with. After the profile construction, participants’ names were removed and replaced in the dataset with an alternative ID code participants provided during the prescreener. This alternative ID code is a combination of the participants’ initials and birth date and was created (with consent) to link their prescreener responses to data in related follow-up studies in the human subjects’ pool system. The lead investigator provided de-identified paper survey packets for research assistants, who administered surveys directly to participants in the lab. The front of these survey packets was temporarily time-stamped to correspond with the research timeslot each participant signed up for, and during data collection the research assistants could not see the profile page constructed for each participant.

After providing consent, participants were told they would be reviewing a randomly selected student’s prescreener responses. Unbeknownst to the participant, this “other” person had the exact same behaviors that the participant provided on the prescreener. For example, if a participant indicated in the pre-screener that they “avoided an interaction with someone because they spoke with a thick foreign accent,” “worn the same underwear two days in a row,” but never “lied to get out of a gathering with friends or family,” they were later shown these exact same behaviors under the guise that it was reported by another student. All participants were exposed to their own behaviors under the guise that these were someone else’s. After reviewing the comparison other’s responses, participants rated this other in regards to several undesirable trait ratings. Participants then created an alternative ID using the same parameters of the prescreener ID code to link their data (and during data entry was also used as a check to confirm that each participant received their own responses). Finally, participants provided demographic information and were given an open-ended question asking what they thought was the purpose of the current study. This final question was used to probe for potential suspicion that participants were aware that they received their own prescreener responses. After completing the task, participants were debriefed and explained the true intentions of the study. During the open-ended suspicion-probe, 10 participants expressed suspicion that the comparison

other's responses were in fact their own. One participant did not complete the study. All were removed from the analyses, resulting in $N = 88$, and removing these participants did not impact the overall results.

Materials

Self-other comparison measure

In line with BTAE research, participants were specifically asked to compare this target to the average student (e.g., "Compared to the average fellow college student, how *racist* is this person?"). Next, participants rated themselves in the same manner (e.g., "Compared to the average fellow college student, how *racist* are you?"). All ratings were made using a -4 (much less *racist* than average) to 4 (much more *racist* than average).

We chose the verbiage of these items to be consistent with prior research (e.g., Alicke et al., 1995, 2001; Brown, 2012) but also to avoid a potential methodological confound. We could have simply asked participants to rate how racist they were compared to the comparison other, but doing so creates a problem because prior research suggests that people automatically give more weight to target in a question (i.e., self) than the referent (i.e., comparison other; Windschitl, Conybeare, & Krizan, 2008). To avoid this possible difference in weighting, we asked two separate questions, each with the person of interest serving as the target of the question.

In addition to the racist trait comparison, participants evaluated themselves and the target compared to the average student on dishonest and disgusting trait ratings to determine whether the racist comparisons differed from comparing other undesirable behaviors.

Results and discussion

A series of paired sample t -tests were conducted to compare participants' racist ratings of the comparison other to ratings of themselves. A post hoc power analysis indicated we had sufficient power (0.956) to detect the observed difference in self-other racist ratings ($d = 0.36$). As predicted, there was a significant difference between the other-rating and self-rating, such that participants rated themselves as less racist ($M = -1.53$, $SD = 1.70$) than the target other ($M = -0.97$, $SD = 1.75$), $t(87) = 3.45$, $p = .001$, $d = 0.36$ (Figure 1). Participants considered themselves less racist than the comparison other, despite the fact that the comparison other engaged in the exact same racist behaviors as the participants.

Participants also rated themselves as less dishonest ($M = -1.48$, $SD = 1.48$) than the target ($M = -0.91$, $SD = 1.59$), $t(87) = 3.20$, $p = .002$, $d = 0.34$, and less disgusting ($M = -1.59$, $SD = 1.65$) than the target ($M = -1.01$, $SD = 1.76$), $t(87) = 3.12$, $p = .002$, $d = 0.33$. Across all domains, participants demonstrated a better-than-myself effect. Because participants' self-other evaluations are similar across the three undesirable behaviors this may be due to an automatic tendency to self-enhance rather than a unique motivation to deny oneself as racist.

One possible explanation for these results is that participants' definition of what constitutes being racist changed from the time of the prescreener to the time of the in-lab study. To test for this possibility, we compared participants' racist self-ratings during the prescreener to their self-ratings during the in-lab session. As expected, there was no significant difference between these two racist evaluations, $t(80) = 0.50$, $p = .617$, $d = 0.056$. Participants' tendency to rate another as more racist than themselves even though this "other" person was in fact themselves is not due to a change in definitions of racism over time. There were also no significant differences for the disgusting or dishonest evaluations; $t(81) = -0.26$, $p = .799$, $d = 0.03$ and $t(80) = -1.27$, $p = .207$, $d = 0.14$, respectively. These findings suggest that people who deny undesirable traits likely do so not out of a difference in definition of what it means to exhibit (or possess) these traits, but rather out of a self-enhancing tendency to see themselves in a positive light.

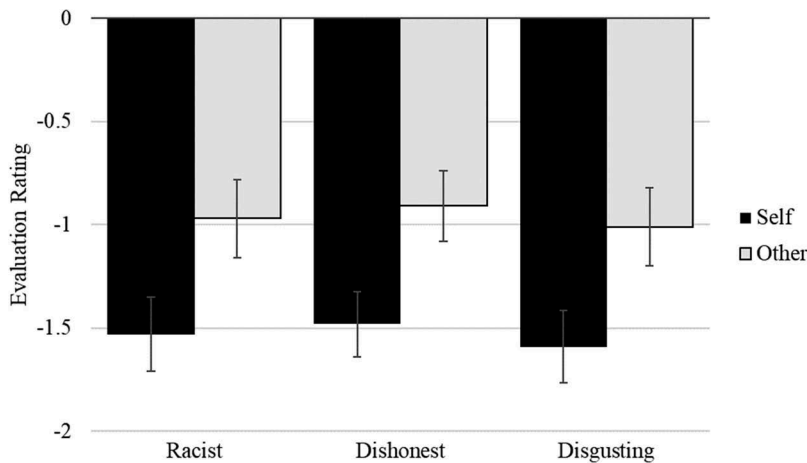


Figure 1. Racist, dishonest, and disgusting evaluations for study 1. The lower the bar, the less racist, dishonest, or disgusting, the target (self or identical other) was rated compared to the average.

Studies 2a and 2b

Study 1 provided initial evidence that the better-than-myself effect extends to evaluations of racism. Studies 2a and 2b were designed to replicate this effect and attempt to understand the conditions that may weaken or strengthen the self-other racism comparison. To do this, we followed an approach used by Alicke et al. (1995; Study 7). In one of their studies, Alicke and colleagues sought to demonstrate that a cognitively effortful process did not underlie the BTAE. To determine this, they added a cognitive load manipulation to their typical BTAE study and showed that the BTAE held across both high and low load conditions. For our set of studies, we sought to demonstrate that an explicit motive to control prejudice is not underlying the biased self-other comparisons. As such, studies 2a and 2b incorporated two different but well-established manipulations of prejudiced motives (social pressure and social consensus information). In both cases, we predicted that the manipulation would not have an impact, such that the biased racism comparison would even occur when motivations to control prejudice are low. Such a result would demonstrate that the desire to appear non-prejudiced is not a necessary condition for this comparison bias and would provide further evidence that this effect is routed in a more basic heuristic, akin to the BTAE.

It is important to note that our prediction that the prejudice motives manipulations would have little impact on our effect is equivalent to predicting a null effect. Within psychological science, there appears to be an aversion—intentionality aside—to constructing paradigms that assess the falsifiability of theories (i.e., testing null hypotheses; Ferguson & Heene, 2012; Heene, 2013). The infrequency of such designs does not negate the importance of falsifiability to the scientific enterprise. Although null predictions are not necessarily the norm, there are contexts when they are justified (Cashen & Geiger, 2004; Cohen, 1990; Cortina & Dunlap, 1997; Greenwald, 1993; Nickerson, 2000). This is especially the case when a researcher is trying to reject a theoretical explanation for an effect (Greenwald, 1993), such as is the case with the present work. However, when researchers make a null prediction, certain recommendations should be followed. First, it is imperative that the study has a large enough sample to provide adequate statistical power (Cashen & Geiger, 2004). To address this, we report power analyses for studies 2a and 2b and ensure that our sample sizes exceed this number. Second, in addition to the predicted null effect of one independent variable, the study should also include another independent variable that is known to have an effect on the dependent variable, thereby demonstrating the presence of a null effect and a significant effect within the data (Cashen & Geiger, 2004). In our studies, we included the self-other rating variable from Study 1 in addition to the predicted null variable. This allowed us to attempt to replicate the effect seen in Study 1 in addition to testing the null effect of prejudice motives.

Study 2a

The purpose of this study was to test the influence of social pressure on comparative evaluations of racism. Social proof is a persuasive influence on behavior (Cialdini, Reno & Kallgren, 1990), and prior research has more specifically demonstrated the influence of social pressure on expressions of prejudice (Buzinski & Kitchens, 2017; Crandall, Eshleman, & O'Brien, 2002; Monteith, Deneen, & Tooman, 1996). A main assumption would be that increased social pressure to act non-prejudiced or “politically correct” would only strengthen the established comparison bias, or low social pressure to act non-prejudiced would de-bias the effect. However, as previously stated, we predicted that participants will continue to evaluate their behaviors as significantly less racist than the comparison other, regardless of the allowance or relief to justify their own racist behaviors, because the automatic tendency to self-enhance subjugates external motives to behave.

Method

Participants and design

One hundred and thirty-one students from the same Southwestern University participated in this study for partial course credit. Participants (mean age 18.66 years) were predominantly women (77%). Racial background was as follows: 76% European American/White, 6% African American/Black, 5% Native American/Alaskan Native, 3% Latino/Hispanic, and 9% unspecified or multiracial. Analyses indicated that gender, age, race, and the distinction between being White or a participant of color did not influence the results. Participants were randomly assigned to one of two experimental conditions: high pressure or low pressure.

A power analysis (G*Power; Erdfelder, Faul, & Buchner, 1996) indicated that a sample size of 128 participants would be needed to detect a significant interaction between the self-other rating variable (within-subjects) and the social pressure manipulation variable (between-subjects manipulation), assuming a medium effect size ($f = .25$) and a power level of .80 (Cohen, 1992).

Materials and procedure

The first phase of the study was identical to that of Study 1. The second phase was nearly identical to Study 1, but it included an experimental manipulation designed to influence an external motivation to control prejudice. Specifically, participants were randomly assigned to read information known to increase or decrease social pressure to avoid expressing prejudice. Similar to Study 1, the lead investigator randomly assigned participants to conditions, and the manipulated instructions were obscured from the research assistants administering surveys during data collection. Participants in the low social pressure condition were encouraged to express their own attitudes even if they were not politically correct and were reminded their responses were completely confidential. Participants in the high social pressure condition were told to keep in mind society's norms for being politically correct when expressing their own attitudes. The full manipulation instructions adapted from Payne, Burkley, and Stokes (2008; Study 4) are available in our supplementary materials.

Participants completed a manipulation check to assess perceptions of social pressure, “To what extent do you believe these instructions emphasize that you can express potential prejudices?” This question was evaluated on a 9-point scale ranging from -4 (*not at all*) to $+4$ (*extremely*). Then, participants rated how racist the comparison other was and themselves in a manner identical to that of Study 1 (e.g., “Compared to the average fellow college student, how racist are you?”). To conceal the purpose of the study, participants rated the target and themselves on a number of non-relevant traits (e.g., cooperative, aggressive, sophisticated, intelligent, polite, etc.; taken directly from Alicke et al., 2001). All ratings occurred on a 9-point scale ranging from -4 (much less *racist* than average) to 4 (much more *racist* than average). Participants also completed the perspective-taking subscale of the Interpersonal Reactivity Index (IRI; Davis, 1983) as an exploratory moderator of the self-enhanced social comparison. Due to a lack of a priori hypotheses for this measure, and non-

significant main effects and a non-significant interaction between the IRI and the social pressure manipulation, we discuss this measure (and the exploratory moderators used in Study 2b) in greater detail in our supplementary materials. Finally, participants were debriefed and told the true intentions of the study. One participant expressed suspicion they received their own answers during the open-ended suspicion probe, and another did not complete the study. These participants were removed from the analyses, resulting in an $N = 129$, but this removal did not influence the overall results.

Results

An independent samples t -test assessed the effectiveness of the social pressure manipulation. There was a significant difference in perceived social pressure between the high- and low-pressure conditions, $t(127) = 3.24$, $p = .002$, $d = 0.57$. Participants in the low social-pressure condition perceived greater freedom in expressing bias ($M = 1.49$, $SD = 1.82$) than those in the high social-pressure condition ($M = 0.34$, $SD = 2.21$). Thus, our social pressure manipulation was successful.

Next, a 2 (target: self vs. other rating) \times 2 (social pressure: high vs. low) mixed model ANOVA was conducted on the racist evaluations. Once again, there was a main effect for target, $F(1, 127) = 49.46$, $p < .001$, $\eta_p^2 = .28$, such that participants rated themselves as significantly less racist ($M = -1.85$, $SD = 1.61$) than the comparison other ($M = -0.71$, $SD = 1.71$). In regards to the social pressure manipulation, there was no main effect of pressure, $F(1, 127) = 0.096$, $p = .757$, $\eta_p^2 = .001$, nor a target \times pressure interaction, $F(1, 127) = 0.97$, $p = .326$, $\eta_p^2 = .008$ (Figure 2). The biased effect remained even when people felt free to express their prejudiced biases (i.e., low social pressure).

Study 2b

Study 2a replicated the biased racism comparison found in Study 1 and also demonstrated it persists despite a low motivation to control one's prejudiced responses (i.e., high social pressure), providing preliminary evidence of this effect's robustness. Study 2b was similar to Study 2a, but it manipulated another external motive to control prejudice by providing consensus information. Prior research indicates that when people receive consensus information indicating their prejudiced beliefs and behaviors are uncommon (i.e., low consensus), they are reluctant to express their own prejudices and often alter their responses to be more egalitarian (e.g., Klein & Goethals, 2002; Puhl, Schwartz, &

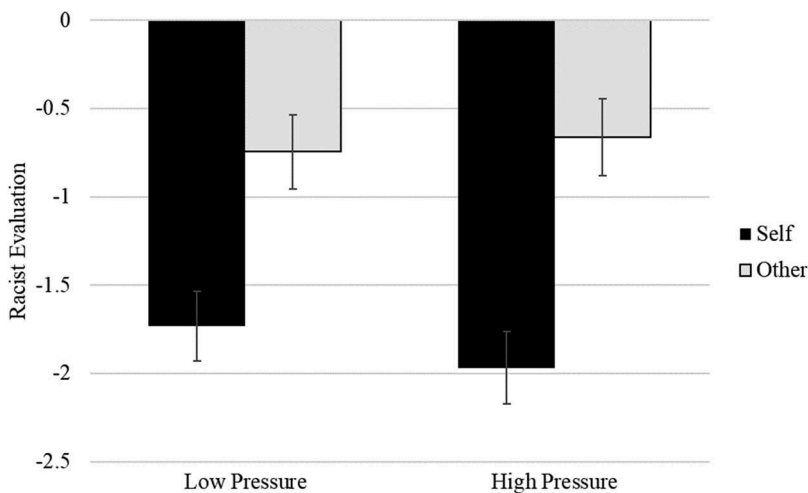


Figure 2. Racist evaluations for study 2a as a function of target (self vs. other) and social pressure (low vs. high).

Brownell, 2005; Sechrist & Milford, 2007; Stangor, Sechrist, & Jost, 2001). Research on the false consensus effect demonstrates that people have an automatic tendency to believe that others share their opinions, beliefs, and values (Alicke & Largo, 1995). Further, this effect emerges when people compare perceptions of the self to how others perceive them (Sechrist & Stangor, 2001; Tabachnik, Crocker, & Alloy, 1983). For example, participants were less likely to endorse obesity stereotypes when told their fellow peers held favorable attitudes toward obese people, and more likely to endorse obesity stereotypes when told their peers held negative attitudes (Puhl et al., 2005). Similarly, participants who were told the majority of their peers held favorable attitudes toward African Americans were more willing to help an African American in need (Sechrist & Milford, 2007).

We also selected this manipulation to address a possible alternative explanation for the null results of Study 2a. It is possible that the null effect of the pressure variable was not caused by the robustness of our effect but rather because social pressure alone may be an insufficient motive to moderate the effect. Because consensus feedback has shown to be effective in changing people's attitudes and behavior in comparison to others, it is of particular interest to the present study to test the robustness of the biased racism-based comparison against another robust manipulation.

Method

Participants

Students ($N = 285$) from the same Southwestern University participated in this study in return for course credit. Participants were predominantly women (69%; 30% male, 1% unspecified) and traditionally college-aged (mean age = 19.31 years). Ethnic/racial background was as follows: 71% White/European American, 7% Black/African American, 4% Native American/Alaskan Native, 4% Latino/Hispanic, 2% Asian American, 11% unspecified or multiracial, and one participant did not report. Similar to Study 2a, a sample size of 128 participants would be needed to detect a significant interaction between the self-other rating and the social pressure manipulation, assuming a medium effect size ($f = .25$) and a power level of .80. To ensure we had an adequate sample size given our null prediction, and to account for possible drop-out of participants due to suspicion, we made sure our sample sufficiently exceeded this number.

Materials and procedure

The first phase of the study was identical to that of Study 1. The second phase was nearly identical to Study 1 and 2a but included the consensus manipulation. Participants were randomly assigned to receive high or low consensus information regarding the racist behaviors listed. The wording for this manipulation was taken directly from Sechrist and Stangor (2001). Specifically, participants were told: "Based on extensive research at the present university, the percentage of fellow students who also engaged in these behaviors listed on the previous questionnaire is 81% (*high consensus*)/19% (*low consensus*)."

After the consensus information, participants rated how racist the comparison other was and themselves in a manner identical to that of Study 1 and 2a (including additional positive and negative filler trait ratings). Participants then completed a number of individual differences that could potentially moderate the effect (e.g., external and internal motivations to control prejudice, Plant & Devine, 1998; life satisfaction, Diener, Emmons, Larsen, & Griffin, 1985; trait self-esteem, Rosenberg, 1965; private and public self-consciousness, Scheier & Carver, 1985; and social anxiety, Fenigstein, Scheier, and Buss, 1975; and self-monitoring, Snyder, 1974). Finally, participants were debriefed and told the true intentions of the study.

Five participants expressed suspicion that the comparison other's responses were their own and were removed from the analysis, which did not influence the overall results. Thirteen participants failed the attention check question (e.g., What was the percentage of [fellow university] students that engaged in the behaviors listed on the previous questionnaire?) and were removed from the analyses. Removing all 18 participants resulted in $N = 267$.

Results

A 2 (target: self vs. other rating) \times 2 (consensus: high vs. low) mixed model analysis of variance (ANOVA) was conducted on the racist evaluations. Consistent with Study 1, there was a main effect for target, $F(1, 265) = 43.43, p < .001, \eta_p^2 = .141$; participants rated themselves as significantly less racist ($M = -1.53, SD = 1.73$) than the comparison other ($M = -0.84, SD = 1.83$). In regards to the consensus manipulation, there was no main effect of consensus, $F(1, 265) = .144, p = .705, \eta_p^2 = .001$, nor a target \times consensus interaction, $F(1, 265) = 3.48, p = .063, \eta_p^2 = .013$ (Figure 3). Thus, the effect persisted even when people were given high consensus information that suggested racist behaviors are common.

It is possible the effect of consensus would be stronger for people high in external and internal motivations to control prejudice, life satisfaction, trait self-esteem, public self-consciousness, and self-monitoring. Due to a data collection error, we only have data for these exploratory measures for 74 participants and are underpowered for the regression analysis. However, we analyzed potential moderating variables that might impact the consensus \times less-racist comparison effect for preliminary results and found that none of the potential moderators had a significant main effect or an interaction with consensus. See supplementary materials for more details.

General discussion

The present work provides evidence for the consistency and robustness of the biased self-enhanced evaluations of racism. Across three experiments, participants received a list of racist behaviors that purportedly were enacted from a fellow student but in fact were based on the participants' own behaviors. People consistently evaluated themselves as less racist than this comparison other, even though this other's racist behaviors were identical to their own. Studies 2a and 2b demonstrate this effect is quite robust and even occurs under conditions in which participants feel free to express their racial biases. Taken together, this work suggests that people are less likely to base their racist trait ratings on behavioral evidence when evaluating themselves compared to when they are evaluating another. By doing so, people are able to maintain the self-perception that they are not racist even in the face of contradictory behavioral evidence (i.e., people are less racist than themselves).

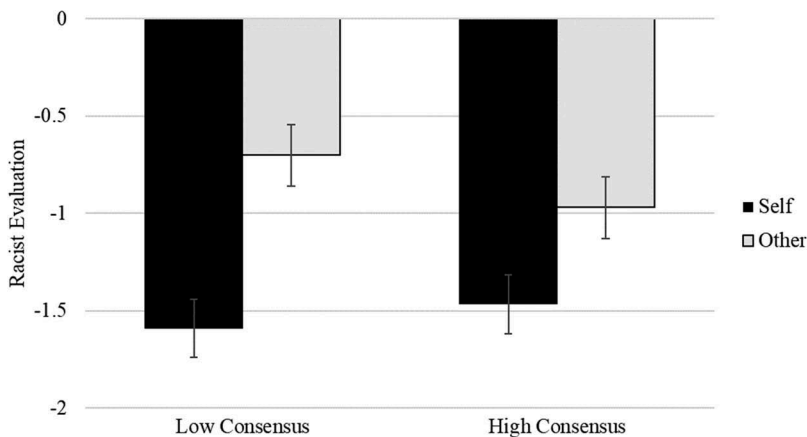


Figure 3. Racist evaluations for study 2b as a function of target (self vs. other) and consensus (low vs. high).

Limitations and future directions

A potential limitation to this research was that the prejudiced attitudes and behavior evaluation measures were vague enough that they applied to all ethnic minorities. People may view themselves as *more racist* than the average person (or themselves) when evaluating racist behaviors toward a single target group, compared to the Gestalt-approach of the current investigation (e.g., an individual could rate themselves as more racist towards African Americans but not more racist toward all ethnic/racial minorities). Additional measures for prejudice (i.e., feeling thermometers) should be included to assess participants' attitudes toward minorities belonging to specific ethnic groups, to see if there are any differences if the ethnic group is specified. Prior work shows that the more specific an assessment is, the more likely it will predict behavior (Fishbein & Ajzen, 1974), so the effects present in the present studies might be stronger if more specific ethnic groups were targeted in the measures. However, it may be that this effect is just too robust to be debiased by a simple in-lab manipulation. This effect is theorized to stem from the larger motivation of self-enhancement and because self-enhancement is so fundamental, it may be difficult to make the effect disappear.

Another potential limitation of the present studies is specific to Study 2b and concerns the large number of exploratory measures used. A data collection error limited our analyses for these measures to a small subsample and meant we were underpowered to examine their moderating effects on our manipulation. The number of exploratory measures that we used in Study 2b would normally require a statistical adjustment to the critical-alpha value, but because they were given to only a small subset of participants, they were omitted from our full sample analyses, and no such adjustment was made. Further, the subset of participants given these exploratory measures completed them *after* the critical measures reported in Study 2b, thereby removing potential confounding effects on our results when comparing those who did and did not complete the exploratory measures.

Now that the better-than-myself effect has been replicated and extended to the domain of prejudice, a logical next step in this line of research would be to identify the reasons why this effect occurs. The fact that the effect occurred even when people felt free to express racial bias suggests a general, heuristic-based process likely causes it, although more research is needed to confirm this assertion. Future research that does explore the exact mechanism that underlies this effect may again benefit by relying on prior BTAE and self-enhancement studies. For instance, Williams and Gilovich (2012) found that people are more likely to base their own trait estimates on their peak behavioral performances but base their estimates of others' traits on their average performances. The same may be true when it comes to judgments of racism. Participants may be more likely to think of a few instances when they behaved in an egalitarian manner and weight these instances more heavily when judging their own racism than when judging another's. Or, it may be that participants use situational attributions to discount their own racist behaviors but not those of others (i.e., self-serving bias; Miller & Ross, 1975; Shepperd, Malone, & Sweeny, 2008).

Future research should also explore factors that may moderate the asymmetry in self-other ratings of prejudice. The results of studies 2a and 2b suggest this effect is particularly robust and difficult to extinguish, which is not particularly surprising given that the effect is thought to be rooted in broader heuristics and motivations that are difficult to extinguish, such as self-enhancement and the BTAE (Sedikides, Gregg, & Hart, 2007). As a result, future researchers may want to explore factors that would strengthen, rather than weaken, this effect. For example, prior research shows that self-enhancement effects, including the BTAE, become stronger after a threat to self-esteem (Boney-McCoy, Gibbons, & Gerrard, 1999; Brown, 2012; Dunning, Leuenberger, & Sherman, 1995). Beyond self-enhancement, research on self-affirmation has found that negatively stereotyping out-groups buffers self-esteem (e.g., Fein & Spencer, 1997). It may be that self-esteem threats would increase self-enhanced comparisons as well. Prior research indicates that the BTAE is more likely to occur for highly important traits (e.g., honesty) rather than less important traits (e.g., imaginative; Brown, 2012). Therefore, it may be that self-enhanced evaluations of racism are more likely to occur for prejudices perceived as highly offensive (e.g., racism) rather than prejudices that are often perceived as less offensive (e.g., sexism; Czopp & Monteith, 2003) or are potentially more normalized (e.g., homophobia).

Future research could explore these and other motivational moderators, such as those related to the Justification-Suppression Model (JSM; Crandall et al., 2002). According to JSM, expression of prejudice is affected by motivational factors related to one's ability to justify and suppress said prejudice. As such, one is likely to suppress their prejudice when the social norms of a situation call behavior and when their prejudice cannot be easily justified. In the event that a motivation to suppress prejudice is absent, one may use a belief system to justify prejudice as a means to absolve shame. Studies 2a and 2b demonstrate that the social pressure and consensus conditions did not signal a social norm of tolerating racist behavior, thus failing to "release" participants from suppression by preventing the ability to justify of their own racist behaviors. While our two social motive manipulations did not affect the self-other evaluations, we were not exhaustive in testing the moderating effect of all other social motives. Selecting social motives that are informed by the JSM may allow future researchers to identify motives that specifically target suppression and/or justification mechanisms.

Implications

The present work adds to a larger body of research that examines the ways that people maintain positive self-perceptions in the face of evidence to the contrary (Critcher et al., 2011; Sedikides & Alicke, 2012). Our investigation into this phenomenon is important for a number of reasons. First, it is vital to understand how people label themselves and others as racist because doing so has important consequences. For instance, being labeled as racist has been shown to result in emotional consequences such as anxiety and guilt (Amodio, Devine, & Harmon-Jones, 2007; Monin & Miller, 2001; Sherman & Gorkin, 1980), financial consequences such as job loss (e.g., Paula Deen), and legal consequences such as discrimination lawsuits (McGeehan, 2013). However, not all of the consequences are negative. When people are labeled as racist, they often attempt to compensate for this label by behaving more egalitarian toward the minority members in question (Dutton & Lake, 1973; Dutton & Lennox, 1974; Sherman & Gorkin, 1980). If we want to encourage people to behave in an egalitarian manner, it is imperative to understand when and why people may or may not label themselves as racist in the face of behavioral evidence. The present work identifies one impediment that may prevent people from seeing their own racism in this way. Second, these studies demonstrate how research on biases in self-perceptions can offer theoretical and methodological guidance for prejudice researchers. Nearly 30 years of work on the better-than-average effect (including the better-than-myself effect) offer theoretical and methodological guidance for how researchers can study if people who engage in racist behaviors are reluctant to label themselves as racist. We suggest that researchers may want to consider other ways that the prejudice literature might overlap with the self-perception and self-enhancement literature. Few studies have examined such overlap, and many have focused on the role that self-esteem plays in prejudice (Aberson, Healy, & Romero, 2000; Crocker, Blaine, & Luhtanen, 1993; Fein & Spencer, 1997, 2000; Hogg & Abrams, 1990). However, as the present work suggests, there are likely other self-mechanisms that play a role in prejudice other than just self-esteem. Understanding why people fail to recognize their own racism—even when confronted with evidence of racism by their own definition—is a necessary step to reduce prejudice.

ORCID

Angela C. Bell  <http://orcid.org/0000-0001-9870-8645>

References

- Aberson, C. L., Healy, M., & Romero, V. (2000). Ingroup bias and self-esteem: A meta-analysis. *Personality and Social Psychology Review*, 4(2), 157–173. doi:10.1207/S15327957PSPR0402_04
- Alicke, M. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49, 1621–1630. doi:10.1037/0022-3514.49.6.1621
- Alicke, M., & Largo, E. (1995). The role of the self in the false consensus effect. *Journal of Experimental Social Psychology*, 31, 28–47. doi:10.1006/jesp.1995.1002

- Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). New York, NY: Psychology Press.
- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68(5), 804–825. doi:10.1037/0022-3514.68.5.804
- Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, 20, 1–48. doi:10.1080/10463280802613866
- Alicke, M. D., Vredenburg, D. S., Hiatt, M., & Govorun, O. (2001). The ‘better than myself effect. *Motivation and Emotion*, 25(1), 7–22. doi:10.1023/A:1010655705069
- Allport, G. W. (1937). *Personality: A psychological interpretation*. Oxford, England: Holt.
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, 18(6), 524–530. doi:10.1111/j.1467-9280.2007.01933.x
- Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74(3), 183–200. doi:10.1037/h0024835
- Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, 74(3), 183–200. doi:10.1016/S0065-2601(08)60024-6
- Boney-McCoy, S., Gibbons, F. X., & Gerrard, M. (1999). Self-esteem, compensatory self-enhancement, and the consideration of health risk. *Personality and Social Psychology Bulletin*, 25(8), 954–965. doi:10.1177/01461672992511004
- Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin*, 38(2), 209–219. doi:10.1177/0146167211432763
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review*, 90(2), 105–126. doi:10.1037/0033-295X.90.2.105
- Buzinski, S. G., & Kitchens, M. B. (2017). Self-regulation and social pressure reduce prejudiced responding and increase the motivation to be non-prejudiced. *The Journal of Social Psychology*, 157(5), 629–644. doi:10.1080/00224545.2016.1263595
- Cashen, L. H., & Geiger, S. W. (2004). Statistical power and the testing of null hypotheses: A review of contemporary management research and recommendations for future studies. *Organizational Research Methods*, 7(2), 151–167. doi:10.1177/1094428104263676
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. doi:10.1037/0022-3514.58.6.1015
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45(12), 1304–1312. doi:10.1037//0003-066x.45.12.1304
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. doi:10.1037/0033-2909.112.1.155
- Cortina, J. M., & Dunlap, W. P. (1997). On the logic and purpose of significance testing. *Psychological Methods*, 2(2), 161–172. doi:10.1037/1082-989X.2.2.161
- Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, 82, 359–378. doi:10.1037//0022-3514.82.3.359
- Critcher, C. R., Helzer, E. G., & Dunning, D. (2011). Self-enhancement via redefinition: Defining social concepts to ensure positive views of the self. In M. D. Alicke & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 69–91). New York, NY: Guilford Press.
- Crocker, J., Blaine, B., & Luhtanen, R. (1993). Prejudice, intergroup behaviour and self-esteem: Enhancement and protection motives. In M. A. Hogg & D. Abrams (Eds.), *Group motivation: Social psychological perspectives* (pp. 52–67). Hertfordshire, HP2 7EZ: Harvester Wheatsheaf.
- Cross, P. K. (1977). Not can, but will college teachers be improved? *New Directions for Higher Education*, 17, 1–15. doi:10.1002/he.36919771703
- Czopp, A. M., & Monteith, M. J. (2003). Confronting prejudice (literally): Reactions to confrontations of racial and gender bias. *Personality and Social Psychology Bulletin*, 29(4), 532–544. doi:10.1177/0146167202250923
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126. doi:10.1037/0022-3514.44.1.113
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49, 71–75. doi:10.1207/s15327752jpa4901_13
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment. *Psychological Science*, 5, 69–106. doi:10.1111/j.1529-1006.2004.00018.x
- Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces? *Journal of Personality and Social Psychology*, 69(1), 58–68. doi:10.1037/0022-3514.69.1.58
- Dutton, D. G., & Lake, R. A. (1973). Threat of own prejudice and reverse discrimination in interracial situations. *Journal of Personality and Social Psychology*, 28(1), 94–100. doi:10.1037/h0035582

- Dutton, D. G., & Lennox, V. L. (1974). Effect of prior "token" compliance on subsequent interracial behavior. *Journal of Personality and Social Psychology*, 29(1), 65–71. doi:10.1037/h0035719
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments & Computers*, 28(1), 1–11. doi:10.3758/BF03203630
- Faber, J. (2010). 'Kramer' apologizes, says he's not racist. *CBSNews.com*. Retrieved July 7, 2015, from <http://www.cbsnews.com/news/kramer-apologizes-says-hes-not-racist/>
- Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, 73(1), 31–44. doi:10.1037/0022-3514.73.1.31
- Fein, S., & Spencer, S. J. (2000). Prejudice as self-image maintenance: Affirming the self through derogating others. In C. Stangor (Ed.), *Psychology Press*. New York, NY.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43(4), 522–527.
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555–561. doi:10.1177/1745691612459059
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. doi:10.1177/001872675400700202
- Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, 81, 59–74. doi:10.1037/h0035872
- Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science*, 2(1), 31–39. doi:10.1111/j.1467-9280.1991.tb00093.x
- Giladi, E. E., & Klar, Y. (2002). When standards are wide of the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. *Journal of Experimental Psychology: General*, 131, 538–551. doi:10.1037/0096-3445.131.4.538
- Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, 69, 227–236.
- Goetz, T., Ehret, C., Jullien, S., & Hall, N. C. (2006). Is the grass always greener on the other side? Social comparisons of subjective well-being. *The Journal of Positive Psychology*, 1(4), 173–186. doi:10.1080/17439760600885655
- Greenwald, A. G. (1993). Consequences of prejudice against the null hypothesis. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 419–448). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. doi:10.1037/h0076157
- Guenther, C. L., & Alicke, M. D. (2010). Deconstructing the better-than-average effect. *Journal of Personality and Social Psychology*, 99(5), 755–770. doi:10.1037/a0020959
- Heene, M. (2013). Additive conjoint measurement and the resistance toward falsifiability in psychology. *Frontiers in Psychology*, 4. doi:10.3389/fpsyg.2013.00246
- Hogg, M. A., & Abrams, D. (1990). Social motivation, self-esteem and social identity. In D. Abrams & M. A. Hogg (Eds.), *Social identity theory: Constructive and critical advances* (pp. 28–47). New York, NY: Springer-Verlag.
- Howell, J. L., & Ratliff, K. A. (2017). Not your average bigot: The better-than-average effect and defensive responding to implicit association test feedback. *British Journal of Social Psychology*, 56(1), 125–145. doi:10.1111/bjso.12168
- Klein, S. B., Cosmides, L., Murray, E. R., & Tooby, J. (2004). On the acquisition of knowledge about personality traits: Does learning about the self engage different mechanisms than learning about others. *Social Cognition*, 22(4), 367–390. doi:10.1521/soco.22.4.367.38295
- Klein, S. B., Loftus, J., & Sherman, J. W. (1993). The role of summary and specific behavioral memories in trait judgments about the self. *Personality and Social Psychology Bulletin*, 19(3), 305–311. doi:10.1177/0146167293193007
- Klein, S. B., Loftus, J., Trafton, J. G., & Fuhrman, R. W. (1992). Use of exemplars and abstractions in trait judgments: A model of trait knowledge about the self and others. *Journal of Personality and Social Psychology*, 63, 739–753. doi:10.1037/0022-3514.63.5.739
- Klein, S. B., Sherman, J. W., & Loftus, J. (1996). The role of episodic and semantic memory in the development of trait self-knowledge. *Social Cognition*, 14(4), 277–291. doi:10.1521/soco.1996.14.4.277
- Klein, W., & Goethals, G. (2002). Social reality and self-construction: A case of "bounded irrationality?". *Basic and Applied Social Psychology*, 24(2), 105–114. doi:10.1207/S15324834BASP2402_3
- Malle, B. F. (2005). Self-other asymmetries in behavior explanations: Myth and reality. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgments* (pp. 155–180). New York, NY: Psychology Press.
- Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The Journal of Finance*, 60(6), 2661–2700. doi:10.1111/j.1540-6261.2005.00813.x
- McGehean, P. (2013). Merrill Lynch in big payout for bias case. *NYTimes.com*. Retrieved July 7, 2015, from <http://dealbook.nytimes.com/2013/08/27/merrill-lynch-in-big-payout-for-bias-case/>
- Miller, D. T., & Ross, M. (1975). Self-serving biases in attribution of causality: Fact or fiction? *Psychological Bulletin*, 82, 213–225. doi:10.1037/h0076486
- Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, 81(1), 33–43. doi:10.1037/0022-3514.81.1.33

- Monteith, M. J., Deneen, N. E., & Tooman, G. D. (1996). The effect of social norm activation on the expression of opinions concerning gay men and blacks. *Basic and Applied Social Psychology*, 18, 267–288. doi:10.1207/s15324834basp1803_2
- Moraski, L. (2013). Paula Deen on “Today”: I’m not a racist. *CBSNews.com*. Retrieved July 6, 2015, from <http://www.cbsnews.com/news/paula-deen-on-today-im-not-a-racist/>
- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, 5(2), 241–301. doi:10.1037/1082-989X.5.2.241
- O’Brien, L. T., Crandall, C. S., Horstman-Reser, A., Warner, R., Alsbrooks, A., & Blodorn, A. (2010). But I’m no bigot: How prejudiced White Americans maintain unprejudiced self-images. *Journal of Applied Social Psychology*, 40(4), 917–946. doi:10.1111/j.1559-1816.2010.00604.x
- Odean, T. (1998). Volume, volatility, price, and profit when all traders are above average. *The Journal of Finance*, 53(6), 1887–1934. doi:10.1111/0022-1082.00078
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, 94(1), 16–31. doi:10.1037/0022-3514.94.1.16
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, 75, 811–832. doi:10.1037/0022-3514.75.3.811
- Pronin, E., Gilovich, T., & Lee, R. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, 111(3), 781–799. doi:10.1037/0033-295X.111.3.781
- Puhl, R. M., Schwartz, M. B., & Brownell, K. D. (2005). Impact of perceived consensus on stereotypes about obese people: A new approach for reducing bias. *Health Psychology*, 24(5), 517–525. doi:10.1037/0278-6133.24.5.517
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Saucier, D. A. (2002). Self-reports of racist attitudes for oneself and for others. *Psychologica Belgica*, 42, 99–105.
- Scheier, M. F., & Carver, C. S. (1985). The self-consciousness scale: A revised version for use with general populations. *Journal of Applied Social Psychology*, 15(8), 687–699. doi:10.1111/j.1559-1816.1985.tb02268.x
- Sechrist, G. B., & Milford, L. R. (2007). The influence of social consensus information on intergroup helping behavior. *Basic and Applied Social Psychology*, 29(4), 365–374. doi:10.1080/01973530701665199
- Sechrist, G. B., & Stangor, C. (2001). Perceived consensus influences intergroup behavior and stereotype accessibility. *Journal of Personality and Social Psychology*, 80(4), 645–654. doi:10.1037/0022-3514.80.4.645
- Sedikides, C., & Alicke, M. D. (2012). Self-enhancement and self-protection motives. In R. M. Ryan (Ed.), *The oxford handbook of human motivation* (pp. 303–322). New York, NY: Oxford University Press.
- Sedikides, C., Gregg, A. P., & Hart, C. M. (2007). The importance of being modest. In C. Sedikides & S. J. Spencer (Eds.) *Psychology Press*. New York, NY.
- Shepperd, J., Malone, W., & Sweeny, K. (2008). Exploring causes of the self-serving bias. *Social and Personality Psychology Compass*, 2(2), 895–908. doi:10.1111/j.1751-9004.2008.00078.x
- Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology*, 16(4), 388–403. doi:10.1016/0022-1031(80)90030-X
- Sinclair, H. (2017). Man pictured at Charlottesville rally wants you to know he’s ‘not an angry racist,’ just marches with them. *Newsweek*. Retrieved January 3, 2017, from <http://www.newsweek.com/man-pictured-charlottesville-rally-wants-you-know-hes-not-angry-racist-650236>
- Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 30, 526–537. doi:10.1037/h0037039
- Stangor, C., Sechrist, G. B., & Jost, J. T. (2001). Changing racial beliefs by providing consensus information. *Personality and Social Psychology Bulletin*, 27(4), 486–496. doi:10.1177/0146167201274009
- Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, 11(5), 159–163. doi:10.1111/1467-8721.00191
- Tabachnik, N., Crocker, J., & Alloy, L. B. (1983). Depression, social comparison, and the false-consensus effect. *Journal of Personality and Social Psychology*, 45(3), 688–699. doi:10.1037/0022-3514.45.3.688
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193. doi:10.1037/0033-2909.103.2.193
- Vazire, S. (2010). Who knows what about a person? The self–Other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98(2), 281–300. doi:10.1037/a0017908
- Williams, E. F., & Gilovich, T. (2012). The better-than-my-average effect: The relative peak and average performances in assessments of the self and others. *Journal of Experimental Social Psychology*, 48(2), 556–561. doi:10.1016/j.jesp.2011.11.010
- Windschitl, P. D., Conybeare, D., & Krizan, Z. (2008). Direct-comparison judgments: When and why above- and below-average effects reverse. *Journal of Experimental Psychology: General*, 137(1), 182–200. doi:10.1037/0096-3445.137.1.182