A Meta-Analysis of Change in Implicit Bias

July 1, 2017

Patrick S. Forscher*[1], Calvin K. Lai*[2], Jordan R. Axt[3], Charles R. Ebersole[3], Michelle Herman[4], Patricia G. Devine[5], and Brian A. Nosek[3,6]

[1]Dept. of Psychological Science, University of Arkansas, [2]Dept. of Psychological & Brain Sciences, Washington University in Saint Louis, [3]Dept. of Psychology, University of Virginia, [4]Dept. of Psychology, University of North Carolina at Chapel Hill, [5]Dept. of Psychology, University of Wisconsin – Madison, [5]Center for Open Science, Charlottesville, VA.

*Authors contributed equally to this manuscript. Order was determined by coin flip.

**Author notes:**

Conceived research: Forscher & Lai, Devine, Nosek; Designed research: Forscher & Lai; Coordinated data collection: Lai; Coded articles: Forscher & Lai, Axt, Ebersole, Herman; Analyzed data: Forscher; Wrote paper: Forscher & Lai; Revised paper: all authors.

We thank Katie Lancaster, Diana Abrego, Amy Bisker, Isabelle Gigante, Julie Lee, Lauren Loffredo, Ryan Massopust, Margot Mellon, Kelci Straka, and Nicole Sather for their assistance with the early phases of article coding. We also thank Mike W.-L. Cheung and Ian White for their assistance and advice about the analyses presented in this paper.

Address correspondence to Calvin K. Lai, Dept. of Psychological & Brain Sciences, Washington University in Saint Louis, One Brookings Drive, Saint Louis, MO, 63130 (Email: cklai4@gmail.com) and/or Patrick S. Forscher, Dept. of Psychological Science, University of Arkansas, 216 Memorial Hall, Fayetteville, AR 72701 (Email: schnarrd@gmail.com).

Data and materials for this project can be found at https://osf.io/awz2p/

## Abstract

Using a technique known as network meta-analysis that is new to psychological science, we synthesized evidence from 494 studies (80,356 participants) to investigate the effectiveness of different procedures to change implicit bias, and their effects on explicit bias and behavior. We found that implicit bias can be changed, but the effects are often weak ($|ds| < .30$). Procedures that associate sets of concepts, invoke goals or motivations, or tax mental resources changed implicit bias the most, whereas procedures that induced threat, affirmation, or specific moods/emotions changed implicit bias the least. Most procedures were brief and were tested within a single experimental session, and funnel plot analyses suggested that the effects could be inflated relative to their true population values. Many procedures changed explicit bias, but to a smaller degree than they changed implicit bias. We found no evidence of change in behavior. Finally, changes in implicit bias did not mediate changes in explicit bias or behavior. Our findings suggest that changes in measured implicit bias are possible, but those changes do not necessarily translate into changes in explicit bias or behavior. We discuss potential interpretations of these findings, including the possibility that current manipulations change non-associative aspects of implicit measures and the possibility that the automatic retrieved associations do not influence explicit biases or behavior.


Keywords: meta-analysis, implicit bias, intervention, social cognition

## A Meta-Analysis of Change in Implicit Bias

What we intend to do often conflicts with what we actually do. We may plan to diet but find ourselves reaching for a chocolate bar over an apple. Or, we might try to quit smoking but find the temptation of cigarettes too difficult to resist. We can even value racial equality but choose to hire a White job candidate over a similarly qualified Black job candidate (Bertrand & Mullainathan, 2004). These gaps between what we intend and what we actually do characterize many societal problems, such as intergroup discrimination (Devine, 1989), depression (Beevers, 2005; Haeffel et al., 2007), and addiction (Wiers et al., 2010).

The prevalence of unwanted behaviors across many areas of human life suggests that mental processes outside of one's conscious awareness or control influence behavior (Smith & DeCoster, 2000). Based on this reasoning, researchers have developed dual-process theories that distinguish between automatic mental processes which are relatively fast, efficient, and unintentional, and deliberate mental processes which are relatively slow, controlled, and intentional. By this logic, the same underlying mental construct can be retrieved either automatically or deliberately. For example, the association between the concepts "Flowers" and "Good" can be retrieved automatically, as when a person spots a vase of flowers and feels good, or deliberately, as when a person is asked to think about how much they like flowers.

Dual process theories posit that deliberate processes are more influential on behavior when people have sufficient motivation, awareness, and the ability to reflect before acting, whereas automatic processes are more influential when motivation, awareness, or the ability to reflect are compromised (Devine, 1989; Fazio & Olson, 2014; cf. Greenwald et al., 2009). Dual process theories also predict that dissociations between intentions and behavior are most likely to occur when the output of automatic and deliberate processes are opposed. Given opposing automatic and deliberate processes, lack of motivation, awareness, or the ability to reflect can cause people to act against their intentions.

Dual process theories are attractive on both theoretical and practical grounds. Theoretically, they provide a parsimonious approach for explaining dissociations between intentions and behavior and between mental phenomena more broadly. Dual process theories are used to account for such wide-ranging phenomena as attention (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), reasoning (Evans, 1989; Sloman, 1996; Stanovich & West, 2000), decision-making (Barbey & Sloman, 2007; Kahneman, 2011), memory (Jacoby & Dallas, 1981; Roediger, 1990), attitudes (Wilson, Lindsey, & Schooler, 2000), stereotypes and prejudice (Devine, 1989), the self (Schnabel & Asendorpf, 2010), motivation (Chartrand & Bargh, 2002), and emotion regulation (Mauss, Bunge, & Gross, 2007). Practically, dual-process theories suggest a solution to problems caused by unintentionally biased behavior: change the automatic processes and changes in behavior will follow (Lai, Hoffman, & Nosek, 2013; Forscher & Devine, 2014).

Implicit and explicit measures of mental associations between concepts have been a particular interest for dual process theorists.[1] Implicit measures assess associations through a

[1] In the present article, we describe implicit and explicit measures as assessing an association that is retrieved automatically or deliberately. We are theoretically uncommitted to whether implicit and explicit measures assess a common representation or categorically different representations, and whether the measures are assessing stored

behavior that does not require deliberate retrieval of the target association (e.g., the speed of sorting words into different categories relevant to the association). In contrast, explicit measures assess the target association through a behavior that does require deliberate retrieval (e.g., answers to a questionnaire). On an implicit measure, comparisons between behavior that results from pairings between one set of concepts relative to the resulting behavior from a different pairing is termed *implicit bias*; similar comparisons on explicit measures are termed *explicit bias*. For example, differences in the time to identify the word "flower" when it is preceded by the words "good" vs. "bad" can serve as an indicator of implicit bias, whereas differences in the ratings from ratings of the degree to which flowers are good and bad can serve as an indicator of explicit bias.

Implicit and explicit biases are assumed to indicate the presence of automatically or deliberately retrieved associations, respectively. However, like all psychological measures, implicit and explicit measures are not process-pure. Implicit and explicit measures are both prone to measurement error, and implicit and explicit biases can be influenced by many processes other than automatically and deliberately retrieved associations (e.g., task-switching ability and impulse inhibition for implicit measures, social desirability and acquiescence bias for explicit measures; Calanchini et al., 2013; Conrey et al., 2005; Cronbach, 1946; Blanton et al., 2006; Crowne & Marlowe, 1960).

Implicit and explicit biases are correlated, but the extent to which they correlate varies (Cameron, Brown-Iannuzzi, & Payne, 2012; Greenwald et al., 2009; Hofmann et al., 2005; Nosek & Hansen, 2008). These correlations range from very low ($r$ = .07; e.g., attitudes toward approaching vs. avoiding) to very high ($r$ = .70; e.g., attitudes toward Democrats vs Republicans; Nosek & Hansen, 2008). Half of the variation in implicit-explicit relations can be accounted for with four aspects of the social and mental context: the social sensitivity of the target concepts, the extent to which people have thought about the concepts, the degree to which the concepts in the implicit measures are diametrically opposed (e.g., pro-choice vs. pro-life) or not (e.g., dog vs. furniture), and the degree to which people view their opinions about the concepts to be distinct from others (Nosek, 2005; 2007). The predictability of the relation between implicit and explicit bias measures suggest underlying processes that are causally related and/or influenced by third variables.

When automatic and deliberate processes are not aligned discrepancies between intentions and behavior may arise, such as intending to be unbiased in selection of candidates for an honor society but showing racial discrimination anyway (Axt, Ebersole, & Nosek, 2016). Consistent with dual process theories, some evidence suggests that implicit biases are more strongly correlated with behavior in socially sensitive issues (Greenwald et al., 2009; cf. Oswald et al., 2013), whereas explicit biases are relatively strongly correlated with behavior when the situation demands a more deliberate response (Devine, 1989; Fazio & Olson, 2014; cf. Greenwald et al., 2009). Alternatively, when automatic and deliberate processes are aligned, these processes mutually reinforce each other to guide behavior. Consistent with this claim,

---

representations or active constructions (Greenwald & Nosek, 2008). Likewise, we use "association" with a theory-uncommitted view (Greenwald et al., 2005). We do not assert a commitment to a particular understanding of what the underlying constructs or processes are (e.g., associative or propositional; Gawronski & Bodenhausen, 2006). Various accounts of the underlying constructs / processes can be adapted to accommodate the changes in implicit bias observed in the present meta-analysis.

behavior is most consistent with both implicit and explicit biases when implicit and explicit biases are more strongly correlated (Greenwald et al., 2009).

**Implicit bias change**

Of course, correlation is not causation, so understanding the causal importance of automatically retrieved associations first requires procedures that can change observed levels of implicit bias. At first, the prospect of changing implicit bias through randomized experiments was dim. Approaches such as cognitive dissonance reduction and persuasive appeals were successful changing self-reported attitudes but often had limited impact on implicit bias (for reviews, see Cooper, 2007; Gawronski & Strack, 2012; Petty & Cacioppo, 1986). The apparent rigidity of automatic processes led the social psychologist John Bargh to portray them as a "cognitive monster" (Bargh, 1999) that is deep-rooted, immune to social pressure, and resistant to the influences of deliberate processes.

Yet this understanding changed with the discovery that brief experiences can change implicit bias without affecting explicit bias (Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000). Over the past sixteen years, the accumulated evidence suggests that implicit biases can be changed, but doing so often relies on mechanisms that are ineffective for shifting explicit biases (for reviews, see Blair, 2002; Dasgupta, 2009; Gawronski & Bodenhausen, 2011; Gawronski & Sritharan, 2010; Lai et al., 2013; Lenton et al., 2009; Sritharan & Gawronski, 2010). For example, the mere presence of a Black experimenter changed implicit bias without affecting explicit bias (Sinclair, Lowery, Hardin, & Colangelo, 2005). More recently, some studies suggest that approaches that affect explicit bias can also affect implicit bias, such as intergroup contact, social threat, and cognitive balance (Bradley et al., 2012; Shook & Fazio, 2008; Smith, De Houwer, & Nosek, 2012). Further, some strategies highlight the process-impurity of implicit measures by changing aspects of performance in implicit measures that are unrelated to underlying associative processes (e.g., instruction to fake on an implicit measure; Kim, 2003; Fiedler & Bluemke, 2005).

Inspired by social problems characterized by unintentional or unwanted behavior, many studies aim to change implicit bias with the goal of changing behavior. Many of these studies occur in domains, such as race relations or addiction, where automatic and deliberate associations are often thought to be at odds and where deliberate processes are either resistant to change or theorized to have a limited influence on behavior (e.g., Wiers et al, 2010; Mann & Kawakami, 2012). If interventions on deliberate processes are of limited utility, perhaps approaches that change automatically retrieved associations will be more effective.

Despite the proliferation of many approaches to changing implicit bias, little is known about their relative effectiveness (Lai et al., 2013; 2014; 2016). At the same time, there is also little understanding about what approaches are *consistently effective* across a wide range of phenomena, and what kinds of approaches are inconsistently effective and are contextually dependent on the population, study methodology, or topic of study. Advances in these areas of knowledge would inform a basic understanding of the mental mechanisms that are most influential on implicit biases and a practical understanding of what interventions would be most effective for changing implicit biases.

**Overview of present research**

We conducted a meta-analytic review to understand the relative effectiveness of different procedures to change implicit bias and whether the changes in implicit bias generalize to changes in explicit bias and behavior. The diversity in research goals means that research on implicit bias change spans many disciplines, theoretical perspectives, and methodological approaches. Study designs range from two-condition single-session laboratory experiments (e.g., Rudman & Lee, 2002) to multiple-condition longitudinal studies (Sportel, de Hullu, de Jong, & Nauta, 2013). They also differ in what kinds of manipulations are used, from minimal manipulations that prime a concept in memory (Dasgupta & Greenwald, 2001) to intensive long-term interventions that unfold over several weeks (O'Brien et al., 2010). The studies are also diverse in their dependent measures, ranging from popular measures such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998; Nosek, Greenwald, & Banaji, 2007) to less popular measures such as the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart, 2010; Hussey et al., 2015).

This research diversity poses two unique analytic issues for meta-analysis. First, different studies often compare different sets of procedures. The diversity in procedures is a challenge for conventional meta-analytic methods that synthesize two-group studies because conventional methods assume all studies use a common comparison. Second, studies in this literature sometimes compare the effects of three or more procedures within the same design. Conventional meta-analytic methods assume that each effect size is independent and thus cannot accommodate these non-independent comparisons.

We used a technique from the medical sciences called multivariate network meta-analysis to address these issues (Lu & Ades, 2004; Caldwell, Ades, & Higgins, 2005; Salanti, 2012). Compared to conventional meta-analytic methods, network meta-analysis synthesizes information from many procedures simultaneously to better addresses research literatures where there are many studies that compare distinct procedures (Lumley, 2002). A multivariate implementation of network meta-analysis addresses the problem of single studies making multiple comparisons by modeling the non-independence between multiple comparisons extracted from the same study (White et al., 2012; Mavridis & Salanti, 2012). Multivariate network meta-analysis therefore allows us to use all information from studies comparing many procedures to change implicit bias, rather than having to simplify the information available when a study has more than one possible contrast (e.g., via averaging, dummy-codes, or data exclusions).

Our meta-analysis of procedures for changing implicit bias was guided by 5 central questions:

1. What approaches to changing implicit bias are most influential? We developed a taxonomy of procedures to change implicit bias and compared the effectiveness of procedures within that taxonomy.
2. Are the sample, methodology, or topic of a study associated with the magnitude of implicit bias change? We assessed whether any of these characteristics were associated with the degree of implicit bias change.

3. How do changes in implicit bias correspond with changes in explicit bias? We compared the relative size of explicit bias change to implicit bias change. We also examined whether implicit bias change mediated explicit bias change.
4. How do changes in implicit bias correspond with changes in behavior? We compared the relative size of behavioral change to implicit bias change. We also examined whether implicit bias change mediated behavioral change.
5. Is there evidence that the size of reported effects is biased? We examined whether reported effect sizes are likely to be larger than the true effect sizes (e.g., due to publication bias or *p*-hacking; Rosenthal, 1979; Simmons, Nelson, & Simonsohn, 2011).

# Method

## Inclusion criteria

Valid meta-analysis requires careful consideration of which studies are relevant to the research question and which studies are not. We set the following inclusion criteria:

(1) *The study is a between-subjects experiment*. We excluded studies that used correlational or quasi-experimental designs (e.g., Rudman, Ashmore, & Gary, 2001) and manipulations that were exclusively within-subjects (e.g., Wheeler & Fiske, 2005). We also excluded studies that experimentally manipulated the stimuli or categories in an implicit measure (e.g., by manipulating whether pictures of animals and plants in an animal/plant pleasant/unpleasant IAT are positively or negatively valenced; Govan & Williams, 2004) because the conditions assessed categorically different associations rather than changing a particular set of associations.

(2) *The study includes an implicit measure that is administered after the onset of the experimental manipulation.* Implicit measures were defined as measures of associations between concepts that do not require the participant to actively bring to mind the target association. This definition included measures that are both widely used (e.g., the IAT; Greenwald, et al., 1998; Nosek et al., 2007) and less widely used (e.g., Stereotypic Explanatory Bias, Sekaquaptewa et al., 2003). Measures for which the manipulation began during task instructions or practice trials (e.g., Foroni & Mayr, 2005) and for which the manipulation extended into the measure (e.g., Huntsinger et al., 2010) were also considered eligible.

(3) *The implicit measure assesses a pre-existing association*. We defined a "pre-existing association" as an association that either should theoretically be present, or has been empirically demonstrated as present, within the target population before the study began. For example, most non-Black people implicitly associate Black people with bad and White people with good more easily than the reverse (Nosek, Smyth, et al., 2007).[2]

[2] In making these decisions, we assumed that people tend to associate positive attributes with both themselves and with their own groups, and that people tend to possess associations that are commonly present in their culture (e.g., Black people with the attribute "musical"). When we could not make a clear determination, we sought data collected from the target population and/or examined whether a pre-existing association was present in a control condition for the study in question.

Based on the nature of the pre-existing association, we defined pairings that strengthen (e.g., Black-bad and White-good) and weaken (e.g., Black-good and White-bad) the measured association. We excluded studies that formed a new association (e.g., about fictitious people or social groups, McConnell, Rydell, Strain, & Mackie, 2008) and studies of ambivalent or unelaborated associations (e.g., Petty et al., 2006) based on this criterion.

(4) *The study is reported in English*. We excluded studies that were not written in English.

**Article retrieval**

Our article retrieval procedure was conducted in three phases between September 2012 and July 2015 and again between August 2016 and October 2016. In the first phase (September 2012 to June 2014; August 2016), we retrieved articles that potentially matched our inclusion criteria. We searched PsycINFO and Web of Science using the following search terms: (names of implicit constructs, measures, and acronyms, e.g., implicit self-esteem*, affect misattribution procedure, GNAT) AND (malleab* OR chang* OR influenc* OR moderat* OR reduc* OR increas* OR shift* OR alter*) AND (1995 TO 2015). We created the list of eligible implicit measures and acronyms by compiling lists from published reviews of implicit measurement (Nosek, et al., 2011; Gawronski & Payne, 2010), and from discussions among the lead authors (for the full list, see https://osf.io/awz2p/). We supplemented these results with direct requests for relevant studies through email and the Society for Personality and Social Psychology listserv, and an additional 115 articles from an unpublished meta-analysis of the malleability of implicit intergroup bias. Our search procedure resulted in approximately 5,238 articles that potentially matched our inclusion criteria.[3]

In the second phase (September 2012 to October 2014; August 2016 to October 2016), trained coders inspected each article and eliminated articles that did not contain a study matching our inclusion criteria. This thinned our database to 418 articles, 594 studies, and 692 independent samples.

Finally, for any studies that did not report sufficient data to calculate effect sizes and variance components, we sent emails to the corresponding authors requesting the required statistics (November 2014 to July 2015; October 2016). If the authors did not respond, we sent two follow-up reminder emails. If the data required to calculate effect sizes on the implicit measure could not be retrieved for a study, we eliminated that study from the meta-analysis. After eliminations, our final sample represented 80,356 participants and included 343 articles, 494 studies, and 573 independent samples.

**Article coding**

---

[3] This number reflects the number of articles from searches of PsycINFO and Web of Science. Web of Science yielded 4,979 articles and PsycINFO yielded 4,161 articles, most of which were also in Web of Science. Articles retrieved through email requests were not tracked systematically, meaning that the actual number of potential articles is somewhat higher.

Coders underwent extensive training to reliably apply the coding scheme. We adopted an iterative process to maximize reliability and validity of the coding scheme (Lipsey & Wilson, 2001). When coders encountered an ambiguity, they added the ambiguity to the agenda for a weekly coding meeting. During these meetings, we discussed each ambiguity until we reached a consensus for resolution. Some ambiguities revealed issues with the coding scheme. In these cases, we revised the coding scheme and rolled out any required coding changes to all other studies. We have made our coding scheme, data, and analysis scripts publicly available (https://osf.io/awz2p/). Anyone who is interested can delve into these materials to assess how the results change with different coding decisions.

We tested the reliability of our coding scheme by choosing a random sample of 50 fully coded articles and assigning each coder 10 articles to double-code. The random sample was stratified by topic to ensure that each coder received articles that varied in topic and coding difficulty. Where appropriate, Cohen's $\kappa$ calculated using these data is shown in the descriptions below.

*Experimental procedures* ($\kappa = .71$). Each experimental procedure was categorized into one of fourteen categories. We developed these categories based on preliminary searches of the literature and prior reviews of malleability and change in implicit bias (e.g., Blair, 2002; Dasgupta, 2009; Gawronski & Bodenhausen, 2011; Gawronski & Sritharan, 2010; Lai et al., 2013; Lenton et al., 2009; Sritharan & Gawronski, 2010) with the goal of capturing the breadth of approaches that researchers have employed. Two of the fourteen categories (physiological deprivation and satiation) were excluded from the final dataset because there were not enough procedures that fit the description (four and two procedures respectively across four papers).

Researchers often disagree about the likelihood and mechanism by which a manipulation might change implicit bias. To address this issue and maximize agreement between coders, our coding scheme prioritized procedural elements of the study conditions over theoretical expectations regarding the impact of these procedural elements. For example, conditions from two studies that both give participants instructions to show no bias on an IAT would be placed same category, regardless of whether the authors of the studies differ in their predictions as to whether this condition would produce change in IAT scores (e.g., Kim, 2003; Fiedler & Bluemke, 2005). If a given experimental condition fit into multiple coding categories or did not find into a category clearly, that condition was excluded from the meta-analysis. As shown in Table 1, our final coding scheme included twelve categories:

*(1) Strengthen associations directly* ($k = 128$) / *Weaken associations directly* ($k = 155$). Some efforts to change implicit biases create experiences that directly affirm or counter one's own biases (e.g., Blair et al., 2001; Dasgupta & Greenwald, 2001). These two categories involve creating pairings of the concepts used in the implicit measure that either strengthen or weaken the measured implicit bias. For example, exposing people to pictures of admired Black people and despised White people in a study assessing associations between Black people/White people and good/bad would go in the "weaken associations directly" category. In contrast, exposing people to admired White people and disliked Black people would go in the "strengthen associations directly" category (Dasgupta & Greenwald, 2001).

*Table 1.* Taxonomy of experimental procedures.

| Procedure | Samples | Description | Examples |
|---|---|---|---|
| Weaken associations directly<br>Strengthen associations directly | 155<br>128 | Direct pairing of concepts in implicit measure | Evaluative conditioning *(Olson & Fazio, 2006)*<br>Persuasive argument *(Horcajo et al., 2010)*<br>Counterstereotypical exemplars *(Dasgupta & Greenwald, 2001)* |
| Weaken associations indirectly<br>Strengthen associations indirectly | 154<br>86 | Activating ideas/mindsets not in implicit measure | Perspective-taking for attitudes *(Todd et al., 2011)*<br>Inducing feelings of power *(Guinote et al., 2010)*<br>Approach/avoid training for attitudes *(Kawakami et al., 2008)* |
| Goals to weaken bias<br>Goals to strengthen bias | 92<br>37 | Inducing a goal related to implicit measure | Implementation intentions *(Stewart & Payne, 2008)*<br>Making anti-prejudiced norms salient *(Wyer, 2010)*<br>Subtly priming a goal *(Ferguson, 2008)* |
| Threat | 72 | Putting one's identity at risk | Mortality salience *(Jong et al., 2002)*<br>Giving a speech *(Rabbitt, 2012)*<br>Stereotype threat *(Frantz et al., 2004)* |
| Affirmation | 23 | Maintaining adequacy of one's identity | Self-affirmation *(Rudman et al., 2007)*<br>Group affirmation *(Peach et al., 2011)*<br>Success feedback *(Brown, 2010)* |
| Positive affective state<br>Negative affective state | 26<br>27 | Positive/negative moods or emotions | Listening to happy/sad music *(Birch et al., 2008)*<br>Watching a funny movie *(Cain, 2012)*<br>Writing about a disgusting event *(Dasgupta et al., 2009)* |
| Depletion | 26 | Depletion of mental resources | Cognitive load *(Allen et al., 2009)*<br>Thought suppression *(Hooper et al., 2010)*<br>Ego-depletion *(Govorun & Payne, 2006)* |
| Neutral | 429 | No features relevant to implicit measure | Baseline control condition<br>Exposure to unrelated stimuli *(Dasgupta & Rivera, 2008)* |

*(2) Strengthen associations indirectly* ($k$ = 86) / *Weaken associations indirectly* ($k$ = 154). A related approach to the first category is creating experiences that bring to mind an idea or mindset that will indirectly affirm or counter one's biases (Blair, 2002). These categories were similar to the "strengthen / weaken associations directly" categories except that the concepts used to create the pairings are not the ones used in the implicit measure, but instead are assumed to lead to their mental activation through some intermediate idea or mindset. For example, taking the perspective of a Black person is theorized to create overlap between a person's self-concept and Black people (Galinsky & Moskowitz, 2000;

Todd et al., 2011). As most people evaluate themselves positively (Taylor & Brown, 1988), linking Black people to the self creates an indirect link between Black people and positivity that changes implicit racial attitudes. Other examples include taking an abstract construals to associate a temptation with negativity (Fujita & Han, 2009) and changing approach/avoid tendencies to change implicit attitudes toward math versus arts (Kawakami, Steele, Cifa, Phills, & Dovidio, 2008).

*(3) Goals to strengthen bias* ($k$ = 37) */ Goals to weaken bias* ($k$ = 92). Implicit biases may be sensitive to motivations, goals, and habits (e.g., Fishbach, Friedman, & Kruglanski, 2003; Sinclair et al., 2005). Procedures in these categories gave participants goals to respond on an implicit measure in ways that either strengthen or weaken the measured implicit bias. These goals could be created directly, such as by instructing participants to appear non-shy on a measure of implicit shy/non-shy self-concept (Asendorpf, Banse, & Mucke, 2002). These goals could also created indirectly, such as by making anti-prejudiced norms salient prior to measuring implicit bias toward Black people (Wyer, 2010).

*(4) Threat* ($k$ = 72). Threat involves putting the integrity of a person's identity at risk. Threat plays a powerful role in shifting attention (Mogg, Bradley, De Bono, & Painter, 1997), evaluations of one's self (Taylor & Lobel, 1989), and evaluations of others (Stephan & Stephan, 2000). The threats included in this category were diverse, including the threat of confirming a negative stereotype (e.g., Frantz et al., 2004), mortality salience (e.g., Jong, Halberstadt, & Bluemke, 2012), and the threat of giving a speech in front of a panel of judges (e.g., Rabbitt, 2012).

*(5) Affirmation* ($k$ = 23). Affirmation involves procedures that sought to maintain the adequacy of a person's identity, which may buffer against acute or chronic experiences of threat (Cohen & Sherman, 2014; Steele, 1988). Examples in this category included procedures in which the participants were given feedback that they were competent, moral, or unbiased (Frantz et al., 2004), and procedures where the participants were instructed to think about a value important to a social group to which they belonged (Peach et al., 2011).

*(6) Positive affective state* ($k$ = 26) */ Negative affective state* ($k$ = 27). According to an affect-as-information account, positive affect affirms chronically accessible concepts and negative affect rejects them (Huntsinger, Isbell, & Clore, 2014). These categories involved procedures that induced a mood or emotion without placing the manipulation in the "threat" or "affirmation" categories. Although manipulations that threaten or affirm a person's identities are likely to induce affect, we reasoned that threat and affirmation are the primary characteristics of these conditions and take precedence.[4] Examples of manipulations in these categories included both positive or negative mood inductions (e.g., Birch et al, 2008) and inductions of specific emotions like anger, disgust, or moral elevation (Dasgupta, DeSteno, Williams, & Huntsinger, 2009; Lai, Haidt, & Nosek, 2014).

*(7) Depletion* ($k$ = 26). Depleting mental resources may lead to increased reliance on social-cognitive biases (e.g., Bodenhausen, 1990; Gilbert & Hixon, 1991; Stangor & Duan, 1991). This category involves manipulations that reduced the amount of mental resources

[4] We placed anger inductions into the positive affective state category as anger is more cognitively and neurally similar to positive emotions than negative ones (Lerner & Tiedens, 2006; Harmon-Jones, 2003; Carver & Harmon-Jones, 2009).

available to the participant during the implicit measure. Conditions in this category often instructed people to complete a mentally effortful task, such as holding a multi-digit number in their heads (Allen et al., 2009), prior to or during the implicit measure.

*(8)* *Neutral* ($k$ = 429). This category involves conditions where nothing happened that was potentially relevant to the concepts assessed by the implicit measure (e.g., control conditions). This category did not necessarily contain conditions that a specific research tradition would predict are ineffective. For example, on the basis of past evidence (Dasgupta & Greenwald, 2001), some researchers might predict that exposure to images of admired White people and disliked Black people does little to change implicit racial attitudes because admired White people are already chronically accessible prior to the exposure experience. Although this may be the case, exposure to liked White people pairs White people with positivity, and thus this condition would be placed in the "strengthen associations directly" category.

*Implicit, explicit, and behavioral measures*. Measures were considered implicit if they did not require the target association to be actively brought to mind. For example, the Black/White good/bad IAT requires participants to categorize Black faces, White faces, positive words, and negative words, but it does not require them to introspect about their feelings about Black people relative to White people. Measures were considered explicit if they required the target association to be actively brought to mind. For example, a survey item asking "How warm do you feel toward Black people?" requires participants to actively assess their personal feelings about Black people. Measures were considered behavioral if they involved the participant's actual, hypothetical, or intended behavior in relation to the target association. Behavioral measures assessed a wide range of outcomes, such as seating distance from a Black or White confederate (Todd et al., 2011), willingness to participate in a hypothetical beer pong game (Goodall & Slater, 2010), intentions to drink in the future (Glock, Klapproth, & Müller, 2015; Lindgren et al., 2015), reported chocolate consumption (Kroese, Adriaanse, Evers, & De Ridder, 2011), and intentions to vote for gay and lesbian civil rights referenda (Dasgupta & Rivera, 2008).

Explicit and behavioral measures were included only if coders judged that they assessed the same association as the implicit measure selected from the study. For example, a questionnaire assessing Black stereotypes would be eligible for a measure of implicit Black/White stereotyping but not a measure of Black/White implicit attitudes. This inclusion criterion was notably stricter than past meta-analyses that included explicit/behavioral measures which did not narrowly tap into the same constructs (e.g., physiological or neural activity for IATs in Greenwald et al., 2009; stereotype measures for attitude IATs in Oswald et al., 2013). As with the implicit measures, explicit and behavioral measures were only eligible if they were administered after the onset of the manipulation. If multiple measures in a sample met our definition of an implicit, explicit, or behavioral measure, we selected the measure that was most widely used in the meta-analysis (i.e., if a study included both an IAT and a Lexical Decision Task assessing implicit self-esteem, we selected the IAT) or the measure that best matched the implicit measure conceptually (e.g., for a relative implicit measure of stereotypes, we prioritized relative explicit stereotyping measures over absolute stereotyping measures).

All measures were scored such that higher numbers represent greater levels of the pre-existing bias. Implicit measures that assessed associations between two sets of concepts were scored by creating a difference score that reflected the underlying association. For example, in a study where researchers measured participant reaction times (RT) to categorize positively and negatively valenced words with Black and White face primes, we created the following difference score: (Black/good RT - Black/bad RT) - (White/good RT - White/bad RT). If a score computed from a *D* score algorithm (Greenwald et al., 2003) was used, we chose that over a reaction time difference score. If the explicit and behavioral measures were composed of multiple parts (e.g., separate assessments of feelings of warmth toward Black people and White people), we scored these measures to be most correspondent with the implicit measure. In a study using the aforementioned priming measure that also contained separate feelings thermometer ratings of Black people and White people, we created the following difference score: White thermometer rating - Black thermometer rating.

*Multiple study subsamples*. If a study reported their results separately for groups with a given individual-difference characteristic (e.g., a median split of a questionnaire measure), we collapsed across the target individual difference. If, however, participants were recruited on the basis of that individual difference characteristic (e.g., from the top and bottom quartile of a scale), we treated these groups as separate subsamples for the purposes of the meta-analysis to avoid confounding (Glass, 1977). In some cases, we analyzed groups separately even if they were not recruited on a specific characteristic if the meaning of the measure or manipulation was unambiguously different for different subgroups. For example, the meaning of a Bill Clinton/George Bush good/bad IAT is likely different for Democrats and Republicans because Democrats share a party affiliation with Bill Clinton, whereas Republicans share a party affiliation with George Bush (Albertson, 2011). Finally, studies were split into subsamples if the study randomly assigned participants to different implicit measures in addition to randomly assigning them to different manipulations (e.g., by assessing the effects of reading a counter-stereotypical vs. neutral scenario on the personalized vs. original IAT, Han et al., 2010).

*Sample characteristics*

*Sample population* ($\kappa$ = .92). University student samples tend to be more compliant and more easily socially influenced (Sears, 1986), and may be more susceptible to psychological manipulations than non-student samples (e.g., Lai et al., 2016). Student and non-student samples may also differ because of issues related to the publication process (e.g., reviewers may be less critical of small effects if the study does not use an undergraduate convenience sample). To assess these possibilities, we coded whether the sample was drawn from a university student or a non-university-student population (e.g., hazardous drinkers, elementary school children).

*Demographic characteristics*. We coded the racial and gender distribution of each sample to examine the generalizability of results to different demographic groups. Coders recorded the number of participants who were White, non-White, or whose race was not reported. Coders followed a similar process for gender distribution: male, female, or gender not reported.

*Methodological characteristics*

*Design* (implicit $\kappa$ = .86; explicit $\kappa$ = .89; behavior $\kappa$ = .96). The effects of procedures on measured implicit bias may depend on whether participants completed an implicit measure before the intervention (e.g., Lai et al., 2014). Thus, we assessed whether the implicit, explicit, and behavioral measures were administered in a fully between-subjects design or in a mixed design with between-subjects and within-subjects (i.e., pre-test and post-test) components.

*Implicit measure* ($\kappa$ = .90). Different implicit measures may tap different constructs. Implicit measures also vary in measurement reliability, which can depress the relationship between manipulations and their effects on implicit bias. Thus, we coded whether a study's implicit measure was an IAT or not (e.g., the Affect Misattribution Procedure, Go/No-Go Association Task, etc.).

*Longitudinal* ($\kappa$ = .87). This variable assessed whether the implicit measure was administered longitudinally (i.e., at least one of the measurements occurred after a delay that is longer than one experimental session). As only 38 (6.6%) of 598 samples were longitudinal, we did not use this variable for inferential analyses.

*Manipulation length* ($\kappa$ = .64). This variable assessed whether the manipulation occurred in a single experimental session or in multiple sessions. Only 17 (3.0%) of the 598 samples had procedures occurring over multiple sessions, so we did not use this variable for inferential analyses.

*Topic characteristics*

*Evaluative vs. conceptual associations* ($\kappa$ = .85). Implicit associations vary in whether their content is more evaluatively (e.g.,, good/bad) or conceptually (e.g., mental/physical) focused. Because different neural substrates are associated with affective and semantic memory (Amodio & Devine, 2006; Amodio & Ratner, 2011), it is possible that the same procedure will produce different effects on conceptual and evaluative associations. We therefore coded whether the concepts involved in the target association were primarily evaluative (e.g., good/bad in a self/other-good/bad IAT) or conceptual (e.g., science/humanities in a male/female-science/humanities IAT). Some associations had both evaluative and conceptual content (e.g., a Lexical Decision Task where the primes are pictures of Black people and the targets are negative Black stereotypes). We handled these on a case-by-case basis.

*Self-associations* ($\kappa$ = .85). The self is one of the most fundamental constructs in psychology (James, 1890), and has long been an important construct in research on automatic processes (Greenwald & Banaji, 1995). Whether self-associations should be more or less easy to change than other associations is unclear. To assess the role of the self in implicit malleability, we coded whether or not the concepts involved in the target association were related to the self.

*Association domain* ($\kappa$ = .97). The topics studied by experiments in the meta-analysis were diverse, ranging from anti-Arab/Muslim prejudice to dieting and exercise. Coders judged whether the study's topic was related to intergroup relations, health psychology, personality, clinical psychology, political preferences, consumer preferences, or close relationships.

*Article characteristics*

*Publication status.* Larger significant effects are more likely to be published than smaller non-significant effects (Stern & Simes, 1997). We assessed whether this was the case in this literature by coding whether a study had been published in an academic journal or book at the time of analysis. Many of the unpublished studies were dissertations and/or studies in a researcher's "file-drawer," but some unpublished studies were studies that were in the process of being prepared for publication.

*Publication year*. The effect size of early published studies is often larger than effect sizes of later published studies on the same topic (Jennions & Møller, 2002), a result popularly known as the decline effect. There are multiple possible reasons for the decline effect, including publication bias, increasing sample heterogeneity, and loss of adherence to intervention quality over time. We coded the year a study was published to see if a decline effect exists in this literature. Unpublished studies were not included in any analyses involving publication year.

*Study characteristics*

*Geographic region of sample* ($\kappa$ = .92). Published effect sizes from the United States in the behavioral sciences tend to be larger than those published in other countries, perhaps due to publication pressures (Fanelli & Ioannidis, 2013). To investigate whether this was the case in this literature, we coded whether the studies were conducted in the United States, Europe, Israel, Canada, Australia and New Zealand, Asia, Africa and Latin America, or multiple countries. For analysis, we compared the effect sizes of studies published in the US and elsewhere.

*Number of experimental groups* ($\kappa$ = .67). This variable represented the numbers of groups in a study's design. Sometimes this with synonymous with the number of conditions in an experiment, but other times it was not (e.g., when a condition was excluded, when multiple conditions were merged together for analysis).

**Meta-analytic computations**

Meta-analysis involves the synthesis of one or more effect sizes and the variance components associated with those effect sizes. The breadth of this project demanded special procedures to do so.

*Standardized mean differences.* Given our interest in the degree to which quantitative measures differ between groups, the primary effect size computed for this project was the standardized mean difference. For each comparison between procedures on the implicit, explicit, and behavioral measures, we estimated Hedge's *g* (Hedges & Olkin, 1985), which is a measure of the standardized mean difference similar to Cohen's *d* that corrects for small-sample bias. We estimated Hedge's *g* using the means, standard deviations, and number of participants within each cell of a given sample's design. If the total sample size was available but the number of participants per group was not, we assumed equal sample sizes within each group. If the means and/or standard deviations were missing, we attempted to back-calculate the missing descriptive statistics or the standardized mean difference from other statistics reported in the article (see Lipsey & Wilson, 2001). If this was not possible, we requested the required information directly from the authors.

In multi-group designs (i.e., designs with more than two groups), we designated one group the "reference group" and computed multiple effect sizes relative to this reference group (Salanti, 2012; White et al., 2012). This yielded ($g$ - 1) effect sizes, where $g$ is the number of groups in a study. Where possible, this reference group was a neutral condition. In studies that lacked a neutral condition, we calculated effect sizes relative to a virtual neutral condition that had an effect size of 0 and a standard error of 1000 (Higgins & Whitehead, 1996; White et al., 2012). This computational device ensures that studies that lack a neutral condition will contribute information during model fitting (Higgins & Whitehead, 1996) without directly influencing meta-analytic estimates involving neutral conditions (White et al., 2012). The virtual neutral conditions therefore play a similar role as continuity corrections to avoid divide-by-zero errors when analyzing odds ratios: they allow estimation to proceed without inappropriately impacting results.

We handled experiments with pre-test post-test designs by using the mean differences from pre-test to post-test as the means within each condition and the pre-test standard deviations as our standard deviations within each condition (Morris & DeShon, 2002; Morris, 2008). If the pre-test standard deviations were unavailable but the standard deviations of the differences from pre-test to post-test were available, we used the standard deviations of the differences instead, then transformed this change score metric into one comparable to the pre-test standard deviation metric (Morris & DeShon, 2002). If we were unable to obtain either the pre-test or difference score data, we computed effect sizes with post-test data only. Some studies used dichotomous outcomes to measure behavior. For these outcomes, we calculated log-odds ratios that we then transformed into standardized mean differences (Cox & Snell, 1989; Sánchez-Meca et al., 2003).

*Variance components*. The variances of Hedge's $g$ in post-test only designs were estimated using formulas developed by Hedges and Olkin (1985). In experiments with pre-test post-test designs, we estimated the effect size variances using formulas that correct for the correlation between pre-test and post-test (Morris & DeShon, 2002; Morris, 2008). For studies missing the correlation between pre-test and post-test (27/84 implicit correlations; 11/35 explicit correlations, 3/14 behavioral correlations), we imputed the missing correlation with its meta-analytic estimate calculated from the rest of the sample (implicit $r = .35$, $k = 57$, 95% CI = [.29, .41]; explicit $r = .74$, $k = 24$, 95% CI = [.68, .79]; behavioral $r = .72$, $k = 11$, 95% CI = [.66, .78]). We estimated the variance components for effect sizes for dichotomous measures using a formula described by Cox and Snell (1989).

Effect sizes extracted from a single study are typically non-independent, either because they share a common reference group in multi-group studies or because the same participants complete multiple measures (i.e., when participants take a measure of implicit bias and a measure of explicit bias and/or behavior). Thus, in addition to the variances typically estimated in pairwise meta-analyses, we also estimated covariances between each pair of effect sizes derived from a given study in studies that yielded multiple effect sizes. For multi-group studies, estimating the covariance between effect sizes only requires the number of people per condition and the means and standard deviations of the outcome measure (Gleser & Olkin, 2009). For studies that use multiple measures (i.e., an explicit and/or behavioral measure in addition to an implicit measure), the calculation of these covariances requires the correlation between the two types of measures. In studies where this correlation was unavailable (26/256 implicit-explicit correlations; 12/92 implicit-behavioral correlations), we imputed the correlation using the meta-analytic estimate from the remaining studies (implicit-explicit $r = .14$, $k = 230$, 95% CI = [.12, .16]; implicit-behavioral $r = .10$, $k = 80$, 95% CI = [.06, .14]). We calculated the covariances between different measures using formulas derived by Wei and Higgins (2013).

*Indirect effects*. For both explicit bias and behavior, we estimated effects from both a mediation model and a reverse mediation model. The mediation model estimated the degree to which the effects of procedures on the target outcome is mediated by change in implicit bias, and the reverse mediation model estimated the degree to which the effects of procedures on implicit bias is mediated by change in the target outcome. We constructed a series of 3 by 3 correlation matrices representing the bivariate relationships between manipulations, implicit bias, and explicit bias/behavior. The correlations between manipulations and other variables were extracted for each study report by transforming the standardized mean differences on implicit bias and explicit bias/behavior into correlation coefficients. These correlations were combined with the correlation between implicit bias and explicit bias/behavior.[5] We only included two-condition studies when constructing these correlation matrices because of ambiguity in how to define the direct and indirect effects in multi-condition studies. We then used the delta method to extract the standardized indirect effects and their asymptotic variances from these correlation matrices (Cheung, 2009).[6]

## Results

### Network meta-analysis

We performed most of the analyses using a multivariate implementation of network meta-analysis (Lu & Ades, 2004; Caldwell et al., 2005; Salanti, 2012). Multivariate network meta-analysis treats each study in the meta-analysis as having multiple outcomes. Each of these outcomes is a potential comparison between the two of the 12 categories of procedures coded for the meta-analysis. Because studies that contain more than two categories of procedures yield more than one two-group comparison, multivariate network meta-analysis also explicitly models the interdependence between these multiple comparisons.

More formally, given $k$ studies comparing $g$ conditions, multivariate network meta-analysis represents each study as a set of comparisons between one of the conditions (the reference group $r$) and each other condition. Thus, study $i$ yields a vector of $(g - 1)$ effect sizes, labeled $y_i$, along with a $(g - 1)$ by $(g - 1)$ matrix of variances and covariances between the effect sizes within study $i$, labeled $S_i$. Given effect sizes $y_i$ and covariance matrices $S_i$, one can estimate coefficients α and the between-studies variance-covariance matrix Σ using the following multivariate model (White et al., 2012):

$$y_i \sim N(\alpha X_i, \Sigma + S_i)$$

where $X_i$ is a matrix of study covariates. If there are no study covariates and α and Σ are assumed to be the same across studies, α represents the meta-analytic effect size estimates of comparisons between the reference group and each other condition and Σ represents the between-studies variance-covariance matrix for those effect sizes.

---

[5] Although we imputed this correlation for the analysis of the consistency between effects on implicit bias and explicit bias/behavior, we did not impute this correlation for the analysis of the indirect effects.

[6] We also estimated the direct effects, their asymptotic variance, and the asymptotic covariance between the direct and indirect effect so as to not bias the indirect effect estimates by fixing them to 0. We only report the indirect effects here.

An advantage of this meta-analytic model is that it uses both direct information from the comparisons within each study and indirect information from the pattern of comparisons across studies (Higgins & Whitehead, 1996; Lu & Ades, 2004). For example, taking the difference between the effect of the comparisons between procedures A & B and procedures A & C allows for the indirect estimation of the comparison of procedures B & C. Direct and indirect information can only be combined if a network of comparisons meets the *consistency assumption*, which assumes that each procedure is similar regardless of which other procedures appear alongside it in a given study (Salanti, 2012). We tested the viability of this assumption by testing whether, within single treatment estimates, studies of different designs had different effect sizes (the design by treatment interaction approach; Higgins et al., 2012; White, 2011). They did not, $\chi^2(69, k = 573) = 84.28, p = .102$, indicating the consistency assumption was reasonable for our data.

We fit all multivariate network meta-analytic models using the metaSEM package in R (Cheung, 2015). To ensure model identifiability, we constrained the components of the between-studies variance-covariance matrix $\Sigma$ such that the variances were equal and the covariances were equal (Higgins & Whitehead, 1996; Lu & Ades, 2009).

**Descriptive information**

Descriptive information about the articles, studies, samples, and measures included in the meta-analysis is shown in Table 2. The data primarily came from published articles (80.5%), studies conducted in the United States (53.0%), and from studies of intergroup relations (63.4%). The participants in the meta-analysis reflect the demographics of students in Introductory Psychology classes: 81.7% of samples were composed entirely of university students, and samples were majority White (76.2%) and female (65.6%). The majority of the samples used evaluative measures (65.1%), usually with an IAT (64.7%), and usually in a single-session, post-test only design (16.1%). Only 38 (6.6%) of the samples used a longitudinal design to assess change over time, and only 18 (3.1%) used intense, multi-session procedures. Finally, 45.7% of the samples included an explicit measure, and 16.2% of the samples contained a behavioral measure.

Some study characteristics were weakly to moderately related to each other. The strongest relationships were that health/clinical studies were more likely to use a pre-test post-test design ($r = .41$) and include a behavioral measure ($r = .38$) than studies in other domains. For a complete correlation matrix of study characteristics, see https://osf.io/awz2p/.

The network of comparisons between the 12 categories of procedures is shown in Figure 1. The most common procedure most frequently used in a study was the neutral category. Indeed, most studies (74.9%) compared neutral procedures with one or more comparison procedures. When studies made other types of comparisons, they most often (86.8%) compared a procedure and its conceptual opposite (e.g., positive and negative affective states). Few studies that made non-neutral comparisons used procedures in conceptually different categories (13.2%) (e.g., weaken associations directly vs. threat).

*Table 2*. Characteristics of the final meta-analysis sample.

**Methodological characteristics**

| | | | |
|---|---|---|---|
| *.Proc. length* | Single session | 555 | 97.0% |
| | Multiple sessions | 17 | 3.0% |
| *.Longitudinal* | Longitudinal | 38 | 6.6% |
| | Non-longitudinal | 535 | 93.4% |
| *.Design* | Post-test only | 481 | 83.9% |
| | Pre-test post-test | 92 | 16.1% |
| *.Measure* | IAT | 371 | 64.7% |
| | Priming | 60 | 10.5% |
| | SC-IAT/ST-IAT | 27 | 4.7% |
| | Other | 115 | 20.1% |
| *.Explicit* | Explicit measure | 262 | 45.7% |
| | None | 311 | 54.3% |
| *.Behavior* | Behavioral measure | 93 | 16.2% |
| | None | 480 | 83.8% |

**Topic characteristics**

| | | | |
|---|---|---|---|
| *Domain* | Intergroup | 313 | 63.4% |
| | Personality | 64 | 13.0% |
| | Health/clinical | 69 | 14.0% |
| | Other | 48 | 9.7% |
| *Type* | Evaluative | 373 | 65.1% |
| | Conceptual | 200 | 34.9% |
| *Self-related* | Non-self | 470 | 82.0% |
| | Self | 103 | 18.0% |

**Sample characteristics**

| | | | |
|---|---|---|---|
| *Population* | University student | 468 | 81.7% |
| | Not university student | 105 | 18.3% |
| *Gender* | Female | 52,631 | 65.6% |
| | Male | 27,601 | 34.4% |
| | Not reported | 15,205 | |
| *Race* | White | 42,687 | 76.2% |
| | Non-White | 13,362 | 23.8% |
| | Not reported | 39,158 | |

**Study characteristics**

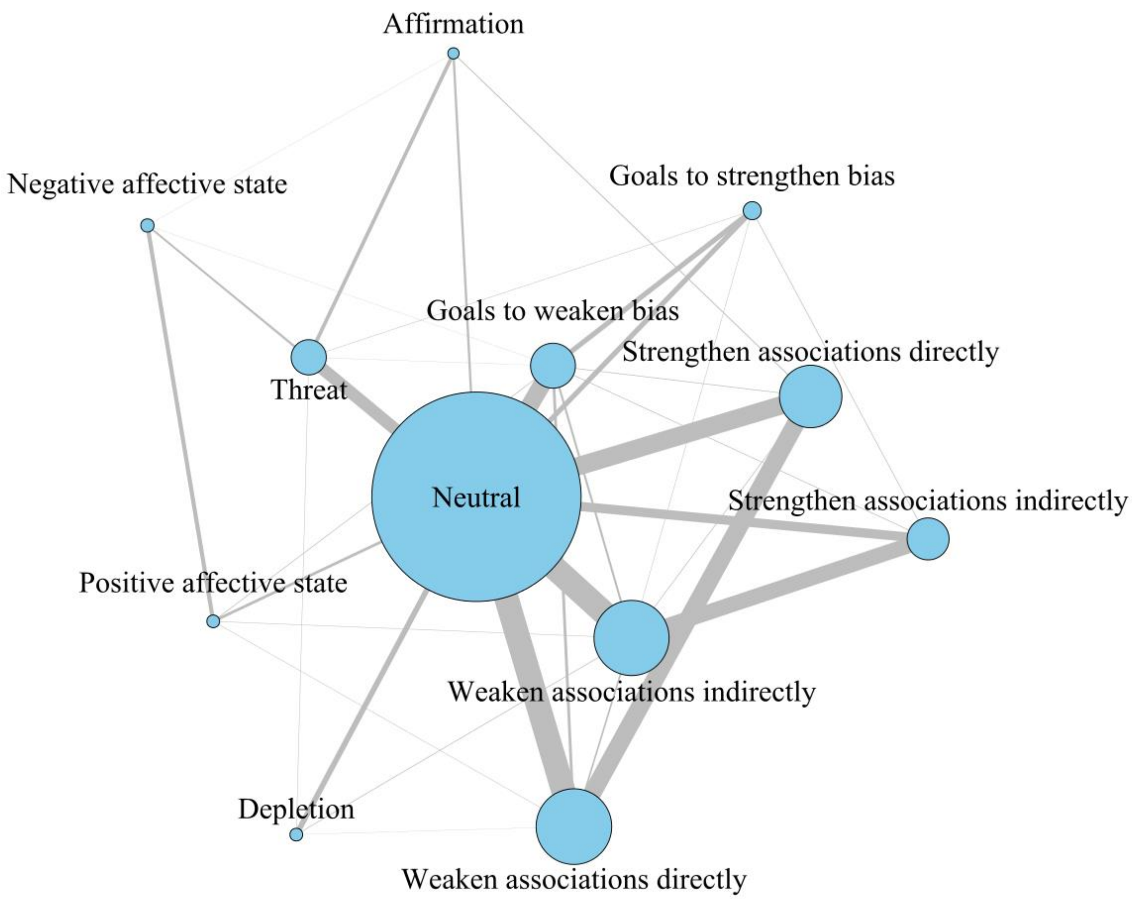
| | | | |
|---|---|---|---|
| *Location* | US | 262 | 53.0% |
| | Europe | 134 | 27.1% |
| | Canada | 41 | 8.3% |
| | Other | 32 | 6.5% |
| | Multiple | 25 | 5.1% |
| *Conditions* | Two | 237 | 48.0% |
| | Three | 90 | 18.2% |
| | Four | 102 | 20.6% |
| | Five+ | 65 | 13.2% |

**Article characteristics**

| | | | |
|---|---|---|---|
| *Status* | Published | 277 | 80.5% |
| | Unpublished | 67 | 19.5% |
| *Date* | 1995 - 2000 | 3 | 1.1% |
| | 2001 - 2005 | 31 | 11.3% |
| | 2006 - 2010 | 87 | 31.6% |
| | 2011+ | 154 | 56.0% |

*Note*. Methodological, topic, and sample characteristics are presented in # of samples. Gender/Race are presented in # of participants. Study characteristics are presented in # of studies. Publication status is presented in # of papers, and publication date is presented in # of published papers.

*Figure 1.* Network plot of procedures included in the meta-analysis. The radius of the category circles = the number of procedures in that category, line width = the number of samples in which a pair of conditions were directly compared.

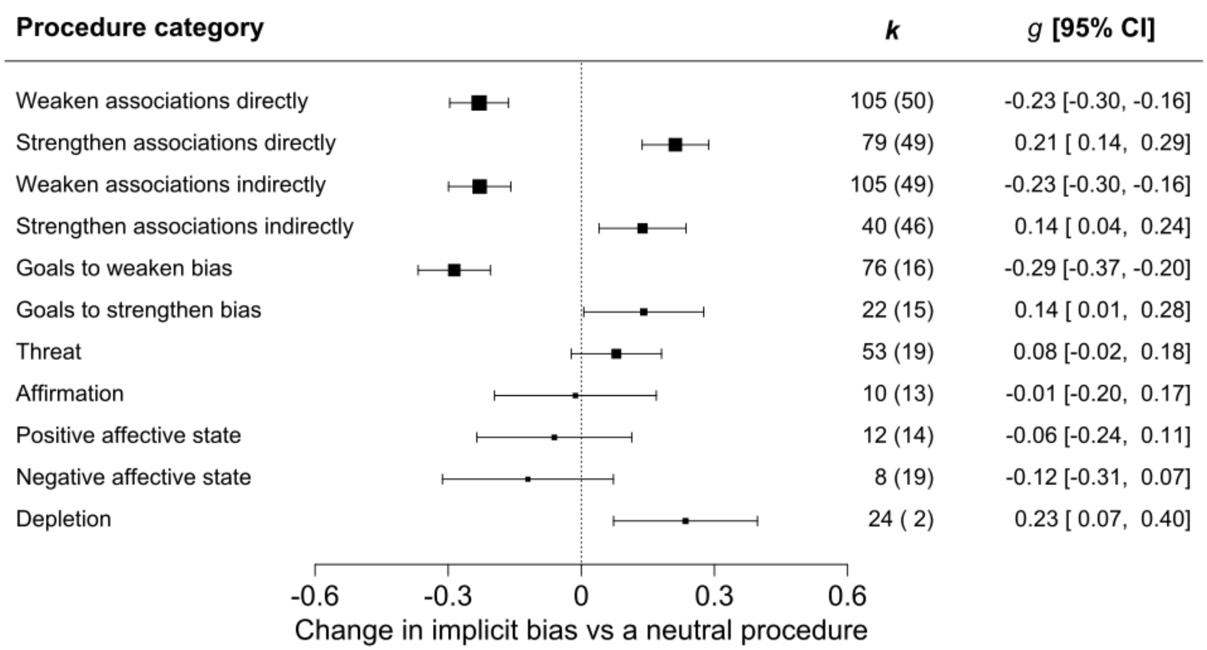



**What approaches to changing implicit bias are most influential?**

We compared the effectiveness of procedures to change implicit bias by fitting a multivariate network meta-analytic model with the neutral group as the reference category. As shown in Figure 2, seven categories changed implicit bias relative to a neutral condition: procedures that strengthen or weaken associations, either directly ($g_{\text{strengthen}}$ = .21, 95% CI = [.14, .29]; $g_{\text{weaken}}$ = -.23, 95% CI = [-.30, -.16]) or indirectly ($g_{\text{strengthen}}$ = .14, 95% CI = [.04, .24]; $g_{\text{weaken}}$ = -.23, 95% CI = [-.30, -.16]), that induce goals ($g_{\text{strengthen}}$ = .14, 95% CI = [.01, .28]; $g_{\text{weaken}}$ = -.29, 95% CI = [-.37, -.21]), and that deplete mental resources ($g$ = .23, 95% CI = [.07, .40]). In all cases, the effects were small by conventional standards ($|d|$ < .35; Hyde, 2005) and smaller than the typical effects reported in social psychology papers (median $d$ = .37; Richard, Bond, & Stokes-Zoota, 2003). Compared to a neutral procedure, procedures that that produce threat ($g$ = .08, 95% CI = [-.02, .18]), affirmation ($g$ = -.01, 95% CI = [-.20, .17]), positive affective states ($g$ = -.06, 95% CI = [-.23, .11]), and negative affective states ($g$ = -.12, 95% CI = [-.31, .07]) produced effects that were small and not distinguishable from zero.

We estimated the variation in effect sizes due to substantive differences between studies using the multivariate *R*-based statistic developed by Jackson, White, and Riley (2011). This statistic revealed high between-study variation ($I^2$ = .798), a finding mirrored by the large estimated effect size standard deviation ($\tau$ = .303). This reflects the diversity of disciplines, theoretical approaches, and methodological approaches in this area.[7]

*Figure 2*. Forest plot of the comparisons between each procedure and a neutral procedure. $k$ gives the number of studies that directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure. $g$ gives the estimated standardized mean difference and its 95% CI. Higher effect sizes reflect greater increases in implicit bias relative to a neutral procedure.




**Are the sample, methodology, or topic of a study associated with the magnitude of implicit bias change?**

We tested whether effect sizes varied according to the sample, design, or topic of a study. We did this by using Wald $\chi^2$ tests that compared moderator models to models without any moderators. Treating $p < .05$ as a strict criterion, there was evidence of variation based on whether the sample was a student sample, $\chi^2(9, k = 573) = 27.38$, $p = .001$, the racial composition

---

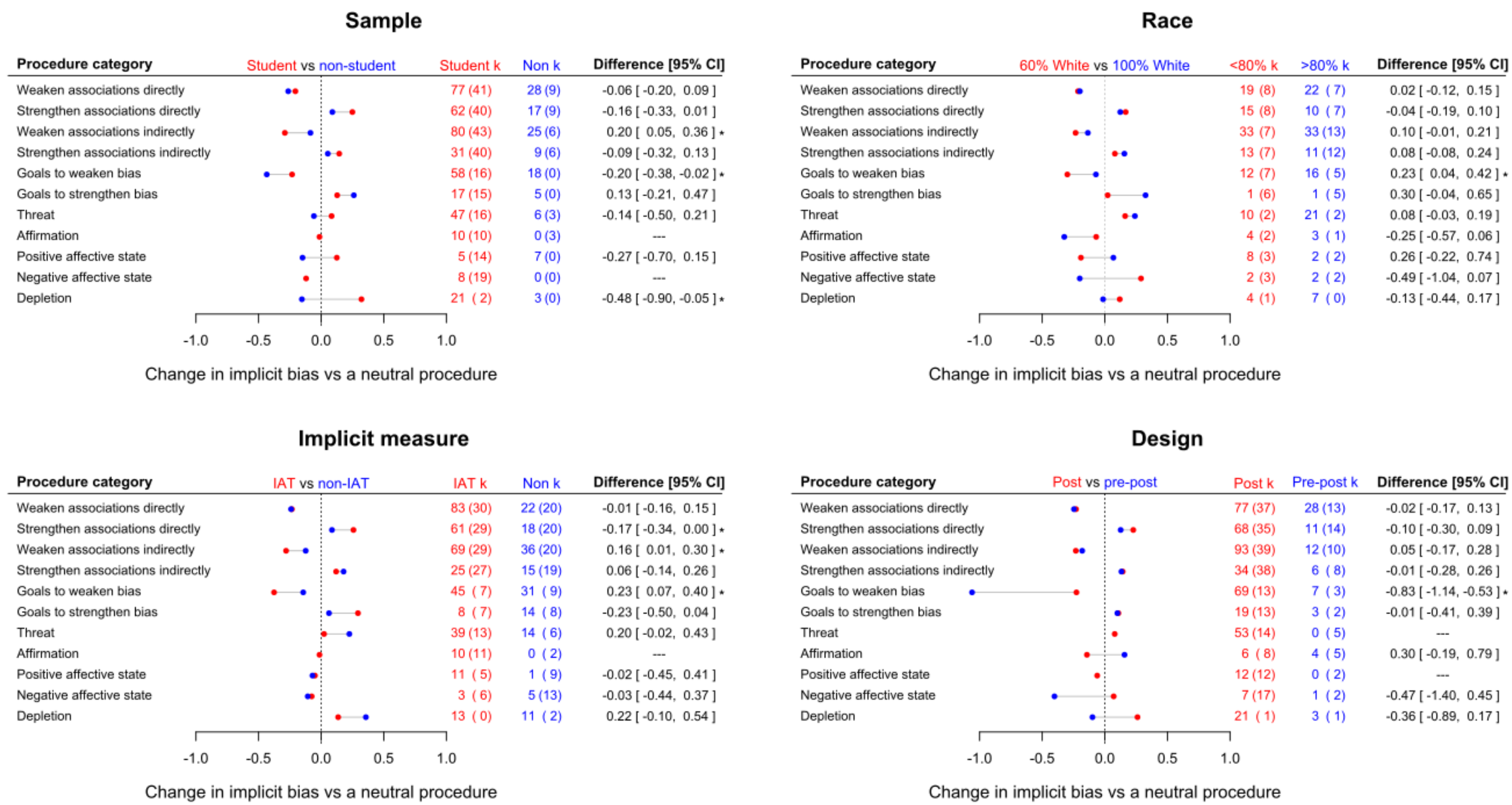
[7] Because the between-studies variances of the meta-analytic comparisons were constrained to be equal (Higgins & Whitehead, 1996; Lu & Ades, 2009), we could not estimate how the between-studies variance differed for each comparison between procedure categories.

of the sample, $\chi^2(11, k = 248) = 20.59$, $p = .038$, the measure used to assess implicit bias, $\chi^2(10, k = 573) = 27.62$, $p = .002$, and whether the design included a pre-test assessment of implicit bias, $\chi^2(9, k = 573) = 36.44$, $p < .001$. There was little evidence of variation by the number of conditions compared within the study, $\chi^2(11, k = 573) = 13.43$, $p = .266$, the gender composition of the sample, $\chi^2(11, k = 483) = 14.94$, $p = .185$, whether the target association was evaluative or conceptual, $\chi^2(11, k = 573) = 19.58$, $p = .050$, whether the target association was an intergroup association, $\chi^2(11, k = 573) = 18.01$, $p = .081$, whether the target association was related to health or clinical issues, $\chi^2(9, k = 573) = 8.16$, $p = .518$, or whether the target association was related to the self, $\chi^2(8, k = 573) = 15.46$, $p = .051$.

The specific differences for the significant moderators are shown in Figure 3. Most of the moderator differences are driven by the effects of procedures that induce goals to weaken associations. These procedures produced stronger effect sizes in non-student samples ($g_{\text{non-student}} = -.43$, $g_{\text{student}} = -.23$), samples with proportionately fewer White people ($g_{\text{60\% White}} = -.30$, $g_{\text{100\% White}} = -.07$), studies that used an IAT ($g_{\text{IAT}} = -.38$, $g_{\text{non-IAT}} = -.14$), and studies with a pre-test assessment of implicit bias ($g_{\text{pre-test}} = -1.06$, $g_{\text{post-test only}} = -.23$). These results suggest that there are strong sample and methodological differences between studies that show a strong effect of goals to weaken associations and studies that do not.

Student and non-student samples also tended to produce different effect sizes. In addition to the difference between student and non-student samples for studies using weaken goals procedures, student and non-student samples produced different effect sizes in studies that weakened associations indirectly ($g_{\text{non-student}} = -.09$, $g_{\text{student}} = -.29$) and that depleted cognitive resources ($g_{\text{non-student}} = -.15$, $g_{\text{student}} = .32$). Finally, in addition to these sample differences, compared to studies that used a different implicit measure, studies using an IAT produced stronger effects than non-IAT studies when they strengthened associations directly ($g_{\text{IAT}} = .26$, $g_{\text{non-IAT}} = .09$) and weakened associations indirectly ($g_{\text{IAT}} = -.28$, $g_{\text{non-IAT}} = -.12$).

*Figure 3*. Moderation analyses. *k* gives the number of studies that directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure for the displayed levels of the moderator. “Difference” represents the difference between the two moderator levels and its 95% CI. Higher effect sizes reflect greater increases in implicit bias compared to a neutral procedure. Where there was not enough data in one of the moderator levels for estimation, the overall model estimate is shown instead.



**How do changes in implicit bias correspond with changes in explicit bias?**

To test whether the effects on implicit bias are consistent with effects on explicit bias, we fit a network meta-analytic model that allows the simultaneous analysis of two correlated outcomes (Efthimiou et al., 2015; Achana et al., 2014). This model revealed that effects on implicit bias differed from effects on explicit bias, $\chi^2(11, k = 572) = 34.63, p < .001$.[8] As shown in Figure 4, although effects on explicit bias were non-zero, $\chi^2(11, k = 572) = 68.87, p < .001$, they tended to be smaller than effects on implicit bias. Three of the eleven procedures had effects on implicit bias that were significantly larger than their effects on explicit bias: weaken associations directly, weaken associations indirectly, and weaken goals. The rest of the procedures except for threat, affirmation, and negative affect had non-significantly larger effects on implicit bias. Explicit effect sizes tended to be less variable than implicit effect sizes, both in terms of the percentage of between-studies heterogeneity ($I^2_{\text{implicit}} = .786$, $I^2_{\text{explicit}} = .751$) and the effect size standard deviations ($\tau_{\text{implicit}} = .282$, $\tau_{\text{explicit}} = .231$).

To test whether implicit bias change mediated the effects of procedures on explicit bias and whether explicit bias change mediated the effects of procedures on implicit bias, we synthesized the indirect effects extracted from the correlation matrices from each study using two-stage meta-analytic structural equation modeling (Cheung & Chan, 2005; Cheung & Cheung, 2016). We modeled the differences between the indirect effects resulting from different procedure comparisons using a contrast-based approach, which represents direct comparisons using dummy codes and indirect comparisons using treatment contrasts (Salanti, Higgins, Ades, & Ioannidis, 2008). Because we only conducted these analyses with two-condition studies for which we knew the implicit effect size, explicit effect size, and the correlation between implicit and explicit bias, the results are based on fewer studies ($k = 189$) than the full set of studies that contain an explicit measure ($k = 262$). All values from this analysis can be interpreted as the product of a correlation and a semi-partial correlation.

As shown in Figure 5, the indirect effects are all quite small. A Wald $\chi^2$ test suggested that we could not reject the null hypothesis that the indirect effects of procedures on explicit bias through implicit bias change were zero, $\chi^2(10, k = 189) = 17.94, p = .083$. The same was true of the reverse mediation estimating the effects of procedures on implicit bias through explicit bias change, $\chi^2(10, k = 189) = 7.74, p = .736$. None of the estimates for the indirect effects of the specific procedures was different from zero. There is little evidence in our data that is consistent with a causal relationship between automatically and deliberately retrieved associations. There was so little variation between studies in the magnitude of the indirect effects that the variation had to be fixed to zero for the models to converge.

---

[8] One study was removed from this analysis because its within-studies variance-covariance matrix of effects on implicit and explicit bias was degenerate.

*Figure 4*. Forest plot of the consistency between effects on implicit and explicit bias. *k* gives the number of studies with implicit and explicit measures that directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure. "I - E" gives the difference between the implicit and explicit effect sizes, "$\chi^2$" gives the 1 *df* Wald $\chi^2$ test of the difference, and "*p*" gives its *p*-value.

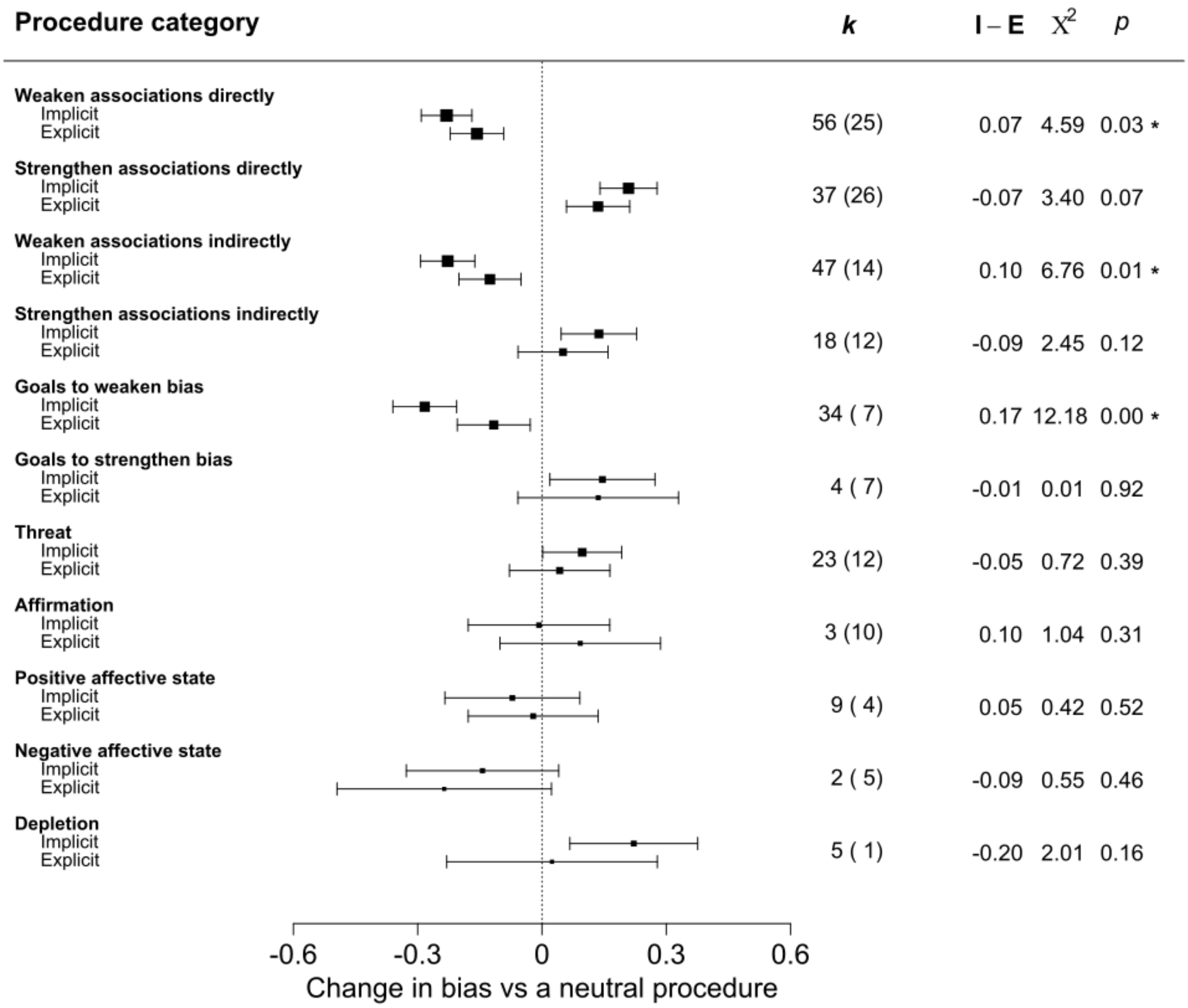
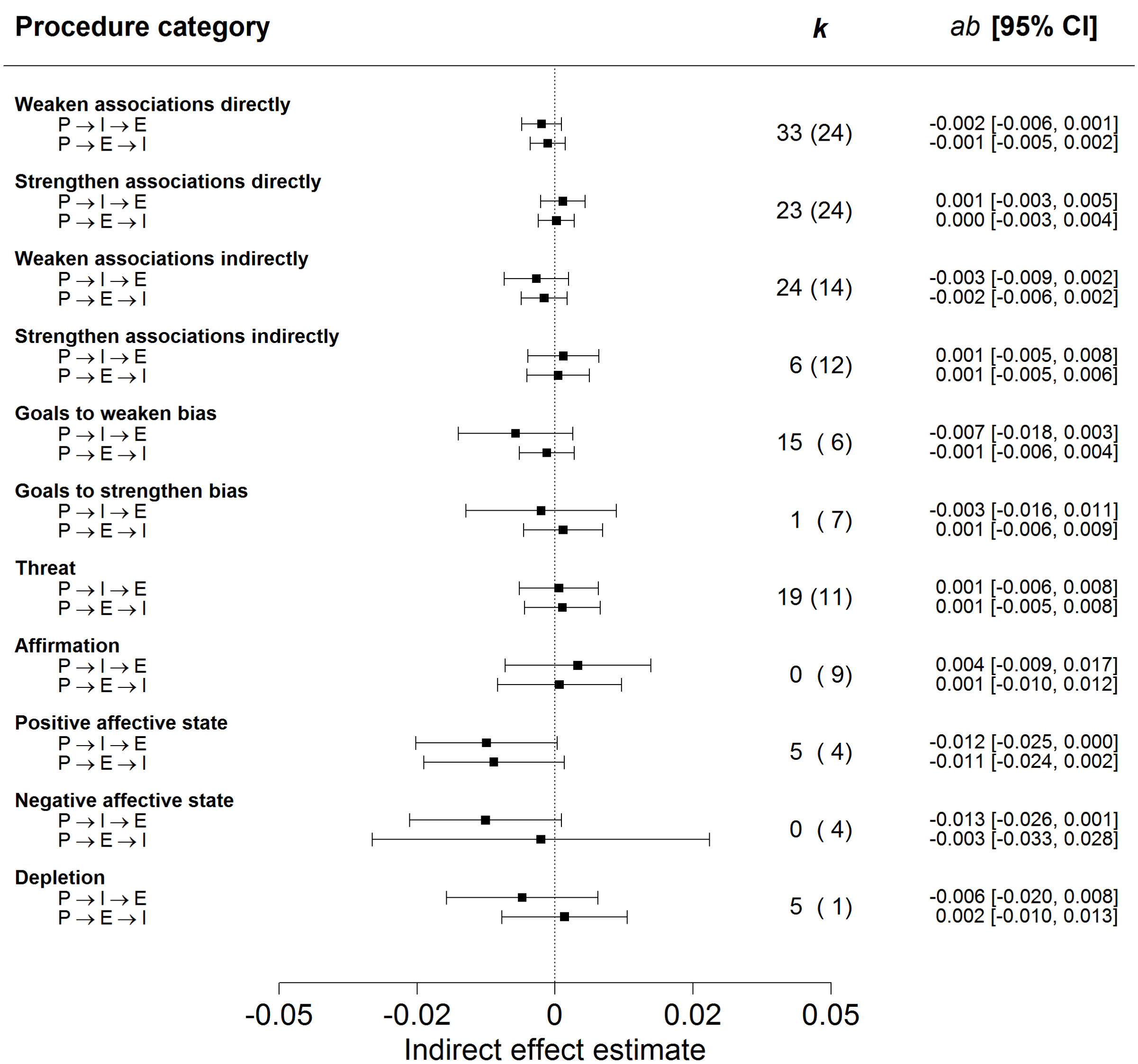


*Figure 5*. Indirect effects (in the conventional mediation framework, the effect *ab*) of procedures on explicit bias through changes in implicit bias ($P \rightarrow I \rightarrow E$) and implicit bias through changes in explicit bias ($P \rightarrow E \rightarrow I$). $k$ gives the number of studies that directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure.



**How do changes in implicit bias correspond with changes in behavior?**

We performed a similar set of analyses on behavior as we did on explicit bias.[9] As with explicit bias, effects on implicit bias differed from effects on behavioral outcomes, $\chi^2(7, k = 489) = 23.11$, $p = .002$. As shown in Figure 6, procedures (with the exception of threat) had much smaller effects on behavior than on implicit bias. Unlike with explicit bias, we did not reject the null hypothesis that all behavioral effects were zero, $\chi^2(7, k = 489) = 13.38$, $p = .063$. Behavioral effects were somewhat less variable than implicit effects, both measured in terms of the percentage of between-studies heterogeneity ($I^2_{\text{implicit}} = .777$, $I^2_{\text{behavior}} = .716$) and the effect size standard deviations ($\tau_{\text{implicit}} = .299$, $\tau_{\text{behavior}} = .274$).

As shown in Figure 7, we estimated both whether implicit bias change mediated the effects of procedures on behaviors and whether behavior change mediated the effects of procedures on implicit bias. As with our mediation analyses involving explicit bias, this analysis is based on a set of samples ($k = 62$) that is smaller than the set of samples that contain a behavioral measure ($k = 93$) because it only includes two-condition studies that had complete data. There was no evidence that procedures had non-zero indirect effects, either on behavior through implicit bias change, $\chi^2(7, k = 62) = 11.20$, $p = .130$, or on implicit bias through behavior change, $\chi^2(7, k = 62) = 4.69$, $p = .698$. Only one of the indirect effects had a 95% confidence interval that excluded zero, the reverse mediation effect for procedures that induce goals to strengthen bias, and the indirect effect was in the opposite direction of the effect of the procedure on implicit bias ($g = .14$, 95% CI = [.01, .28]). There is little evidence in our data that is consistent with a causal relationship between automatically retrieved associations and behavior. As with the indirect effects on explicit bias, there was so little variation between studies in the size of the indirect effects that the variation had to be fixed to zero for the models to converge.

[9] Studies with affirmation, positive or negative affect, or depletion procedures were excluded from this analysis because there were no studies with behavioral measures that used these procedures. An additional study was removed from this analysis because its within-studies variance-covariance matrix of effects on implicit and behavioral bias was degenerate.

*Figure 6*. Forest plot of the consistency between effects on implicit bias and behavior. *k* gives the number of studies with implicit and behavioral measures that also directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure. "I - B" gives the difference between the implicit and behavioral effect sizes, "$\chi^2$" gives the 1 *df* Wald $\chi^2$ test of the difference, and "*p*" gives its *p*-value.

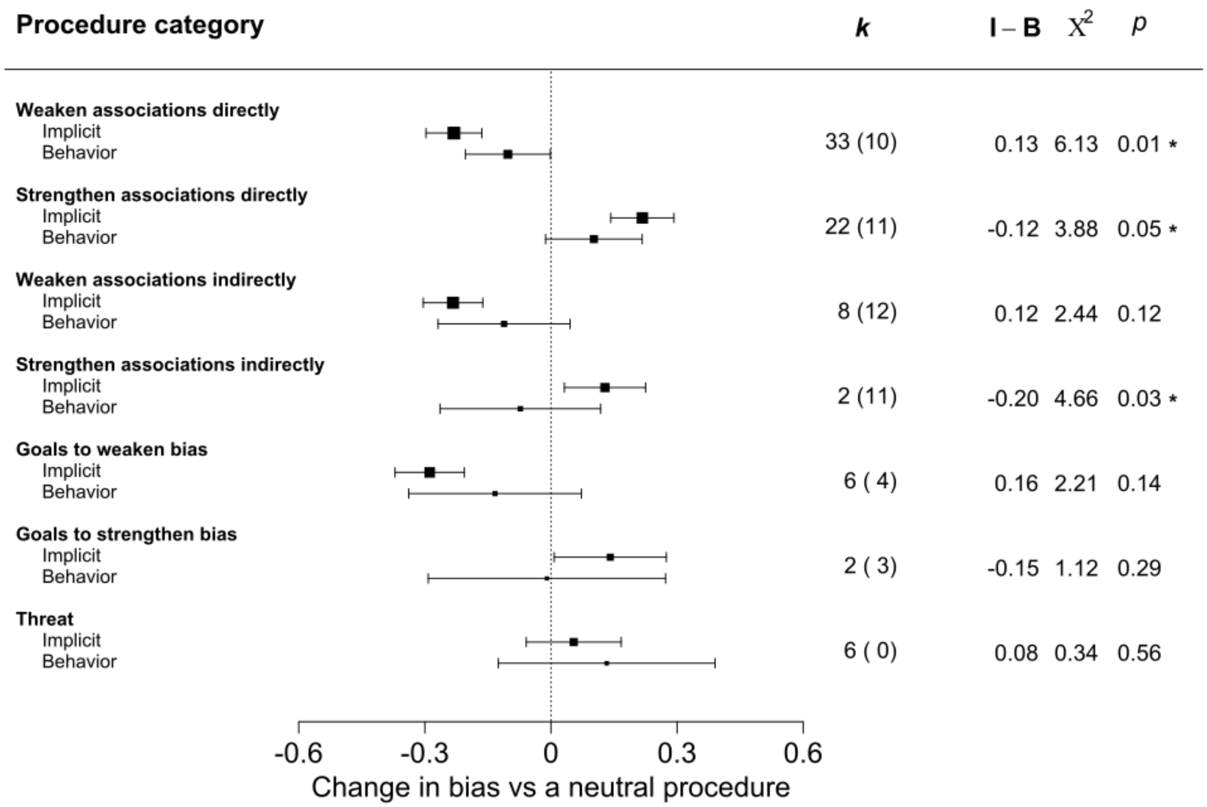
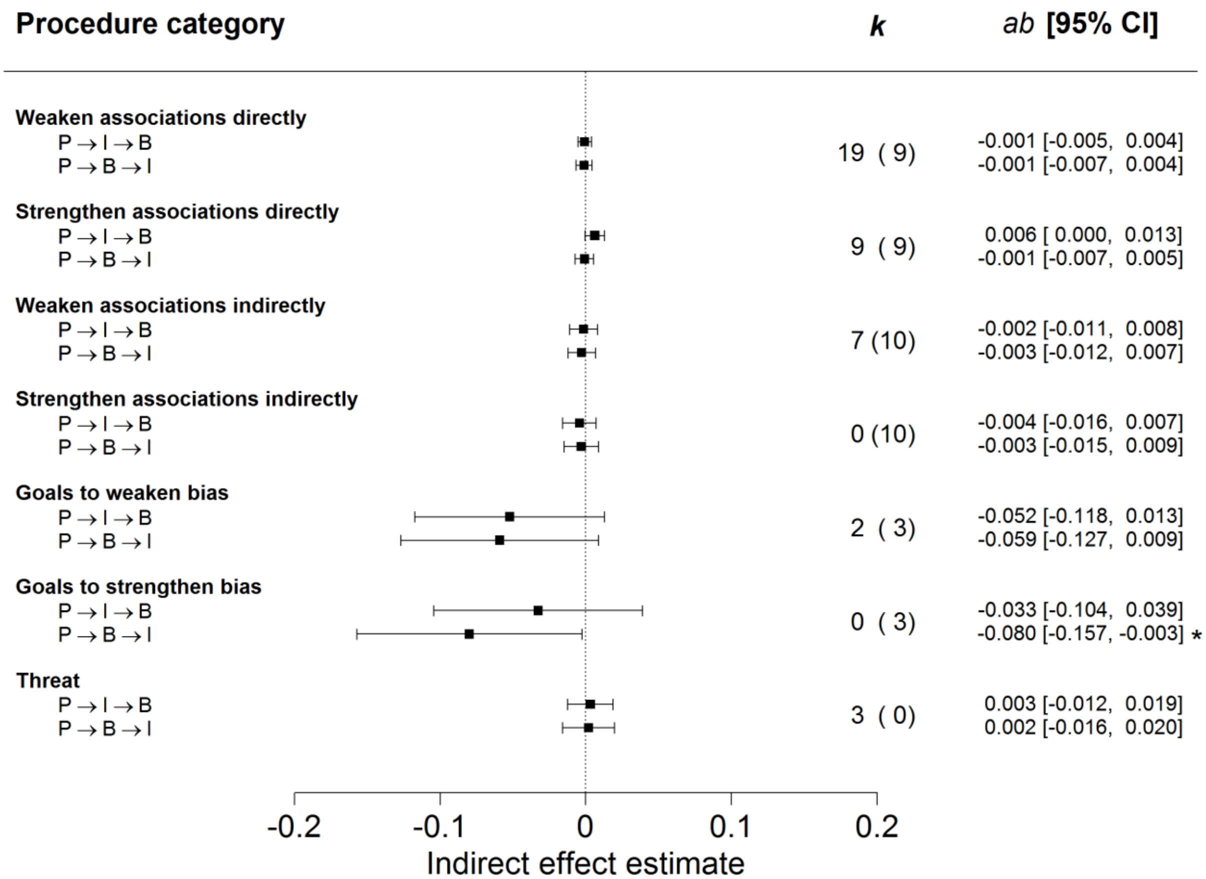


*Figure 7.* Indirect effects (in the conventional mediation framework, the effect *ab*) of procedures on behavior through changes in implicit bias ($P \to I \to B$) and implicit bias through changes in behavior ($P \to B \to I$). *k* gives the number of studies that directly (or indirectly, listed in parentheses) compare the listed procedure and a neutral procedure.



**Is there evidence that the size of reported effects is biased?**

We tested for biases in effect sizes by assessing funnel plot asymmetry[10] and by assessing whether effect sizes varied by publication status, year, or geographic location. Funnel plots show study effect sizes plotted against their standard errors (Egger et al., 1997). Funnel plots of an unbiased literature have a fan shape, with studies centering around a single effect size, regardless of precision, but with a greater scatter around the effect size in low-precision studies. Bias causes asymmetry in funnel plots by preventing a subset of low-precision studies (e.g., those with non-significant results) from entering the meta-analysis.

Comparison-adjusted funnel plots are funnel plots adapted to network meta-analysis (Chaimani et al., 2013). Although they cannot accommodate multiple effects from the same study, they can accommodate studies that examine different sets of comparisons between procedures. They account for these different comparisons by subtracting the relevant meta-analytic comparison estimate (e.g., threat vs. neutral, weaken goals vs. neutral) from each study estimate prior to plotting. As in a normal funnel plot, one can then examine the comparison-adjusted plots for asymmetry, which suggests that some process differentially affected high and low precision studies (e.g., publication bias).

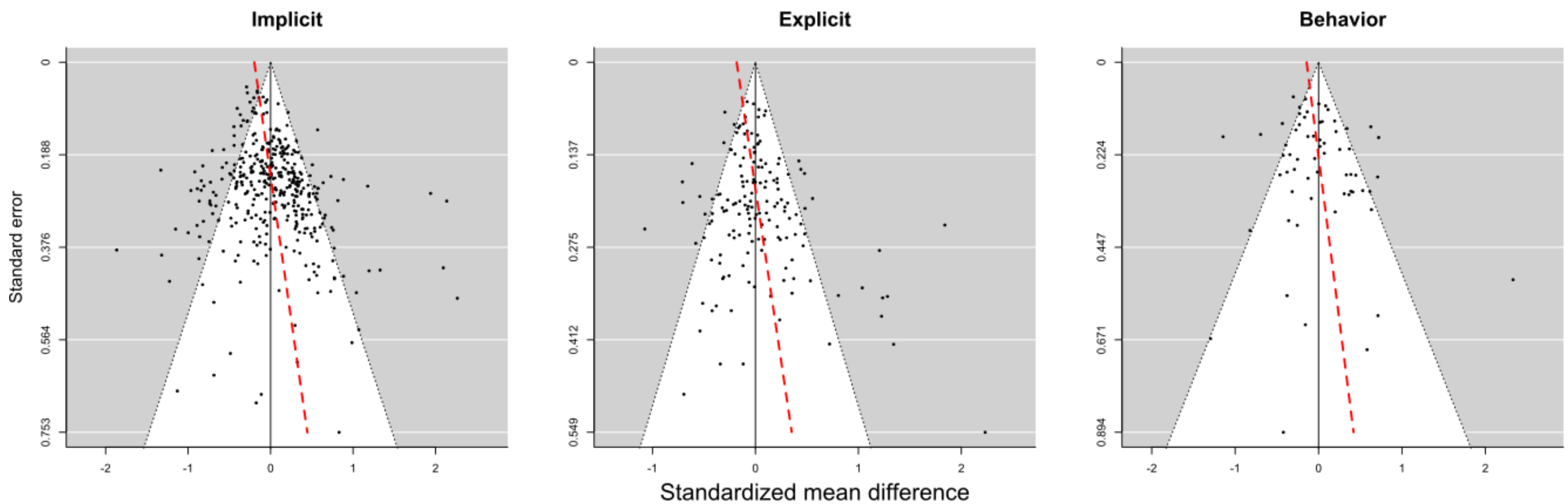
To select a set of two-group studies (published and unpublished) in which most researchers would make similar predictions, we made the following three generic predictions. First, the weaken associations directly, weaken associations indirectly, and weaken goals procedures will result in decreased implicit, explicit, and behavioral bias relative to a neutral procedure. Second, the strengthen associations directly, strengthen associations indirectly, strengthen goals, and deplete resources procedures will result in increased bias relative to a neutral procedure. Third, procedures in the first group will result in less bias than procedures in the second.

The funnel plots of the comparison-adjusted effect sizes for these studies on implicit, explicit, and behavioral measures are shown in Figure 8. The figure reveals asymmetry in all plots in that high-precision effect sizes tended to be smaller than their corresponding overall meta-analytic estimates. This observation was supported by the results of mixed-effect regression analyses (Sterne & Egger, 2005) testing the relationship between implicit standard errors and effect sizes, $z = 3.62$, $p < .001$ and explicit standard errors and effect sizes, $z = 2.99$, $p = .003$. There was no significant relationship between the behavioral standard errors and effect sizes, $z = 1.40$, $p = .163$, though this last relationship was also estimated with much less precision than the implicit and explicit relationships. If the funnel plot asymmetry is caused by processes that systematically prevent small, non-significant effect sizes from entering the meta-analysis (e.g., publication bias, *p*-hacking), this suggests that implicit and explicit effects in this meta-analysis are inflated relative to their population values.

---

[10] Several other methods of detecting bias are available, such as *p*-curve analysis (Simonsohn, Nelson, & Simmons, 2014) and contour-enhanced funnel plots (Peters et al., 2008). We did not use these methods because many of them depend on assumptions of homogeneity and have not been adapted to examining bias in a network of interventions where heterogeneity is expected *a priori* (for a review, see Efthimiou et al., 2016).

*Figure 8*. Comparison-adjusted funnel plots of effect sizes vs standard errors for implicit, explicit, and behavioral measures. Positive numbers are more extreme relative to the meta-analytic comparison a study contributes to and negative numbers less extreme. The red dashed line represents the fit from a mixed-effects regression; a line that departs from the vertical suggests the presence of small-study bias.

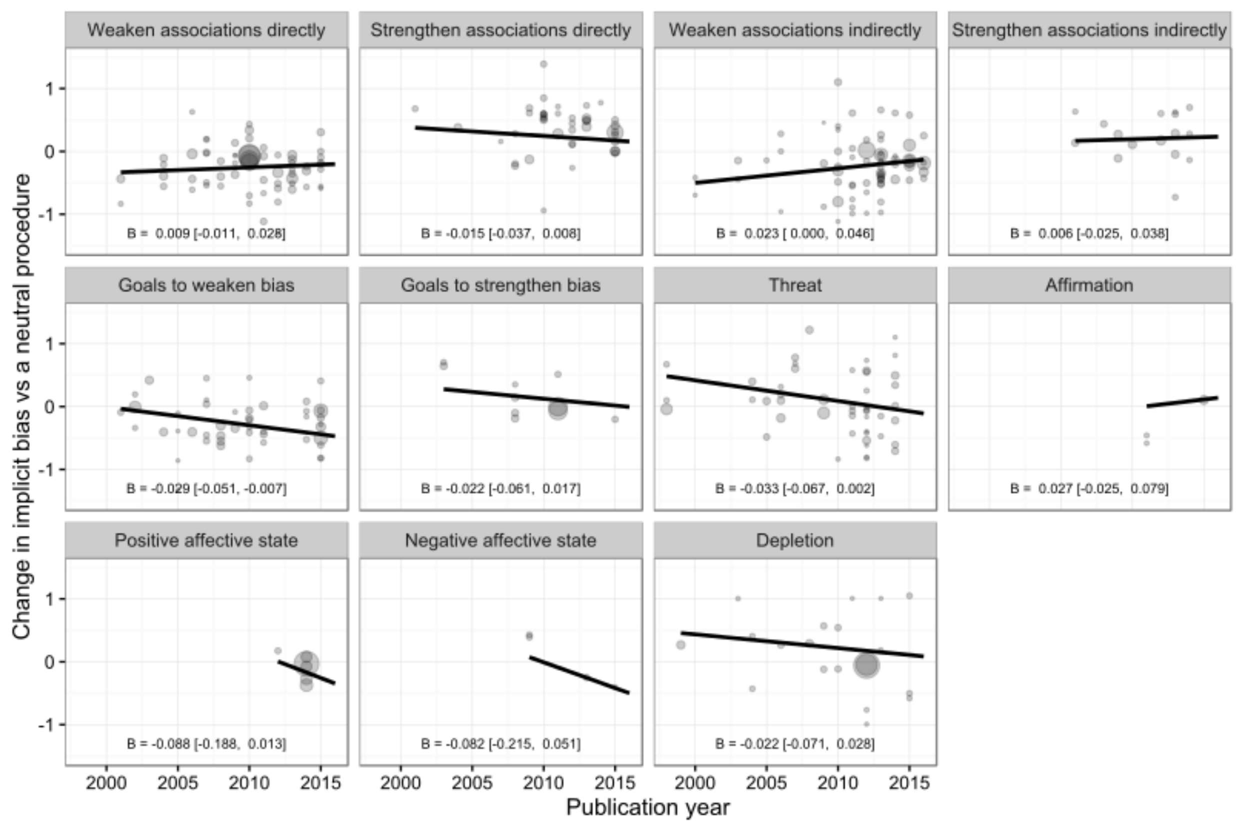


Funnel plots do not distinguish between the many processes that could lead to asymmetry. In contrast, testing whether specific moderators such as publication year, are related to effect sizes is more specific. We therefore supplemented our analysis of funnel plot asymmetry with analyses using publication year, publication status, and geographic region. These analyses allowed us to test for evidence of a decline effect (Jennions & Møller, 2002), publication bias (Stern & Simes, 1997), and a United States bias (Fanelli & Ioannidis, 2013), respectively.

Implicit effect sizes varied by publication year, $\chi^2(11, k = 464) = 25.64$, $p = .007$.[11] As shown in Figure 9, there was a general tendency for more recent studies to yield smaller effect sizes. There were two exceptions: strengthen associations indirectly, for which effect sizes remained constant across all publication years, $B = .006$, 95% CI = [-.025, .038], and goals to weaken bias, for which there was a growth effect rather than a decline effect – more recent studies have larger (more negative) effect sizes, $B = -.029$, 95% CI = [-.062, -.007]. This last relationship may be driven by research on the susceptibility of implicit measures to show no bias or a reverse bias through strategic responding (e.g., implementation intentions to reduce bias on a shooter bias task, Mendoza, Gollwitzer, & Amodio, 2010, instructions to Germans to fake a pro-Turkish IAT score, Fiedler & Bluemke, 2005). Early studies suggested that implicit measures were resistant to strategic responding (Banse, Seise, & Zerbes, 2001; Egloff & Schmukle, 2002; Kim, 2003), whereas more recent studies have suggested that strategic responding is possible, particularly with sufficiently specific instructions (Fiedler & Bluemke, 2005; Lai et al., 2014; 2016; Stewart & Payne, 2008). Effect sizes did not depend on publication status, $\chi^2(11, k = 573) = 18.01$, $p = .081$, or geographic location, $\chi^2(11, k = 573) = 5.63$, $p = .897$.

[11] We did not do any moderator analyses with the explicit or behavioral data because there was not sufficient variability in the moderators across the intervention categories.

*Figure 9.* Relationship between publication year and effect sizes on implicit bias. Larger points represent effect sizes that are estimated with greater precision. Only direct comparisons between each listed procedure and a neutral procedure are shown as points.




## General Discussion

Our meta-analysis is the first large-scale quantitative synthesis of research on change in implicit bias. We found that implicit bias can be changed across many areas of study, populations, implicit measures, and research designs. The type of approach used to change implicit bias mattered greatly. Some procedures were effective at changing implicit bias, whereas others were not. Procedures to change implicit bias produced similar but smaller changes in explicit bias, and there was no strong evidence that they produced any change in behavior at all. Further, changes in implicit bias did not mediate changes in explicit bias or behavior, nor did we find evidence that changes in explicit bias or behavior mediated change in implicit bias.

### Relative effectiveness of procedures to change implicit bias

We developed a taxonomy for understanding how procedures to change implicit bias differed. Using this taxonomy, we found that procedures that directly or indirectly targeted associations, depleted mental resources, or induced goals all changed implicit bias relative to neutral procedures. In contrast, procedures that induced threat, affirmation, or affective states had small and/or inconsistent effects. These results support the theoretical portrayal of automatically retrieved associations as sensitive to pairings of information in the social environment (Gawronski & Bodenhausen, 2006). These results also support the importance of goal-directed motivation and cognitive resources in changing the expression of automatically retrieved associations (Fazio & Olson, 2014; Gawronski & Payne, 2010; Devine, 1989).

However, even the procedures that produced robust effects on implicit bias had "small" effect sizes by conventional standards (Hyde, 2005) and as compared to typical effect sizes in social psychology (Richard, Bond, & Stokes-Zoota, 2003). Our funnel plot analyses suggest that the true population effects of these procedures may be even smaller than our meta-analytic estimates due to publication bias, *p*-hacking, and/or other related processes.

**Generalizability of implicit bias change**

We also uncovered evidence of large variation in the size of the effects produced by procedures to change implicit bias. Some of the sources of this variation reveal complexities in evaluating the impact of the procedures on implicit bias. First, researchers' choices of samples have constrained the generalizability of the available evidence (Henrich, Heine, & Norenzayan, 2010). Most studies have been conducted with samples whose demographic characteristics (students, mostly White, mostly female) strongly resemble those of Introductory Psychology classrooms in the United States. Although the gender composition of the sample was not associated with the size of effects, both the racial composition of the samples and whether the samples were drawn from university student populations were. Student samples in particular produced different effect sizes than non-student samples for three of the nine procedure comparisons that we examined (strengthen associations directly vs. neutral, weaken associations indirectly vs. neutral, goals to weaken bias vs. neutral).

Because studies with university student samples often address different research questions than studies with non-university student samples and because university students are psychologically different from the general population (Henrich et al., 2010; Sears, 1986), the precise cause of these different effect sizes is unclear. Regardless, these results suggest that it would be prudent to directly test whether the effects of manipulations are generalizable to other populations. Combating societal problems such as discrimination and addiction requires exploration of how the problems operate outside of the college campus, and answering questions of human nature depends on sampling from a population that represents humankind.

Another limit to generalizability is a lack of research interest in change beyond the confines of a single experimental session. Only 17 (3.0%) samples used procedures that took longer than one session to complete. Only 38 (6.6%) samples in the meta-analysis collected longitudinal outcomes and therefore had the opportunity to examine whether the procedures they investigated produce long-term changes. Short-term changes in implicit bias do not necessarily generalize to longer-term changes (Devine, Forscher, Austin, & Cox, 2012; Forscher et al., 2017; Forscher & Devine, 2014; Lai et al., 2016; Lai, Hoffman, & Nosek, 2013; Miller, Dannals & Zlatev, 2017). As such, the present meta-analysis speaks more to the processes that change implicit bias in the short-term rather than to processes that change implicit bias in the long-term. These points are particularly salient given theorizing that implicit bias is created and sustained by repeated pairings of information in the social environment. That means that without active efforts to sustain short-term shifts created in the lab, these shifts are likely to be wiped away upon re-exposure to the social environment (Forscher et al., 2017; cf. De Houwer, 2009; Mann &

Ferguson, 2017). In fact, one recent series of studies found that nine interventions that reduced implicit race bias immediately showed little to no lasting impact days later (Lai et al., 2016). What processes determine whether a shift in implicit bias will be temporary or long-lasting? When will a shift in implicit bias translate into a permanent change in orientation? Theory and practice-oriented researchers alike must contend with these open questions.

Effect sizes also differed according to a study's methodological features. Studies using an IAT produced effects that were often larger than studies that did not, and studies with a pre-test post-test design that induced a goal to weaken bias produced larger effects than studies that only included a post test measurement. The large IAT effects could be driven by the IAT's reliability, which is typically higher than the reliability of most other implicit measures (Bar-Anan & Nosek, 2014; Bosson et al., 2000). We speculate that the larger effects for pre-test post-test designs could be driven by effects in studies examining strategic responding on implicit tasks, as such research has revealed that strategic responding is particularly easy with past task experience (Fiedler & Bluemke, 2005).

The effects of interventions did not vary much based on their topic. Studies that targeted evaluative associations did not differ from studies that targeted conceptual associations, and effect sizes did not differ as a function of domain (e.g., intergroup relations, clinical/health).

**Implicit bias and explicit bias**

Most studies of the relationship between the implicit and explicit bias are observational studies that administer implicit and explicit measures within the same study. These relationships can be very low or very high, and are highest – when using the IAT at least – when the target concepts are not socially sensitive, when people's thoughts about the concepts are well-elaborated, when the concepts are diametrically opposed (e.g., liberals vs. conservatives), and when people perceive that their opinions about the concepts are distinct from the opinions of others (Nosek, 2005). Although it was not the primary purpose of our meta-analysis, we found that the correlation between implicit and explicit bias in our sample of experimental studies was quite low ($r_{\text{I-E}} = .14$). This is a marked difference from the median ($r_{\text{I-E}} = .38$) of large-sample studies ($N > 100{,}000$) investigating highly heterogeneous topics in highly heterogeneous samples. In fact, compared to 95 examined topics, the estimate from this meta-analysis was smaller than all but one (Nosek & Hansen, 2008).

There are good reasons expect a different correlation in experimental studies than in observational studies. The correlation between implicit and explicit bias will vary according to a manipulation's causal impact on implicit and explicit bias. However, other features besides a manipulation's causal impact also differ between observational studies of the correlation between implicit and explicit bias and the experimental studies examined in this meta-analysis. The experimental studies tended to use relatively homogeneous White student samples, single-session manipulations, and tended to focus on a limited range of topics. For example, the most common topic in this meta-analysis was intergroup relations (63.4% of studies), an area known for low implicit-explicit correlations in observational studies (Hofmann et al., 2005; Nosek, 2005, 2007). This homogeneous sampling may have constrained the magnitude of the correlation between implicit and explicit bias beyond what might be expected due to the causal impact of experimental manipulations.

Our focus on randomized studies also gave us an opportunity to go beyond correlational evidence by examining whether procedures that attempt to change implicit bias also produce change in explicit bias, and whether change in explicit bias mediated change in implicit bias. We found that many of the procedures that change implicit bias also produce change in explicit bias, though the magnitude of change in explicit bias was weaker and less variable. Simultaneously, there was no evidence that changes in implicit and explicit bias were mediated by each other. One possibility suggested by these data is that there is no relationship between changes in implicit and explicit bias. This possibility would reduce support for theoretical perspectives that posit interdependence between automatic and deliberate processes that are presumed to underlie implicit and explicit bias (e.g., Gawronski & Bodenhausen, 2006; c.f. Smith & DeCoster, 2000). However, even if this is true, we cannot eliminate the possibility that the relationship is stronger in other samples or topics.

It is not possible from these data to determine whether increasing diversity in samples, designs, and topics would yield substantively different mediation results. The most productive next step is to evaluate these possibilities directly. There are some hints that such investigations would yield stronger mediation evidence. For example, Smith, Ratliff, and Nosek (2012) had large samples of participants (*N*'s = 732; 621) form attitudes toward novel policy proposals that were randomly attributed to Democrats or Republicans. Implicit and explicit attitudes toward the plans were strongly correlated (*r*'s = .48, .51/.59) and implicit attitudes fully mediated the effect of the experimental intervention on explicit attitudes, but not the reverse, both immediately and 5 days after the intervention.

This example was not included in this meta-analysis because we only examined studies of pre-existing biases. As a consequence, this and all other studies of the formation of new associations were excluded. This creates an interesting mystery to be solved. The association formation literature provides substantial experimental evidence for the interdependence of automatically and deliberately retrieved associations (e.g., Gawronski & Bodenhausen, 2006, 2011; Gawronski & LeBel, 2008; Gawronski, Rydell, Vervliet, & De Houwer, 2010; Moran, Bar-Anan, & Nosek, 2015; Ranganath & Nosek, 2008). In contrast, this meta-analysis on pre-existing associations provides little evidence of interdependence. Whatever the explanation, resolving the apparent discrepancy between research on new and pre-existing associations provides an exciting opportunity to advance theory about implicit bias.

**Implicit bias and behavior**

Previous investigations of implicit-behavior relations have also relied on observational studies. Meta-analytic estimates of this relationship vary substantially (Greenwald et al., 2009 $r_{\text{I-B}}$ = .27; Cameron et al., 2012 $r_{\text{I-B}}$ = .28; Oswald et al., 2013 $r_{\text{I-B}}$ = .14; Carlsson & Agerström, 2016 $r_{\text{I-B}}$ = .15). The correlations between implicit bias and behavior tend to be smallest for topics in which automatic and deliberate processes are least likely to facilitate each other, such as race relations (Greenwald et al., 2009). The overall correlation between implicit bias and behavior in our meta-analysis was small and closer to the estimates in the meta-analyses on these topics ($r_{\text{I-B}}$ = .10).

On the surface, this research is about prediction, but of course, the interest is also about causation. Indeed, many researchers use evidence of correlations between implicit bias and behavior to argue for the causal importance of automatically retrieved associations (e.g., Dovidio, Kawakami, & Gaertner, 2002; Green et al., 2007; Kang & Banaji, 2006; Devine et al., 2012; Banaji, Bhaskar, & Brownstein, 2015). For example, Devine, Forscher, Austin, and Cox (2012) argue on the basis of correlational studies that "accumulating evidence reveals that implicit biases are linked to discriminatory outcomes ranging from the seemingly mundane, such as poorer quality interactions (McConnell & Leibold, 2001), to the undeniably consequential, such as constrained employment opportunities (Bertrand & Mullainathan, 2004) and a decreased likelihood of receiving life-saving emergency medical treatments (Green et al., 2007). [...] [Implicit bias] leads people to be unwittingly complicit in the perpetuation of discrimination."

Of course, correlations between variables can be produced by many relationships besides ones that are causal. To get closer to questions of causality, we looked at whether changes in implicit bias correspond with and mediate changes in behavior in our sample of randomized experiments. We found no little evidence that procedures that change implicit bias also produce change in behavior. We also found no evidence that changes in implicit bias mediate changes in behavior, nor that changes in behavior mediate changes in implicit bias.

The lack of evidence for mediation is difficult to reconcile with the correlational evidence. As with explicit bias, the fact that the studies in our meta-analysis disproportionately used primarily White student samples, single-session manipulations, and a narrow range of topics provides an important limit to the generalizability of our conclusions. However, even if the relationship between changes in implicit bias and changes in behavior is truly smaller in the domains sampled by this meta-analysis (i.e., intergroup relations, addiction, clinical psychology) than in other domains (e.g., political attitudes), our results suggest a constraint on the conditions under which changing implicit bias will predict or cause corresponding changes in unwanted behavior.

**Potential explanations for implicit bias's relationships with explicit bias and behavior**

Even if we accept that our explanations of our explicit bias and behavior findings do not generalize to all samples and topics, we are left with specifying what those explanations are. We offer four possibilities.

First, our inclusion criteria for explicit and behavioral measures may have led to the inclusion of measures that should not be theoretically expected to change after a change in automatically retrieved associations. We included any measure that we judged to be correspondent with the study's implicit measure, regardless of whether they are expected to change after the manipulation. If the theoretical conditions under which a change in automatically retrieved associations will influence deliberately retrieved associations and behavior are narrower than mere correspondence, our meta-analysis would not be sensitive to those conditions. For example, if the implicit measure was a Black/White good/bad IAT, we included any explicit or behavioral measure that connected race and valence. Correspondent explicit measures ranged from a simple feeling thermometer that assesses perceived warmth toward Whites vs. Blacks (Rudman, Dohn, & Fairchild, 2007) to the Symbolic Racism Scale that assesses the degree to which participants blame Black people for their current social standing (Inzlicht, Gutsell, & Legault, 2012). Correspondent behavioral measures ranged from how close a person sits to a Black confederate (Mann & Kawakami, 2012) to decisions about donating to children in South African vs. Colombian slums (Schwab & Greitemeyer, 2015). Oftentimes, experimental manipulations are designed to create dissociations that target mental processes relevant to implicit biases but not explicit biases or behavior. For example, exposure to counterstereotypical exemplars is often found (and expected) to change implicit racial attitudes without having corresponding effects on explicit racial attitudes (Dasgupta & Greenwald, 2001; Cullen, Barnes-Holmes, & Barnes-Holmes, 2009; Lai et al., 2014; McGrane & White, 2007). Perhaps stricter inclusion criteria that were more sensitive to specific theoretical conditions where implicit change was expected to relate to explicit or behavioral change would have revealed stronger evidence for mediation.

Second, perhaps confounds introduced after the manipulations obscured the evidence for mediation. Statistical mediation analysis relies on the untestable assumption of a lack of confounding of the post-manipulation mediator-outcome relationship (Bullock, Green, & Ha, 2010). Most, but not all, sources of confounding will overstate the evidence for mediation (Bullock et al., 2010). However, confounding that reduces evidence for mediation could explain the null results. That may happen, for example if a second mediator that opposes the causal influence of implicit bias was also changed by many of the procedures examined in the meta-analysis. We cannot rule out this explanation, but we also cannot identify what these confounds would be.

Third, measurement issues may obscure the evidence for mediation within our studies. No measure provides a pure estimate of a latent construct (Borsboom, 2006), and implicit measures are no exception (Conrey et al., 2005; Calanchini & Sherman, 2013; Payne, 2001). Performance on implicit measures reflects the contribution of associative processes, measurement error, and non-associative processes such as task-switching ability, recoding, inhibition of impulses, and guessing (Calanchini et al., 2013; 2014; Klauer & Mierke, 2005). High levels of measurement error, as is characteristic of implicit measures (Buhrmester, Blanton, & Swann, 2011; Olson & Fazio, 2002; Bosson et al., 2000) could obscure evidence that changes in implicit bias mediate changes in other processes.[12]

It is also possible that many of the procedures we examined produced change in measured implicit bias through non-associative processes. At least some of the procedures did do so. For example, a subset of studies that used goals to strengthen or weaken bias gave participants instructions to strategically respond or fake an implicit measure (e.g., Banse, Seise, & Zerbes, 2001; Fiedler & Bluemke, 2005). If many of our procedures produced change through non-associative processes, our analyses would bear on the effectiveness of these non-associative processes for producing change in explicit bias and behavior rather than the effectiveness of automatically retrieved associations. Without tools that isolate the contributions of associative and non-associative processes, we cannot definitively rule this possibility out.

Fourth, perhaps automatically retrieved associations really are causally inert. Accepting this conclusion would force reevaluation of some of the central assumptions that drive implicit

[12] Measurement error in implicit measures would not explain the lack of an overall effect of procedures on behavioral outcomes, although measurement error in behavioral measures might. A recent meta-analysis (Carlsson & Agerström, 2016) found that behavioral measures in research on the IAT and discrimination lacked validity and reliability. Many of the behavioral measures in this meta-analysis appeared to suffer from similar measurement issues. For example, many behavioral outcomes were based on as a single behavior (rather than an aggregate of multiple behaviors) and were not based on standardized procedures where the validity and reliability is well-known.

bias research. For example, instead of acting as a “cognitive monster” that inevitably leads to bias-consistent thought and behavior (e.g., Bargh, 1999; Tajfel, 1982), automatically retrieved associations could reflect the residual “scar” of concepts that are frequently paired together within the social environment. From this view, implicit biases are a side effect of living in a particular social environment. Similarly, a scar interpretation would suggest that changes in implicit bias represent epiphenomenal changes rather than changes in the causal processes that drive deliberately retrieved associations or behavior.

This is not to say that the measurement of implicit bias would be unproductive even under this interpretation. Demographic variables such as life expectancy are often used to predict other consequential outcomes within a population, despite lacking causal force themselves. By the same token, levels of implicit bias could be used to predict the prevalence of certain judgments or behaviors within a population. However, under this interpretation, although the presence of an implicit bias would speak to the structure of the social environment, efforts to change behavior by changing implicit bias would be misguided. It would be more effective to rid the social environment of the features that cause biases on both behavioral and cognitive outcomes (Beaman, Duflo, Pande, & Topalova, 2012) or equip people with strategies to resist the environment’s biasing influence (Devine et al., 2012; Cohen & Sherman, 2014) rather than trying to alter the biases themselves.

A new account of automatically retrieved associations as non-causal requires theoretical integration of findings that do not converge with the results of this meta-analysis. Although the scar interpretation of implicit bias explains correlations between implicit bias, explicit bias, and behavior as resulting from the shared cause of the social environment, this interpretation is nonspecific and does not explain why certain correlations between implicit bias and other variables are stronger than others. For example, well-elaborated concepts have stronger levels of convergence between implicit and explicit bias (Nosek, 2005), and people who have higher levels of working memory have lower levels of convergence between implicit bias and behavior (Hofmann, Gschwendner, Wiers, Friese, & Schmitt, 2008; Perugini, 2005; for a review, see Perugini, Richetin, & Zogmaister, 2010). A non-causality account would also have to integrate studies on novel associations that, as noted above, provide stronger evidence for mediation, at least in the case of explicit bias (e.g., Gawronski & Bodenhausen, 2006, 2011; Gawronski & LeBel, 2008; Gawronski et al., 2010; Moran et al., 2015; Ranganath & Nosek, 2008). Presently, it is unclear how to theoretically integrate the scar interpretation with this evidence.

**Conclusion**

This meta-analysis found that implicit bias can be changed, and identified the approaches that are most successful in doing so. However, we also found little evidence that changes in implicit bias translated into changes in explicit bias and behavior, and we observed limitations in the evidence base for implicit malleability and change.

These results produce a challenge for practitioners who seek to address problems that are presumed to be caused by automatically retrieved associations, as there was little evidence showing that change in implicit biases will result in changes for explicit biases or behavior. This is particularly true for the domains of greatest interest to many practitioners – intergroup bias, health psychology, and clinical psychology. Our results suggest that current interventions that attempt to change implicit bias will not consistently change behavior in these domains.

These results also produce a challenge for researchers who seek to understand the nature of human cognition because they raise new questions about the causal role of implicit bias. The results of the current meta-analysis do not lend themselves to a single interpretation. To better understand what the results mean, future research should innovate with more reliable and valid implicit, explicit, and behavioral measures, intensive manipulations, longitudinal measurement of outcomes, heterogeneous samples, and diverse topics of study.

These innovations may yet reveal stronger evidence for the causal importance of automatically retrieved associations. It would not be the first time that the conclusions of a review were overturned by later advances. Following Wicker's (1969) review showing a weak correlation between explicit attitudes and behavior, better measurement and theory revived the relevance of attitudes for understanding thought and action. As they did in response to Wicker, we hope that researchers take our findings as a challenge to improve theory and method and advance our understanding of human cognition.

References

* Contains data used in the meta-analysis.

* Abedi, F. H., Noorbala, F., & Saeedi, Z. (2010). Stereotypical beliefs about sex role-typed occupations and the mediating effect of exposure to counterstereotypic examples in Iranian students. *Procedia Social and Behavioral Sciences*, *5*, 1425-1428.

Achana, F. A., Cooper, N. J., Bujkiewicz, S., Hubbard, S. J., Kendrick, D., Jones, D. R., & Sutton, A. J. (2014). Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. *BMC Medical Research Methodology, 14*, 92.

* Agustin, A. A., & Francisco, V. G. J. (2008). Analysing the effects of mortality salience on prejudice and decision-taking. In F. M. Olson (Ed.), *New Developments in the Psychology of Motivation* (pp. 53-65). Nova Science Publishers.

* Albertson, B. L. (2011). Religious appeals and implicit attitudes. *Political Psychology*, *32*, 109-130.

Allen, T. J., Sherman, J. W., Conrey, F. R., & Stroessner, S. J. (2009). Stereotype strength and attentional bias: Preference for confirming versus disconfirming information depends on processing capacity. *Journal of Experimental Social Psychology, 45*, 1081–1087.

* Amir, N., Kuckertz, J. M. & Najmi, S. (2012). The effect of modifying automatic action tendencies on overt avoidance behaviors. *Emotion*, *13*, 478-484.

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652–661.

* Amodio, D. M. & Hamilton, H. K. (2012). Intergroup anxiety effects on implicit racial evaluation and stereotyping. *Emotion*, *12*, 1273-1280.

Amodio, D. M., & Ratner, K. G. (2011). A memory systems model of implicit social cognition. *Current Directions in Psychological Science, 20*, 143–148.

* Anderson, A. J., Ahmad, A. S., King, E. B., Lindsey, A. P., Feyre, R. P., Ragone, S., & Kim, S. (2015). The effectiveness of three strategies to reduce the influence of bias in evaluations of female leaders. *Journal of Applied Social Psychology*, *45*, 522-539.

* Aramakis, V. B., Khamba, B. K., MacLeod, C. M., Poulos, C. X., Zack, M. (2012). Alcohol selectively impairs negative self-relevant associations in young drinkers. *Journal of Psychopharmacology*, *26*, 223-231.

* Arendt, F. & Marquart, F. (2015). Corrupt politicians? Media priming effects on overtly expressed stereotypes toward politicians. *Communications-European Journal of Communication Research*, *40*, 185-197.

* Arendt, F. (2012). A newspaper‡EEfs effect on the strength of automatic associations in memory. *Journal of Media Psychology*, *24*, 1-8.

* Arendt, F. (2013). Dose-dependent media priming effects of stereotypic newspaper articles on implicit and explicit stereotypes. *Journal of Communication*, *63*, 830-851.

* Arendt, F., Marquart, F., & Matthes, J. (2015). Effects of right-wing populist political advertising on implicit and explicit stereotypes. *Journal of Media Psychology: Theories, Methods, and Applications*, *27*, 178-189.

* Asendorpf, J. B., Banse, R., & Mucke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, *83*, 380-393.

* Asgari, S. (2004). The malleability of stereotypic beliefs: Combating implicit stereotypes about ingroups and the self. Unpublished dissertation.

* Asgari, S., Dasgupta, N., Stout, J. G. (2012). When do counterstereotypic ingroup members inspire versus deflate? The effect of successful professional women on young women's leadership self-concept. *Personality and Social Psychology Bulletin*, *38*, 370-383.

* Ashby, C. R., & Stritzke, W. G. K. (2013). Is sensitivity to reward associated with the malleability of implicit inclinations toward high-fat food? *Emotion*, *13*, 711-723.

Axt, J. R., Ebersole, C. E., & Nosek, B. A. (2016). An unintentional, robust, and replicable pro-Black bias in social judgment. Social Cognition, 34, 1-39.

* Axt, J. R. & Nosek, B. A. (2012). Effect of depletion on implicit race attitudes. Unpublished dataset.

* Axt, J. R. & Nosek, B. A. (2012). Influence of multiculturalism on the Black-White Brief IAT. Unpublished dataset.

* Axt, J. R. & Nosek, B. A. (2013). Effect of simulated positive contact on implicit race attitudes. Unpublished dataset.

* Baccus, J. R., Baldwin, M. W., Packer, D. J. (2004). Increasing Implicit Self-Esteem Through Classical Conditioning. *Psychological Science*, *15*, 498-502.

Banaji, M. R., Bhaskar, R., & Brownstein, M. (2015). When bias is implicit, how might we think about repairing harm? *Current Opinion in Psychology*, *6*, 183-188.

* Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, *48*, 145-160.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect : From ecological rationality to dual processes, *Behavioural and Brain Sciences, 30*, 241–297.

* Bar-Anan, Y. (2010). Strategic modification of the evaluative priming effect does not reduce its sensitivity to uncontrolled evaluations. *Journal of Experimental Social Psychology*, *46*, 1101-1104.

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods*, *46*, 668–688.

* Bar-Anan, Y., & Zitelny, H. (2014). Homophobic attitudes. Unpublished dataset.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process Theories in Social Psychology* (pp. 361–382). New York, NY: Guilford Press.

* Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The implicit relational assessment procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record*, *60*, 57-66.

* Bassett, J. F. (2005). Does threatening valued components of cultural worldview alter explicit and implicit attitudes about death? *Individual Differences Research*, *3*, 260-268.

* Bauer, M. A., Wilkie, J. E. B., Kim, J. K. & Bodenhausen, G. V. (2012). Cuing consumerism: Situational materialism undermines personal and social well-being. *Psychological Science*, *23*, 517-523.

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India. *Science, 335*, 582–586.

Beevers, C. (2005). Cognitive vulnerability to depression: A dual process model. *Clinical Psychology Review, 25*, 975–1002.
* Bernstein, M. J., Claypool, H. M., Young, S. G., Tuscherer, T., Sacco, D. F. & Brown, C. M. (2013). Never let them see you cry: Self-presentation as a moderator of the relationship between exclusion and self-esteem. *Personality and Social Psychology Bulletin*, *39*, 1293-1305.
* Berry, B. A. (2015). Experimenter characteristics, social desirability, and the Implicit Association Test. *Psi Chi Journal of Psychological Research*, *20*, 247-257.
* Berry, T. R., McLeod, N. C., Pankratow, M. & Walker, J. (2013). Effects of Biggest Loser exercise depictions on exercise-related attitudes. *American Journal of Health Behavior*, *37*, 96-103.
Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*, 991–1013.
* Birch, C. D., Stewart, S. H., Wiers, R. W., Klein, R. M., MacLean, A. D., & Berish, M. J. (2008). The mood-induced activation of implicit alcohol cognition in enhancement and coping motivated drinkers. *Addictive Behaviors*, *33*, 565-581.
Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242–261.
* Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142-1164.
* Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology*, *81*, 828-841.
* Blincoe, S., & Harris, M. J. (2009). Prejudice reduction in white students: Comparing three conceptual approaches. *Journal of Diversity in Higher Education*, *2*, 232-242.
Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology*, *42*, 192-212.
* Bluemke, M., Degner, J., Lotz, J., Ritzenhoefer, L., & Shelliem, L. (2013). Intended and unintended reverberation of traditional and pro-age commercials as a function of viewer age. *Applied Cognitive Psychology*, *27*, 474-482.
* Bluemke, M., Friedrich, M., & Zumbach, J. (2010). The influence of violent and nonviolent computer games on implicit measures of aggressiveness. *Aggressive Behavior*, *36*, 1-13.
Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science, 1,* 319–322.
Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, *71*, 425-440.
Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631–643.
* Boudjemadi, V. & Gana, K. (2012). Effect of mortality salience on implicit ageism: Implication of age stereotypes and sex. *European Review of Applied Psychology*, *62*, 9-17.
* Bowley, C., Faircy, C., Hegarty, B., Johnstone, S. J., Smith, J. L., Kelly, P. J., Rushby, J. A. (2013). The effects of inhibitory control training on alcohol consumption, implicit

alcohol- related cognitions and brain electrical activity. *International Journal of Psychophysiology*, *89*, 342-348.
* Boysen, G. A., Vogel, D. L., & Madon, S. (2006). A public versus private administration of the implicit association test. *European Journal of Social Psychology*, *36*, 845-856.
* Bradley, K. I., Kennison, S. M., Burke, A. L. & Chaney, J. (2012). The effect of mortality salience on implicit bias. *Death Studies*, *36*, 819-831.
* Breines, J. G. (2012). Adjusting the thermometer of race relations: Physical warmth reduces bias. Unpublished dataset.
* Brosschot, J. F., Geurts, S. A. E., Kruizinga, I., Radstaak, M., Verkuil, B., Quirin, M., & Kompier, M. A. J. (2014). Does unconscious stress play a role in prolonged cardiovascular stress recovery? *Stress and Health: Journal of the International Society for the Investigation of Stress*, *30*, 179-187.
* Brown, R. D. (2004). Lateral inhibition goes social: Reducing the expression of the attractiveness stereotype. Unpublished Dissertation.
* Buczny, J., Layton, R. L., & Muraven, M. (2015). The role of implicit affective responses and trait self-control in ego resource management. *Motivation and Emotion*, *39*, 669-679.
Buhrmester, M. D., Blanton, H., & Swann, W. B. (2011). Implicit self-esteem: Nature, measurement, and a new way forward. *Journal of Personality and Social Psychology*, *100*, 365–385.
Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*, 550–558.
* Cadinu, M. Galdi, S. & Mass, A. (2013). Chameleonic social identities: Context induces shifts in homosexuals' self-stereotyping and self-categorization. *European Journal of Social Psychology*, *43*, 471-481.
* Cain, T. R. (2012). Fear bias: The impact of incidental fear on explicit and implicit risk perception. Unpublished dissertation.
* Calanchini, J., Gonsalkorale, K., Sherman, J.W. & Klauer, K.C. (2013). Counter-prejudicial training reduces activation of biased associations and enhances response monitoring. *European Journal of Social Psychology*, *43*, 321-325.
Calanchini, J. & Sherman, J. W. (2013). Implicit attitudes reflect associative, non-associative, and non-attitudinal processes. *Social and Personality Psychology Compass, 7*, 654-667.
Calanchini, J., Sherman, J., Klauer, C., & Lai, C. K. (2014). Attitudinal and non-attitudinal components of IAT performance. *Personality and Social Psychology Bulletin, 40*, 1285-1296.
Caldwell, D. M., Ades, A. E., & Higgins, J. P. T. (2005). Simultaneous comparison of multiple treatments: Combining direct and indirect evidence. *British Medical Journal, 331*, 897–900.
Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review, 16*, 330–350.
* Canadas, E., Lupianez, J., Rodriguez, R., Bayliss, A. P., & Tipper, S. P. (2007). I like you more if you look where I should: Gaze cueing and implicit prejudice change. Unpublished thesis.
Carlsson, R., & Agerström, J. (2016). A closer look at the discrimination outcomes in the IAT literature. *Scandinavian Journal of Psychology*, *57*, 278-287.
* Carpenter, S. J. (2001). Implicit gender attitudes. Unpublished dissertation.

* Castelli, L. & Tomelleri, S. (2008). Contextual effects on prejudiced attitudes: When the presence of others leads to more egalitarian responses. *Journal of Experimental Social Psychology*, *44*, 679-686.

* Castelli, L., Carraro, L., Pavan, G., Murelli, E., Carraro, A. (2012). The power of the unsaid: The influence of nonverbal cues on implicit attitudes. *Journal of Applied Social Psychology*, *42*, 1376-1393.

* Castillo, J. A., Camara, C. P., & Eguizabal, A. J. (2011). Prejudice reduction in university programs for older adults. *Educational Gerontology*, *37*, 164-190.

* Ceschi, G., Banse, R., & Van der Linden, M. (2009). Implicit but stable: Mental imagery changes explicit but not implicit anxiety. *Swiss Journal of Psychology*, *68*, 213-220.

Chaimani, A., Higgins, J. P. T., Mavridis, D., Spyridonos, P., & Salanti, G. (2013). Graphical Tools for Network Meta-Analysis in STATA. *PLoS ONE, 8*, e76654.

Chartrand, T. L., & Bargh, J. A. (2002). Nonconscious motivations: Their activation, operation, and consequences. In A. Tesser, D. Stapel, & J. Wood (Eds.), *Self and motivation: Emerging psychological perspectives* (pp. 13–41). Washington, DC: American Psychological Association.

* Chen, J. M., Cook, M., & Mackie, D. M. (2013). Imagined contact and prejudice reduction. Unpublished dataset.

Cheung, M. W. L. (2009). Comparison of methods for constructing confidence intervals of standardized indirect effects. *Behavior Research Methods, 41*, 425–438.

Cheung, M. W. L. (2015). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, 1521.

Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: a two-stage approach. *Psychological Methods, 10*, 40–64.

Cheung, M. W. L., & Cheung, S. F. (2016). Random-effects models for meta-analytic structural equation modeling: Review, issues, and illustrations. *Research Synthesis Methods, 7*, 140-155.

* Cheung, R. M., Noel, S., & Hardin, C. D. (2011). Adopting the system-justifying attitudes of others: Effects of trivial interpersonal connections in the context of social inclusion and exclusion. *Social Cognition*, *29*, 255-269.

* Cho, J. C. (2015). More than one way to become one: Specifying the relationship between perspective taking, self-other merging, and prosocial outcomes. Unpublished dataset.

* Cicchirillo, V. J. (2009). The effects of priming racial stereotypes through violent video games. Unpublished dissertation.

* Clerkin, E. M., & Teachman, B. A. (2010). Training implicit social anxiety associations: An experimental intervention.. *Journal of Anxiety Disorders*, *24*, 300-308.

* Colonnello, V. & Heinrichs, M. (2014). Intranasal oxytocin enhances positive self-attribution in healthy men. *Journal of Psychosomatic Research*, *77*, 415-419.

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology, 65,* 333–371.

* Columb, C., & Plant, E. A. (2011). Revisiting the Obama effect: Exposure to Obama reduces implicit prejudice. *Journal of Experimental Social Psychology*, *47*, 499-501.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: the quad model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469-487.

Cooper, J. (2007). *Cognitive dissonance: 50 years of a classic theory*. Thousand Oaks, CA: Sage Publications.

* Correll, J. (2005). Context, race and danger: The relationship between threat perception and the decision to shoot. Unpublished dissertation.

* Correll, J. (2008). 1/f noise and effort on implicit measures of bias. *Journal of Personality and Social Psychology*, *94*, 48-59.

* Correll, J., Park, B., & Smith, J. A. (2008). Colorblind and multicultural prejudice reduction strategies in high-conflict situations. *Group Processes & Intergroup Relations*, *11*, 471-491.

* Covin, R. (2008). Beyond the Emotion-Congruent Effect. Unpublished dissertation.

Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd Ed.). Boca Raton, FA: CRC Press.

* Craig, M. A. & Richeson, J. A. (2014). More Diverse Yet Less Tolerant? How the Increasingly Diverse Racial Landscape Affects White Americans' Racial Attitudes. *Personality and Social Psychology Bulletin*, *40*, 750-761.

* Craig, M. A. & Richeson, J. A. (2012). Coalition or Derogation? How perceived discrimination influences intraminority intergroup relations. *Journal of Personality and Social Psychology*, *102*, 759-777.

* Crisp, R. J., Stone, C. H., & Hall, N. R. (2006). Recategorization and subgroup identification: Predicting and preventing threats from common ingroups. *Personality and Social Psychology Bulletin*, *32*, 230-243.

Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, *6*, 475-494.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*, 349-354.

* Culhane, E. N. (2012). The trainability of cross-cultural competence and the examination of the affective component. Unpublished dissertation.

* Curl, L. S. (2002). Can we reduce our latent prejudice? An examination of the Asian Cultural Assimilator with the use of the Implicit Association Test. Unpublished dissertation.

* Dal Cin, S., Gibson, B., Zanna, M. P., Shumate, R., & Fong, G. T. (2007). Smoking in movies, implicit associations of smoking with the self, and intentions to smoke. *Psychological Science*, *18*, 559-563.

* Dasgupta, N. & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology*, *40*, 642-658.

Dasgupta, N. (2009). 13 mechanisms underlying the malleability of implicit prejudice and stereotypes: The role of automaticity and cognitive control. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 267–284). New York, NY: Psychology Press.

Dasgupta, N., DeSteno, D., Williams, L. A., & Hunsinger, M. (2009). Fanning the flames of prejudice: The influence of specific incidental emotions on implicit prejudice. *Emotion, 9*, 585–591.

* Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800-814.

Dasgupta, N., & Rivera, L. M. (2008). When social context matters: The influence of long–term contact and short–term exposure to admired outgroup members on implicit attitudes and behavioral intentions. *Social Cognition, 26*, 112–123.
De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20.
* De Lemus, S., Spears, R., Bukowski, M., Moya, M., & Lupiñez, J. (2013). Reversing implicit gender stereotype activation as a function of exposure to traditional gender roles. *Social Psychology*, *44*, 109-116.
* Delgado, N., Rodriguez-Perez, A., Vaes, J., Leyens, J., & Betancor, V. (2009). Priming effects of violence on infrahumanization. *Group Processes and Intergroup Relations*, *12*, 699-714.
* Dermody, N., Jones, M. K. & Cummin, S. R. (2013). The failure of imagined contact in reducing explicit and implicit out-group prejudice toward male homosexuals. *Current Psychology*, *32*, 261-274.
Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.
* Devos, T. & Yokoyama, Y. (2014). Silent or talking in the classroom: Implicit self-stereotyping among Asian and White students. *Basic and Applied Social Psychology*, *36*, 386-396.
* Devos, T., Gavin, K., & Quintana, F. J. (2010). Say “adios” to the American dream? The interplay between ethnic and national identity among Latino and Caucasian Americans. *Cultural Diversity and Ethnic Minority Psychology*, *16*, 37-49.
* Devos, T., Viera, E., Diaz, P. & Dunn, R. (2007). Influence of motherhood on the implicit academic self-concept of female college students: Distinct effects of subtle exposure to cues and directed thinking. *European Journal of Psychology and Education*, *22*, 371-386.
* Di Lemma, L. C. G., Dickson, J. M., Jedras, P., Roefs, A., & Field, M. (2015). Priming of conflicting motivational orientations in heavy drinkers: Robust effects on self-report but not implicit measures. *Frontiers in Psychology*, *6*.
Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *82*, 62-68.
* Duff, K. J. (1999). Controlling unwanted prejudice: Does practice make perfect? Unpublished dissertation.
* Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S. & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work? *Developmental Cognitive Neuroscience*, *4*, 38-51.
* Ebert, I. D., Steffens, M. C.,von Stulpnagel, R. & Jelenec, P. (2009). How to like yourself better, or chocolate less: Changing implicit attitudes with one IAT task. *Journal of Experimental Social Psychology*, *45*, 1098-1104.
Efthimiou, O., Mavridis, D., Riley, R. D., Cipriani, A., & Salanti, G. (2015). Joint synthesis of multiple correlated outcomes in networks of interventions. *Biostatistics, 16*, 84–97.
Efthimiou, O., Debray, T. P. A., van Valkenhoef, G., Trelle, S., Panayidou, K., Moons, K. G. M., … on behalf of GetReal Methods Review Group. (2016). GetReal in network meta-analysis: A review of the methodology. *Research Synthesis Methods*, *7*, 236–263.
Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

* Egloff, B. & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*, 1441-1455.
* Egloff, B., Weck, F. & Schmukle, S. C. (2008). Thinking about anxiety moderates the relationship between implicit and explicit anxiety measures. *Journal of Research in Personality*, *42*, 771-778.
* Enter, D., Spinhoven, P., & Roelofs, K. (2014). Alleviating social avoidance: Effects of single dose testosterone administration on approach-avoidance action. *Hormones and Behavior*, *65*, 351-354.
Evans, J. (1989). *Bias in human reasoning: Causes and consequences.* Brighton, UK: Erlbaum.
Fanelli, D., & Ioannidis, J. P. A. (2013). US studies may overestimate effect sizes in softer research. *Proceedings of the National Academy of Sciences, 110*, 15031-15036.
* Farmer, H., Maister, L., & Tsakiris, M. (2014). Change my body, change my mind: the effects of illusory ownership of an outgroup hand on implicit attitudes toward that outgroup. *Frontiers in Psychology*, *4*, 1016.
* Farrar, S., Stopa, L., & Turner, H. (2015). Self-imagery in individuals with high body dissatisfaction: The effect of positive and negative self-imagery on aspects of the self-concept. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 42-95.
Fazio, R. H., & Olson, M. A. (2014). The MODE model: Attitude-behavior processes as a function of motivation and opportunity. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual process theories of the social mind* (pp. 155–171). New York, NY: Guilford Press.
* Felix, K. J. (2014). Inter-religious relationships and anxiety in the regulation of automatic inter-religious prejudice. Unpublished dissertation.
* Ferguson, M. J., Zhang, Y., & Fishbach, A. (2015). How motivation shapes meaning: The effects of implicit object construal. Unpublished dissertation.
* Ferguson, M. J. (2008). On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness. *Journal of Personality and Social Psychology*, *95*, 1268-1294.
* Ferrara, K., Burns, J., & Mills, H. (2015). Public attitudes toward people with intellectual disabilities after viewing Olympic or Paralympic performance. *Adapted Physical Activity Quarterly*, *32*, 19-33.
* Fiedler, K. & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology*, *27*, 307-316.
* Field, M., Duka, T., Eastwood, B., Child, R., Santarcangelo, M., & Gayton, M. (2007). Experimental manipulation of attentional biases in heavy drinkers: Do the effects generalise? *Psychopharmacology*, *192*, 593-608.
Fishbach, A., Friedman, R. S., & Kruglanski, A. W. (2003). Leading us not into temptation: Momentary allurements elicit overriding goal activation. *Journal of Personality and Social Psychology, 84*, 296–309.
* Fishbach, A., Zhang, Y., & Trope, Y. (2010). Counteractive evaluation: Asymmetric shifts in the implicit value of conflicting motivations. *Journal of Experimental Social Psychology*, *46*, 29-38.
* Forbes, C. E. (2009). Lessons learned from "a clockwork orange": How retraining implicit attitudes and stereotypes affects motivation and performance under stereotype threat. Unpublished dissertation.

* Foroni, F., & Semin, G. R. (2012). Not all implicit measures of attitudes are created equal: Evidence from an embodiment perspective. *Journal of Experimental Social Psychology*, *48*, 424-427.
* Foroni, F., Mayr, U. (2005). The power of a story: New, automatic associations from a single reading of a short scenario. *Psychonomic Bulletin & Review*, *12*, 139-144.
Forscher, P. S., & Devine, P. G. (2014). Breaking the prejudice habit: Automaticity and control in the context of a long-term goal. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual process theories of the social mind* (pp. 468–482). New York, NY: Guilford Press.
Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T. L., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2017.04.009
* Franklin, J. C. (2014). The nature, development, and reversal of positive attitudes toward nonsuicidal self-injury: Implications for prediction and treatment. Unpublished dissertation.
* Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race Implicit Association Test as a stereotype threat experience. *Personality and Social Psychology Bulletin*, *30*, 1611-1624.
* French, A. R., Franz, T. M., Phelan, L. L. & Blaine, B. E. (2013). Reducing Muslim/Arab stereotypes through evaluative conditioning. *The Journal of Social Psychology*, *153*, 1-6.
* Fuhr, K., Hautzinger, M., & Meyer, T. D. (2015). Are social comparisons detrimental for the mood and self-esteem of individuals with an affective disorder? *Cognitive Therapy and Research*, *39*, 279-291.
* Fujita, K., & Han, H. A. (2009). Moving beyond deliberative control of impulses: The effect of construal levels on evaluative associations in self-control conflicts. *Psychological Science*, *20*, 799-804.
* Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The Roots of Stereotype Threat: When Automatic Associations Disrupt Girls' Math Performance. *Child Development*, *85*, 250-263.
* Galinsky, A. D. (1999). Perspective-taking: Debiasing social thought. Unpublished dissertation.
* Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*, 708-724.
* Gallate, J., Wong, C., Ellwood, S., Chi, R., & Snyder, A. (2011). Noninvasive brain stimulation reduces prejudice scores on an implicit association test. *Neuropsychology*, *2*, 185-192.
* Gan, Y. & Liu, J. (2012). The mechanism by which interpersonal coping flexibility influences self-esteem. *The Psychological Record*, *62*, 735-746.
* Gapinski, K. D., Schwartz, M. B., & Brownell, K. D. (2006). Can television change anti-fat attitudes and behavior? *Journal of Applied Biobehavioral Research*, *11*, 1-28.
* Gawronski, B. (2013). Manipulation of perceived discrimination. Unpublished data.
* Gawronski, B. & Bodenhausen, G. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation and retrieval experiences. *Journal of Personality and Social Psychology*, *89*, 672-685.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 59–128). Waltham, MA: Academic Press.

* Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorisation influence spontaneous evaluations of multiply categorizable objects? *Cognition and Emotion*, *24*, 1008-1025.

* Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology*, *44*, 1355-1361.

Gawronski, B., & Payne, B. K. (2010). *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*. New York, NY: Guilford Press.

Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, *139*, 683–701.

Gawronski, B., & Strack, F. (Eds.). (2012). *Cognitive consistency: A fundamental principle in social cognition*. New York: Guilford.

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 216–240). New York, NY: Guilford Press.

* Gawronski, B. & Ye, Y. (2015). Prevention of intention invention in the affect misattribution procedure. *Social Psychological and Personality Science*, *6*, 101-108.

* Geeraert, N. (2013). When suppressing one stereotype leads to rebound of another: On the procedural nature of stereotype rebound. *Personality and Social Psychology Bulletin*, *39*, 1173-1183.

* Geng, L., Liu, L., Xu, J., Zhou, K. & Fang, Y. (2013). Can evaluative conditioning change implicit attitudes towards recycling? *Social Behavior and Personality*, *4*, 947-956.

* Gibson, B. (2008). Can evaluative conditioning change attitudes toward mature brands? New evidence from the Implicit Association Test. *Journal of Consumer Research*, *35*, 178-188.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509–517.

* Giner-Sorolla, R. (2013). Anticipated and actual affect in prejudice control. Unpublished dataset.

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education, 5*, 351–379.

* Glen, I. S., & Banse, R. (2004). Probing the malleability of implicit and explicit self-esteem: An interview approach. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *22*, 133-158.

Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed., pp. 357–376). New York, NY: Russell Sage Foundation.

* Glock, S., Klapproth, F., & Müller, B. C. N. (2015). Promoting responsible drinking? A mass media campaign affects implicit but not explicit alcohol-related cognitions and attitudes. *British Journal of Health Psychology*, *20*, 482-497.
* Glock, S., Kovacs, C., & Unz, D. (2014). Implicit attitudes toward smoking: How the smell of cigarettes influences responses of college-age smokers and nonsmokers. *Journal of Health Psychology*, *19*, 629-641.
* Glock, S., Muller, B. C. N. & Krolak-Schwerdt, S. (2013). Implicit associations and compensatory health beliefs in smokers: Exploring their role for behaviour and their change through warning labels. *British Journal of Health Psychology*, *18*, 814-826.
* Gonsalkorale, K., Carlisle, K., & von Hippel, W. (2007). Intergroup threat increases implicit stereotyping. *International Journal of Psychology & Psychological Therapy*, *7*, 189-200.
* Goodall, C. E. (2010). Automatic attitude activation: Studies on processing and effects of alcohol advertisements and public service announcements. Unpublished dissertation.

Goodall, C. E., & Slater, M. D. (2010). Automatically activated attitudes as mechanisms for message effects: The case of alcohol advertisements. *Communication Research*, *37*, 620-643.

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology, 40,* 357–365.

* Govorun, O., & Payne, B. K. (2006). Ego-depletion and prejudice: Separating automatic and controlled components. *Social Cognition*, *24*, 111-136.

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine*, *22*, 1231-1238.

* Green, J. S., & Teachman, B. A. (2012). Is "cootie" in the eye of the beholder? An experimental attempt to modify implicit associations tied to contamination fears. *Journal of Experimental Psychopathology*, *3*, 479-495.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.

Greenwald, A. G., & Nosek, B. A. (2008). Attitudinal dissociation: What does it mean? In R. E. Petty, R. H. Fazio, & P. Brinol (Eds.), Attitudes: Insights from the New Implicit Measures (pp. 65-82) . Hillsdale, NJ: Erlbaum.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.

* Grumm, M., Nestler, S., & von Collani, G. (2009). Changing explicit and implicit attitudes: The case of self-esteem. *Journal of Experimental Social Psychology*, *45*, 327-335.
* Guendemir, S., Homan, A. C., de Dreu, C. K. W., & van Vugt, M. (2014). Think Leader, Think White? Capturing and Weakening an Implicit Pro-White Leadership Bias. *PLoS ONE*, *9*, e83915.

* Guinote, A., Willis, G. B., & Martellotta, C. (2010). Social power increases implicit prejudice.. *Journal of Experimental Social Psychology*, *46*, 299-307.
* Gulker, J. E. & Monteith, M. J. (2013). Intergroup boundaries and attitudes: The power of a single potent link. *Personality and Social Psychology Bulletin*, *39*, 943-955.
Haeffel, G., Abramson, L., Brazy, P., Shah, J., Teachman, B., & Nosek, B. (2007). Explicit and implicit cognition: A preliminary test of a dual-process theory of cognitive vulnerability to depression. *Behaviour Research and Therapy, 45*, 1155–1167.

* Haines, E., & Kray, L.J. (2005). Self-power associations: The possession of power impacts women's self-concepts. *European Journal of Social Psychology*, *35*, 643-662.
* Hall, N. R., Crisp, R. J., & Suen, M. (2009). Reducing implicit prejudice by blurring intergroup boundaries. *Basic and Applied Social Psychology*, *31*, 244-254.
* Han, H. A., Czellar, S., Olson, M. A., & Fazio, R. H. (2010). Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, *46*, 286-298.
* Harper, J. C. (2010). Modification of weight bias: Examining the effects of social influence on the expression of anti-fat attitudes. Unpublished dissertation.
* Harris, S. E. (1998). The linguistic intergroup bias: are differential expectations moderated by social identity concerns? Unpublished dissertation.
* Haynes, A., Kemps, E., & Moffitt, R. (2015). Inhibitory self-control moderates the effect of changed implicit food evaluations on snack food consumption. *Appetite*, *90*, 114-122.
Hedges, L. V., & Olkin, I. (1985). Statistical models for meta-analysis. *Psychological Methods, 6*, 203–217.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61–83.
* Hetts, J. J. (1999). Self-evaluations under fire: Implicit self-regard and explicit self-esteem in the face of failure. Unpublished dissertation.
Higgins, J., & Whitehead, A. (1996). Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine*, *15*, 2733-2749.
* Hofmann, W., Deutsch, R., Lancaster, K., & Banaji, M. R. (2010). Cooling the heat of temptation: Mental self-control and the automatic evaluation of tempting stimuli. *European Journal of Social Psychology*, *40*, 17-25.
Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality & Social Psychology Bulletin, 31*, 1369–1385.
Hofmann, W., Gschwendner, T., Friese, M., Wiers, R. W., & Schmitt, M. (2008). Working memory capacity and self-regulatory behavior: Towards an individual differences perspective on behavior determination by automatic versus controlled processes. *Journal of Personality and Social Psychology, 95*, 962-977.
* Hollands, G. J., Prestwich, A. & Marteau, T. M. (2011). Using aversive images to enhance healthy food choices and implicit attitudes: An experimental test of evaluative conditioning. *Health Psychology*, *30*, 195-203.
* Hooper, N., Villatte, M., Neofotistou, E., & McHugh, L. (2010). The effects of mindfulness versus thought suppression on implicit and explicit measures of experiential avoidance. *International Journal of Behavioral Consultation and Therapy*, *6*, 233-244.
* Horcajo, J., Briñol, P., & Petty, R. E. (2010). Consumer persuasion: Indirect change and implicit balance. *Psychology & Marketing*, *27*, 938-963.

* Houben, K. & Jansen, A. (2015). Chocolate equals stop. Chocolate-specific inhibition training reduces chocolate intake and go associations with chocolate. *Appetite*, *87*, 318-323.
* Houben, K., Havermans, R.C. & Wiers, R.W. (2010). Learning to dislike alcohol: Conditioning negative implicit attitudes toward alcohol and its effect on drinking behavior. *Psychopharmacology*, *211*, 79-86.
* Houben, K., Havermans, R.C., Nederkoorn, C., Jansen, A. (2012). Beer ? no-go: learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction*, *107*, 1280-1287.
* Hsueh, M., Yogeeswaran, K., & Malinen, S. (2014). 'Leave your comment below': Can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research*.
* Hugenberg, K., Blusiewicz, R. L., & Sacco, D. F. (2010). On malleable and immalleable subtypes: Stereotype malleability in one subtype does not spill over to other prominent subtypes. *Social Psychology*, *41*, 124-130.
* Huijding, J., & De Jong, P. J. (2007). Beyond fear and disgust: The role of (automatic) contamination-related associations in spider phobia. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 200-211.
* Huijding, J., de Jong, P. J., Wiers, R. W., & Verkooijen, K. (2006). Implicit and explicit attitudes toward smoking in a smoking and a nonsmoking setting. *Addictive Behaviors*, *30*, 949-961.
* Hulme, N., Hirsch, C. & Stopa, L. (2012). Images of the self and self-esteem: Do positive self-images improve self-esteem in social anxiety? *Cognitive Behaviour Therapy*, *41*, 163-173.
* Humphreys, A., & LaTour, K. A. (2013). Framing the game: Assessing the impact of cultural representations on consumer perceptions of legitimacy. *Journal of Consumer Research*, *40*, 773-795.
* Huntsinger, J. R. (2013). Affective incoherence reduces reliance on activated stereotypes. *Social Cognition*, *31*, 405-416.

Huntsinger, J. R., Isbell, L. M., & Clore, G. L. (2014). The affective control of thought: Malleable, not fixed. *Psychological Review, 121*, 600–618.

* Huntsinger, J. R., & Smith, C. T. (2009). First thought, best thought: Positive mood maintains and negative mood degrades implicit-explicit attitude correspondence. *Personality and Social Psychology Bulletin*, *35*, 187-197.
* Huntsinger, J. R., Sinclair, S., & Clore, G. L. (2009). Affective regulation of implicitly measured stereotypes and attitudes: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *45*, 560-566.
* Huntsinger, J. R., Sinclair, S., Dunn, E., & Clore, G. L. (2010). Affective regulation of stereotype activation: It's the (accessible) thought that counts. *Personality and Social Psychology Bulletin*, *36*, 564-577.

Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). Interpreting and inverting with less cursing: A guide to interpreting IRAP data. *Journal of Contextual Behavioral Science*, *4*, 157-162.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581–592.

* Imhoff, R. & Banse, R. (2009). Ongoing victim suffering increases prejudice: The case of secondary anti-semitism. *Psychological Science*, *20*, 1443-1447.

* Inzlicht, M., Gutsell, J. & Legault, L. (2012). Mimicry reduces racial prejudice. *Journal of Experimental Social Psychology*, *48*, 361-365.

* Ito, T. A., Chiao, K. W., Devine, P. G., Lorig, T. S. & Cacioppo, J. T. (2006). The influence of facial feedback on race bias. *Psychological Science*, *17*, 256-261.

Jackson, D., Riley, R., & White, I. R. (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, *30*, 2481-2498.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306–340.

James, W. (1890). *The Principles of Psychology*. New York, NY: Dover Publications.

* Janis, I. B. (2009). Approach/avoidance training to reduce health risk behaviors. Unpublished dissertation.

Jennions, M. D., & Møller, A. P. (2002). Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences, 269*, 43–48.

* Johnson, D. R., Jasper, D. M., Griffin, S. & Huffman, B. L. (2013). Reading narrative fiction reduces arab-muslim prejudice and offers a safe haven from intergroup anxiety. *Social Cognition*, *31*, 578-598.

* Johnson, I. R. (2009). Just say no (and mean it): Meaningful negation as a tool to modify automatic racial prejudice. Unpublished thesis.

* Jones, C. R., Vilensky, M. R., Vasey, M. W., & Fazio, R. H. (2013). Approach behavior can mitigate predominately univalent negative attitudes: Evidence regarding insects and spiders. *Emotion*, *13*, 989-996.

Jong, J., Halberstadt, J., & Bluemke, M. (2012). Foxhole atheism, revisited: The effects of mortality salience on explicit and implicit religious belief. *Journal of Experimental Social Psychology, 48*, 983–989.

* Joy-Gaba, J. A. & Nosek, B. A. (2010). The surprisingly limited malleability of implicit racial evaluations. *Social Psychology*, *41*, 137-146.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kang, J., & Banaji, M. R. (2006). Fair measures: A behavioral realist revision of" affirmative action". *California Law Review*, *94*, 1063-1118.

* Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology*, *81*, 774-788.

* Kashihara, J. (2015). Examination of stigmatizing beliefs about depression and stigma-reduction effects of education by using implicit measures. *Psychological Reports*, *116*, 337-362.

* Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology*, *78*, 871-888.

* Kawakami, K., Phills, C. E., Greenwald, A. G., Simard, D., Pontiero, J., Brnjas, A., Khan, B., Mills, J., & Dovidio, J. F. (2012). In perfect harmony: Synchronizing the self to activated social categories. *Journal of Personality and Social Psychology*, *102*, 562-575.

* Kawakami, K., Phills, C.E., Dovidio, J. F. & Steele, J. R. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, *92*, 957-971.

* Kawakami, K., Steele, J. R., Cifa, C., Phills, C. E. & Dovidio, J. F. (2008). Approaching math increases math = me and math = pleasant. *Journal of Experimental Social Psychology*, *44*, 818-825.
* Kemps, E., Tiggemann, M., Martin, R., & Elliott, M. (2013). Implicit approach/avoidance associations for craved food cues. *Journal of Experimental Psychology: Applied*, *19*, 30-38.
* Kenrick, A. C. (2015). Social tuning and shared reality: Downstream consequences in intergroup attitudes and relations. Unpublished dissertation.
* Kessels, U., Rau, M. & Hannover, B (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, *76*, 761-780.
* Kiefer, A. K. & Sekaquaptewa, D. (2006). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*, 825-832.
* Kiefer, A. K. (2006). The effects of stereotypes on performance attributions: How gender-math stereotypes lead women to internalize failure. Unpublished dissertation.

Kim, D. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66*, 83-96.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin, 31*, 208-217.

* Kleider, H. M. & Parrott, D. J. (2009). Aggressive shooting behavior: How working memory and threat influence shoot decisions. *Journal of Research in Personality*, *43*, 494-497.
* Koranyi, N. & Meissner, F. (2015). Handing over the reins: Neutralizing negative attitudes toward dependence in response to reciprocal romantic liking. *Social Psychological and Personality Science*, *6*, 685-691.
* Kroese, F. M., Adriaanse, M. A., Evers, C., & De Ridder, D. T. D. (2011). "Instant success": Turning temptations into cues for goal-directed behavior. *Personality and Social Psychology Bulletin*, *37*, 1389-1397.
* Kuppens, T., Pollet, T. V., Teixeira, C. P., Demoulin, S., Craig Roberts, S. & Little, A. C. (2012). Emotions in context: Anger causes ethnic bias but not gender bias in men but not women. *European Journal of Social Psychology*, *42*, 432-441.
* Lai, C. K., & Nosek, B. A. (2011). Race, Party, and Policy. Unpublished dataset.
* Lai, C. K., Haidt, J., & Nosek, B. A. (2014). Moral elevation reduces prejudice against gay men. *Cognition & Emotion*, *28*, 781-794.

Lai, C. K., Hoffman, K. M., & Nosek, B. A. (2013). Reducing implicit prejudice. *Social and Personality Psychology Compass, 7*, 315–330.

* Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., Sartori, G., Dial, C. M., Sriram, N., Banaji, M. R., & Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*, 1765-1785.
* Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., Hu, X., McLean, M. C., Axt, J. R., Asgari, S., Schmidt, K., Rubinstein., R, Marini, M., Rubichi, S., Shin,. J. L., &

Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*, 1001-1016.

* Lai, C. K., Todd, A. R., & Nosek, B. A. (2015). Perspective-taking and the Race IAT. Unpublished dataset.

* Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S; R. (2003). Stereotypes as dominant responses: On the "social facilitation" of prejudice in anticipated public contexts. *Journal of Personality and Social Psychology*, *84*, 277-295.

* Lambert, A. E. (2011). On attentional control and the elderly driving stereotype. Unpublished dissertation.

* Lambert, A. E., Seegmiller, J. K., Stefanucci, J. K. & Watson, J. M. (2013). On working memory capacity and implicit associations between advanced age and dangerous driving stereotypes. *Applied Cognitive Psychology*, *27*, 306-313.

* Laurin, K., Kay, A. C. & Fitzsimons, G. M. (2012). Divergent effects of activating thoughts of God on self-regulation. *Journal of Personality and Social Psychology*, *102*, 4-21.

* Lebens, H., Roefs, A., Martijn, C., Houben, K., Nederkoorn, C., & Jansen, A. (2011). Making implicit measures of associations with snack foods more negative through evaluative conditioning. *Eating Behaviors*, *12*, 249-253.

* Legault, L. (2010). Internalizing and automatizing motivation to be nonprejudiced: The role of self-determination in stereotyping, prejudice, and intergroup threat. Unpublished dissertation.

* Legault, L., & Green-Demers, I. (2012). The protective role of self-determined prejudice regulation in the relationship between intergroup threat and prejudice. *Motivation and Emotion*, *36*, 143-158.

* Legault, L., Green-Demers, I., & Eadie, A. L. (2009). When internalization leads to automatization: The role of self-determination in automatic stereotype suppression and implicit prejudice regulation. *Motivation and Emotion*, *33*, 12-24.

Lenton, A. P., Bruder, M., & Sedikides, C. (2009). A meta-analysis on the malleability of automatic gender stereotypes. *Psychology of Women Quarterly, 33*, 183–196.

* Lin, J. H. (2013). Identification matters: A moderated mediation model of media interactivity, character identification, and video game violence on aggression. *Journal of Communication*, *63*, 682-702.

* Lincoln, T. M, Arens, E., Berger, C., & Rief, W. (2008). Can antistigma campaigns be improved? A test of the impact of biogenetic vs psychosocial causal explanations on implicit and explicit attitudes to schizophrenia. *Schizophrenia Bulletin*, *34*, 984-994.

* Lindgren, K. P., Wiers, R. W., Teachman, B. A., Gasser, M. L., Westgate, E. C., Cousijn, J., Enkema, M. C., & Neighbors, C. (2015). Attempted training of alcohol approach and drinking identity associations in US undergraduate drinkers: Null results from two studies. *PLoS ONE*, *10*.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage Publications.

* Lochbuehler, K., Sargent, J. D., Scholte, R. H., Pieters, S., & Engels, R. C. (2012). Influence of Smoking Cues in Movies on Children's Beliefs About Smoking. *Pediatrics*, *130*, 221-227.

* Loggins, G. M. (2009). Can newscasts reduce prejudice? Television's potential impact upon the malleability of implicit attitudes. Unpublished dissertation.

Lu, G., & Ades, A. E. (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine, 23*, 3105–3124.

Lu, G., & Ades, A. (2009). Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics, 10*, 792–805.

* Lueke, A. & Gibson, B. (2015). Mindfulness meditation reduces implicit age and race bias: The role of reduced automaticity of responding. *Social Psychological and Personality Science*, *6*, 284-291.

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine, 21*, 2313–2324.

* Lun, J., Sinclair, S., Whitchurch, E. R., & Glenn, C. (2007). (Why) do I think what you think? Epistemic social tuning and implicit prejudice. *Journal of Personality and Social Psychology*, *93*, 957-972.

* Lybarger, J. E., & Monteith, M. J. (2011). The effect of Obama saliency on individual-level racial bias: Silver bullet or smokescreen? *Journal of Experimental Social Psychology*, *47*, 647-652.

* Macy, J. T., Chassin, L., Presson, C. C., & Sherman, J. W. (2015). Changing implicit attitudes toward smoking: Results from a web-based approach-avoidance practice intervention.. *Journal of Behavioral Medicine*, *38*, 143-152.

* Maister, L., Sebanz, N., Knoblich, G. & Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition*, *128*, 170-178.

* Malinen, S. & Johnston, L. (2007). The influence of an equity statement on perceivers' implicit and explicit associations between males and science. *New Zealand Journal of Psychology*, *36*, 18-24.

* Malinen, S. & Johnston, L. (2013). Workplace ageism: Discovering hidden bias. *Experimental Aging Research*, *39*, 445-465.

* Malinen, S. (2009). Implicit and explicit attitudes towards older workers: Their predictive utility and the role of attitude malleability. Unpublished dissertation.

* Malinen, S., Willis, G. M., & Johnston, L. (2014). Might informative media reporting of sexual offending influence community members' attitudes towards sex offenders? *Psychology, Crime & Law*, *20*, 535-552.

* Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, B., Neuberg, S. L., ... & Schaller, M. (2005). Functional projection: How fundamental social Motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, *88*, 63-78.

* Mann, N. H., & Kawakami, K. (2012). The long, steep path to equality: Progressing on egalitarian goals. *Journal of Experimental Psychology: General*, *141*, 187-197.

Mann, T. C., & Ferguson, M. J. (2017). Reversing implicit first impressions through reinterpretation after a two-day delay. *Journal of Experimental Social Psychology*, *68*, 122-127.

* Marini, M., Rubichi, S., & Sartori, G. (2012). The role of self-involvement in shifting IAT effects. *Experimental Psychology*, *59*, 349-354.

* Markland, D., Hall, C. R., Duncan, L. R., & Simatovic, J. (2015). The effects of an imagery intervention on implicit and explicit exercise attitudes. *Psychology of Sport and Exercise*, *17*, 24-31.

* Marshburn, C. K., & Knowles, E. D. (2015). White out of mind: Identity suppression as a coping strategy among Whites anticipating racially-charged interactions. Unpublished dataset.

* Martijn, C., Sheeran, P., Wesseldijk, L. W., Merrick, H., Webb, T. L., Roefs, A., & Jansen, A. (2012). Evaluative conditioning makes slim models less desirable as standards for comparison and increases body satisfaction. *Health Psychology*, *4*, 433-438.

* Martijn, C., Sheeran, P., Wesseldijk, L.W., Merrick, H., Webb, T.L., Roefs, A & Jansen, A. (2013). Evaluative conditioning makes slim models less desirable as standards for comparison and increases body satisfaction. *Health Psychology*, *32*, 433-438.

* Martin, B. A. S., Lee, M. S. W., & Lacey, C. (2011). Countering negative country of origin effects using imagery processing. *Journal of Consumer Behaviour*, *10*, 80-92.

* Mastro, D., Lapinski, M. K., Kopacz, M. A. & Behm-Morawitz, E. (2009). The influence of exposure of depictions of race and crime in TV news on viewer's social judgments. *Journal of Broadcasting & Electronic Media*, *53*, 615-635.

* Matheus, C. C. (2010). Assessing interventions for reducing gender-based occupational stereotypes: A multi-method study comparing the implicit association test to indirect and explicit measures of stereotypes, and an examination of sex roles and entity versus incremental lay theories of social perception. Unpublished dissertation.

Mauss, I. B., Bunge, S. A., & Gross, J. J. (2007). Automatic emotion regulation. *Social and Personality Psychology Compass, 1*, 146-167.

Mavridis, D., & Salanti, G. (2013). A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research*, *22*, 133-158.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435-442.

McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology*, *94*, 792-807.

* McGrane, J. A. & White, F. A. (2007). Differences in Anglo and Asian Australians' explicit and implicit prejudice and the attenuation of their implicit in-group bias. *Asian Journal of Social Psychology*, *19*, 204-210.

* Mellott, D. S. (2003). Measuring implicit attitudes and stereotypes: Increasing internal consistency reveals the convergent validity of IAT and priming measures. Unpublished dissertation.

* Mendoza, S. A. (2010). Planning Careful Responses to Racial Cues: The Effects of Implementation Intentions on Intergroup Bias. Unpublished dissertation.

* Mendoza, S. A., Gollwitzer, P. M., & Amodio, D. M. (2010). Reducing the expression of implicit stereotypes: Reflexive control through implementation intentions. *Personality and Social Psychology Bulletin*, *36*, 512-523.

Miller, D. T., Dannals, J. E., & Zlatev, J. J. (2017). Behavioral Processes in Long-Lag Intervention Studies. *Perspectives on Psychological Science*, *12*, 454–467.

* Miramontez, D. R. (2010). Examining implicit acculturation and bicultural identity integration. Unpublished dissertation.

Mogg, K., Bradley, B. P., De Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy, 35*, 297–303.

* Moore, S. G., Ferguson, M. J., & Chartrand, T. L. (2011). Affect in the aftermath: How goal pursuit influences implicit evaluations. *Cognition & Emotion*, *25*, 453-465.

Moran, T., Bar-Anan, Y., & Nosek, B. A. (2015). Processing goals moderate the effect of co-occurrence on automatic evaluation. *Journal of Experimental Social Psychology*, *60*, 157–162.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*, 364–386.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105–125.

* Motyl, M., Hart, J., Pyszczynski, T., Weise, D., Maxfield, M., & Siedel, A. (2011). Subtle priming of shared human experiences eliminates threat-induced negativity toward Arabs, immigrants, and peace-making. *Journal of Experimental Social Psychology*, *47*, 1178-1184.

* Mueller, J. S., Melwani, S. & Goncalo, J. A. (2012). The bias against creativity: Why people desire but reject creative ideas. *Psychological Science*, *23*, 13-17.

* Nicolas, G. & Skinner, A.L. (2012). “That's so gay!” Priming the general negative usage of the word gay increases implicit anti-gay bias. *Journal of Social Psychology*, *152*, 654-658.

* Noll, N.E. (2011). Moving situations: Not whether, but when and how arm flexion/extension relate to attitude change. Unpublished dissertation.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134,* 565-584.

Nosek, B. A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16*, 65-69.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press.

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion, 22*, 553-594.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*, 36–88.

* O’Brien, K. S., Puhl, R. M., Latner, J. D., Mir, A. S., & Hunter, J. A. (2010). Reducing anti-fat prejudice in preservice health students: A randomized trial. *Obesity*, *18*, 2138-2144.

* O’Connor, D. B., Hurling, R., Hendrickx, H., Osborne, G., Hall, J., Walklet, E., Whaley, A., & Wood, H. (2011). Effects of written emotional disclosure on implicit self-esteem and body image. *British Journal of Health Psychology*, *16*, 488-501.

Olson, M. A., & Fazio, R. H. (2002). Implicit Acquisition And Manifestation Of Classically Conditioned Attitudes. *Social Cognition*, *20*, 89–104.

* Olson, M. A. & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, *32*, 421-433.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology, 105*, 171-192.

* Park, J., Felix, K. & Lee, G. (2007). Implicit attitudes toward Arab-Muslims and the moderating effects of social information. *Basic and Applied Social Psychology*, *29*, 35-35.

* Park, J. K. & John, D. R. (2011). More than meets the eye: The influence of implicit and explicit self-esteem on materialism. *Journal of Consumer Psychology*, *21*, 73-87.
* Park, S. H. & Kim, H. J. (2015). Assumed race moderates spontaneous racial bias in a computer-based police simulation. *Asian Journal of Social Psychology*, *18*, 252-257.
* Park, S. H., Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice moderates the effect of cognitive depletion on unintended discrimination. *Social Cognition*, *26*, 401-419.

Payne, B. K. (2001). Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181-192.

* Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277-293.
* Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, *38*, 384-396.
* Peach, J. M., Yoshida, E., Spencer, S. J., Zanna, M. P. & Steele, J. R. (2011). Recognizing discrimination explicitly while denying it implicitly: Implicit social identity protection. *Journal of Experimental Social Psychology*, *47*, 283-292.

Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology, 44*, 29-45.

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 255-277). New York, NY: Guilford.

* Peters, K. R. & Gawronski, B. (2011). Mutual influences between the implicit and explicit self-concepts: The role of memory activation and motivated reasoning. *Journal of Experimental Social Psychology*, *47*, 436-442.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, *61*, 991–996.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Vol. 19 (pp. 123–205). New York: Academic.

* Phills, C. E., Kawakami, K., Tabi, E., Nadolny, D., & Inzlicht, M. (2011). Mind the gap: Increasing associations between the self and blacks with approach behaviors. *Journal of Personality and Social Psychology*, *100*, 197-201.
* Pietri, E., Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Roussos, G., & Handelsman, J. (2015). A novel intervention to influence perceptions of women in the sciences. Unpublished dataset.
* Pinkston, K. (2015). The Black-White malleability gap in implicit racial evaluations: A nationally representative study. *The Journal of Social Psychology*, *155*, 189-203.
* Plant, E. A., Peruche, B. M., & Butz, D. A. (2004). Eliminating automatic racial bias: Making race non-diagnostic for responses to criminal suspects. *Journal of Experimental Social Psychology*, *41*, 141-156.
* Preston, J. & Epley, N. (2009). Science and God: An automatic opposition between ultimate explanations. *Journal of Experimental Social Psychology*, *45*, 238-241.

* Quirin, M., Bode, R. C., Luckey, U., Pyszczynski, T., & Kuhl, J. (2014). Profound versus superficial coping with mortality threats: Action orientation moderates implicit but not explicit outgroup prejudice. *Personality and Social Psychology Bulletin*, *40*, 1132-1147.
* Rabbitt, S. M. (2012). The role of stress in parental perceptions of child noncompliance. Unpublished dissertation.
* Ralston, T. E., Palfai, T. P., & Rinck, M. (2015). Implicit Appetitive Responses to Alcohol Stimuli: Examining the Influence of Drinking Status and Mood on Approach and Avoidance Action Tendencies. Unpublished dataset.
* Ramasubramanian, S. (2007). Media-based strategies to reduce racial stereotypes activated by news stories. *J&MC Quarterly*, *84*, 249-264.
* Ramsey. L.R., Betz, D.E. & Sekaqueptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. *Social Psychology of Education*, *16*, 377-397.

Ranganath, K. A., & Nosek, B. A. (2008). Implicit attitude generalization occurs immediately; Explicit attitude generalization takes time. *Psychological Science*, *19*, 249-154.

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7*, 331–363.

* Richeson, J. A. & Ambady, N. (2003). Effects of situational power on automatic racial prejudice. *Journal of Experimental Social Psychology*, *39*, 177-183.
* Richeson, J. A., & Ambady, N. (2001). Who's in charge? Effects of situational roles on automatic gender bias. *Sex Roles*, *44*, 493-512.
* Richeson, J. A., & Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology*, *40*, 417-423.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*, 1043–1056.

* Rohner, J., Schroder-Abe, M., & Schutz, A. (2011). Exaggeration is harder than understatement, but practice makes perfect! Faking success in the IAT. *Experimental Psychology*, *58*, 464-472.
* Rohner, J., Schroder-Abe, M., Schutz, A. (2013). What do fakers actually do to fake the IAT? An investigation of faking strategies under different faking conditions. *Journal of Research in Personality*, *47*, 330-338.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638-641.

* Rothmund, T. & Baumert, A. (2014). Shame on me: Implicit assessment of negative moral self-evaluation in shame-proneness. *Social Psychological and Personality Science*, *5*, 195-202.
* Rotteveel, M. & Phaf, R. H. (2004). Loading working memory enhances affective priming. *Psychonomic Bulletin & Review*, *11*, 326-331.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). “Unlearning” automatic biases: the malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*, 856–868.

* Rudman, L. A., & Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Processes & Intergroup Relations*, *5*, 133-150.
* Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology*, *93*, 798-813.

* Rudman, L. A. & Phelan, J. E. (2010). The effect of priming gender roles on women's implicit gender beliefs and career aspirations. *Social Psychology*, *41*, 192-202.
* Rydell, R.J., Hamilton, D.L. & Devos, T. (2010). Now they are American, now they are not: Valence as a determinant of the inclusion of African Americans in the American identity. *Social Cognition*, *28*, 161-179.
Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods, 3,* 80–97.
Salanti, G., Higgins, J. P., Ades, A., & Ioannidis, J. P. (2008). Evaluation of networks of randomized trials. *Statistical Methods in Medical Research, 17,* 279–301.
* Saleem, M. & Anderson, C.A. (2013). Arabs as terrorists: Effects of stereotypes within violent contexts on attitudes, perceptions and affect. *Psychology of Violence*, *3*, 84-99.
* Saleem, M., Prot, S., & Anderson, C. A. (2013). Americans != Arabs/Muslims: The incompatibility of American and Arab/Muslim identities. Unpublished dataset.
Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8,* 448–467.
* Saporito, J. M., Ryan, C., Teachman, B. A. (2009). Reducing stigma toward seeking mental health treatment. *Stigma Research and Action*, *2*, 9-21.
* Sassenberg, K., & Moskowitz, G. B. (2005). Don't stereotype, think different! Overcoming automatic stereotype activation by mindset priming. *Journal of Experimental Social Psychology*, *41*, 506-514.
* Sassenberg, K., & Wieber, F. (2005). Don't ignore the other half: The impact of ingroup identification on implicit measures of prejudice. *European Journal of Social Psychology*, *35*, 621-632.
* Schlenkermann, R. S. (2007). Negative evaluative conditioning of implicit smoking attitudes as an avenue toward cessation. Unpublished dissertation.
* Schmidt, A. F., Zimmermann, P. S., Banse, R., & Imhoff, R. (2015). Ego depletion moderates the influence of automatic and controlled precursors of reactive aggression: A double dissociation. *Social Psychology*, *46*, 132-141.
Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.
Schnabel, K., & Asendorpf, J. B. (2010). The self-concept: New insights from implicit measurement procedures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition* (pp. 408-425). New York, NY: Guilford Press.
* Schmulke, S.C. & Egloff, B. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance? *European Journal of Personality*, *18*, 483-494.
* Schoel, C., Bluemke, M., Mueller, P., & Stahlberg, D. (2011). When autocratic leaders become an option? Uncertainty and self-esteem predict implicit leadership preferences. *Journal of Personality and Social Psychology*, *101*, 521-540.
* Schultz, P.W. & Tabanico, J. (2007). Self, identity and natural environment: Exploring implicit connections with nature. *Journal of Applied Social Psychology*, *37*, 1219-1247.
* Schwab, A. K. & Greitemeyer, T. (2015). Failing to establish evaluative conditioning effects for indirect intergroup contact on Facebook. *Basic and Applied Social Psychology*, *37*, 87-104.
* Scott-Sheldon, L. A. (2006). Implicit-explicit ambivalent attitudes and health behaviors. Unpublished dissertation.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515–530.

* Seger, C. R. (2010). Literal intergroup contact: Embodied relational cues and the reduction of intergroup bias. Unpublished dissertation.

Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology, 39*, 75–82.

* Sharbanee, J. M., Hu, L., Stritzke, W. G. K., Wiers, R. W., Rinck, M., & MacLeod, C. (2014). The Effect of Approach/Avoidance Training on Alcohol Consumption Is Mediated by Change in Alcohol Action Tendency. *PLoS ONE*, *9*, e85855.

* Shariff, A. F. A., Cohen, A. B., & Norenzayan, A. (2008). The devil's advocate: Secular arguments diminish both implicit and explicit religious belief. *Journal of Cognition and Culture*, *8*, 3-4.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84,* 127–190.

* Shih, M. J., Stotzer, R., & Gutierrez, A. S. (2013). Perspective-taking and empathy: Generalizing the reduction of group bias towards Asian Americans to general outgroups.. *Asian American Journal of Psychology*, *4*, 79-83.

Shook, N. J., & Fazio, R. H. (2008). Interracial roommate relationships: An experimental field test of the contact hypothesis. *Psychological Science*, *19*, 717-723.

* Sibley, C. G. & Harre, N (2009). The impact of different styles of traffic safety advertisement on young drivers' explicit and implicit self-enhancement biases. *Transportation Research Part F*, *12*, 159-167.

* Siegel, E. F., Dougherty, M. R. & Huber, D. E. (2012). Manipulating the role of cognitive control while taking the implicit association test. *Journal of Experimental Social Psychology*, *48*, 1057-1068.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*, 534–547.

Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: the role of affiliative motivation. *Journal of Personality and Social Psychology, 89,* 583–592.

* Skorek, M., & Dunham, Y. (2012). Self-enhancement following exposure to idealized body portrayals in ethnically diverse men: A fantasy effect of advertising. *Sex Roles*, *66*, 655-667.

* Skorinko, J. L. M., Lun, J., Sinclair, S., Marotta, S. A., Calanchini, J., & Paris, M. H. (2015). Reducing prejudice across cultures via social tuning. *Social Psychological and Personality Science*, *6*, 363-372.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22.

* Smith, C. T. & De Houwer, J. (2015). Hooked on a feeling: Affective anti-smoking messages are more effective than cognitive messages at changing implicit evaluations of smoking.. *Frontiers in Psychology*, *6*.

Smith, C. T., De Houwer, J., & Nosek, B. A. (2013). Consider the source: Persuasion of implicit evaluations is moderated by manipulations of source credibility. *Personality and Social Psychology Bulletin, 39,* 193-205.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108–131.

* Smith, P. K., McCulloch, K. C., & Schouwstra, A. (2013). Moving closer to reach the top: approach behavior increases one's sense of power. *Social Cognition*, *31*, 518-529.

* Smith, P. K., Dijksterhuis, A. & Chaiken, S. (2008). Subliminal exposure to faces and racial attitudes: Exposure to Whites makes Whites like Blacks less. *Journal of Experimental Social Psychology*, *44*, 50-64.

Smith, C. T., Ratliff, K. A., & Nosek, B. A. (2012). Rapid Assimilation: Automatically Integrating New Information with Existing Beliefs. *Social Cognition*, *30*, 199–219.

* Smurda, J. D., Wittig, M. A., & Gokalp, G. (2006). Effects of threat to a valued social identity on implicit self-esteem and discrimination. *Group Processes & Intergroup Relations*, *9*, 181-197.

* Soderberg, C. K., & Sherman, J. W. (2011). The hidden costs of multiculturalism: Can celebrating our differences lead to greater discrimination? Unpublished dissertation.

* Spalding, L. R. (1998). The hidden costs of attributional ambiguity: How experiencing discrimination affects implicit self-esteem. Unpublished dissertation.

* Sportel, B. E., de Hullu, E., de Jong, P. J., & Nauta, M. H. (2013). Cognitive bias modification versus CBT in reducing adolescent social anxiety: a randomized control trial. *PLoS ONE*, *8*, e64358.

Sritharan, R., & Gawronski, B. (2010). Changing Implicit and Explicit Prejudice: Insights from the Associative-Propositional Evaluation Model. *Social Psychology, 41,* 113–123.

* Stamm, K. E. (2009). Stereotype threat and implicit attitudes: Implications for the leaky pipeline of women in science. Unpublished dissertation.

Stangor, C., & Duan, C. (1991). Effects of multiple task demands upon memory for information about social groups. *Journal of Experimental Social Psychology, 27*, 357–378.

Stanovich, K. E., & West, R. F. (2002). Individual differences in reasoning: Implications for the rationality debate. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 421–440). New York: Cambridge University Press.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In *Advances in Experimental Social Psychology* (Vol. 21, pp. 261–302). San Diego, CA: Academic Press.

* Steele, J. R., & Ambady, N. (2006). Math is Hard! The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology*, *42*, 428-436.

* Stepanova, E. V., Bartholow, B. D., Saults, J. S., & Friedman, R. S. (2012). Alcohol-related cues promote automatic racial bias. *Journal of Experimental Social Psychology*, *48*, 905-911.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reduce prejudice and discrimination* (pp. 23–46). Mahwah, NJ: Lawrence Erlbaum Associates.

Stern, J. M., & Simes, R. J. (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal, 315,* 640–645.

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 99-110). West Sussex, England: John Wiley & Sons, Ltd.
* Stewart, B. D. & Payne, B. K. (2007). Implementation intentions. Unpublished dataset.
* Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as an efficient means of thought control. *Personality and Social Psychology Bulletin*, *34*, 1332-1345.
* Stewart, T. L., Latu, I. M., Kawakami, K., & Myers, A. C. (2010). Consider the situation: Reducing automatic stereotyping through situational attribution training. *Journal of Experimental Social Psychology*, *46*, 221-225.
* Stewart, T. L., Latu, I. M., Kawakami, K., Walsh, S. P., & Schmidt, A. (2016). Reducing Implicit Racial Stereotyping for More than a Day Through Situational Attribution Training. Unpublished dataset.
* Stieger, S., Goritz, A.S., Hergovich, A. & Voracek, M. (2011). Intentional faking of the single category Implicit Association Test and the Implicit Association Test. *Psychological Reports*, *109*, 219-230.
* Sumner, M. D. (2003). Nonconscious routes for avoiding stereotypic rebound: A focus on chronic egalitarian motivation and implementation intentions. Unpublished dissertation.
Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology, 33,* 1–39.
* Tang, J., Wu, S., & Miao, D. (2013). Experimental test of escape theory: Accessibility to implicit suicidal mind. *Suicide and Life-Threatening Behavior*, *43*, 347-355.
* Taylor, A. (2011). The influence of target race on split-second shooting decisions in simulated scenarios: a Canadian perspective. Unpublished dissertation.
Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.
Taylor, S. E., & Lobel, M. (1989). Social comparison activity under threat: Downward evaluation and upward contacts. *Psychological Review, 96,* 569–575.
* Teachman, B. A., Gapinski, K. D., Brownell, K. D., Rawlins, M., & Jeyaram, S. (2003). Demonstrations of implicit anti-fat bias: The impact of providing causal information and evoking empathy. *Health Psychology*, *22*, 68-78.
* Teachman, B. A., Woody, S. R., & Magee, J. C. (2006). Implicit and explicit appraisals of the importance of intrusive thoughts. *Behaviour Research and Therapy*, *44*, 785-805.
* Terbeck, S. Kahane, G., McTavish, S., Savulescu, J., Cowen, P.J. & Hewstone, M. (2012). Propranolol reduces implicit negative racial bias. *Psychopharmacology*, *222*, 419-424.
* Terbeck, S., Kahane, G., McTavish, S., McCutcheon, R., Hewstone, M., Savulescu, J., Chesterman, L. P., Cowen, P. J., & Norbury, R. (2015). β-Adrenoceptor blockade modulates fusiform gyrus activity to black versus white faces.. *Psychopharmacology*, *232*, 2951-2958.
* Thomas, J. J., Judge, A. M., Brownell, K. D. & Vartanian, L. R. (2006). Evaluating the effects of eating disorder memoirs on readers' eating attitudes and behaviors. *International Journal of Eating Disorders*, *39*, 418-425.
* Tinkler, J. E. (2013). How do sexual harassment policies shape gender beliefs? An exploration of the moderating effects of norm adherence and gender. *Social Science Research*, *42*, 1269-1283.

* Tinkler, J. E., Li, Y., E., & Mollborn, S. (2007). Can legal interventions change beliefs? The effect of exposure to sexual harrassment policy on men's gender beliefs. *Social Psychology Quarterly*, *70*, 480-494.
* Todd, A. R. & Burgmer, P. (2013). Perspective taking and automatic intergroup evaluation change: Testing an associative self-anchoring account. *Journal of Personality and Social Psychology*, *104*, 786-802.
* Todd, A. R. (2014). Perspective taking and automatic evaluation. Unpublished dataset.
* Todd, A. R., & Simpson, A. J. (2014). Target gender moderates the effect of perspective taking on implicit racial bias. Unpublished dataset.
* Todd, A. R., Bodenhausen, G. V., & Galinsky, A. D. (2012). Perspective taking combats the denial of intergroup discrimination. *Journal of Experimental Social Psychology*, *48*, 738-745.
* Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, *100*, 1027-1042.
* Treloar, H.R. & McCarthy, D. (2012). Effects of mood and urgency on activation of general and specific alcohol expectancies. *Addictive Behaviors*, *37*, 115-118.
* Turner, R.N. & Crisp, R.J. (2010). Imagining intergroup contact reduces implicit prejudice. *British Journal of Social Psychology*, *49*, 129-142.
* Uhlmann, E., & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence*, *27*, 41-52.
* Van Dessel, P. , De Houwer, J. , Gast, A., & Smith C. T. (2016). Effects of Approach-avoidance Instructions and Training on Implicit Prejudice. Unpublished dataset.
* Van Dessel, P., De Houwer, J., Gast, A., & Tucker Smith, C. (2015). Instruction-based approach-avoidance effects: Changing stimulus evaluation via the mere instruction to approach or avoid stimuli. *Experimental Psychology*, *62*, 161-169.
* Van Dessel, P., De Houwer, J., Roets, A., & Gast, A. (2016). Failures to change stimulus evaluations by means of subliminal approach and avoidance training. *Journal of Personality and Social Psychology*, *110*, 1-15.
* van Nunspeet, F., Ellemers, N., Derks, B., & Nieuwenhuis, S. (2014). Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social Cognitive and Affective Neuroscience*, *9*, 141-149.
* Vezzali, L., Adrighetto, L., Trifiletti, E., & Visintin, E. P. (2012). Perceiving status (in)stability in a low-status group: The effects of identification on explicit and implicit intergroup attitudes. *Social Psychology*, *43*, 33-40.
* Vezzali, L., Capozza, D., Giovannini, D., & Stathi, S. (2011). Improving implicit and explicit intergroup attitudes using imagined contact: An experimental intervention with elementary school children. *Group Processes & Intergroup Relations*, *15*, 203-212.
* Vianello, M. (2014). Replication of Dasgupta and Greenwald 2001 Procedure. Unpublished dataset.
* Vrijsen, J. N., Oostrom, I., Speckens, A., Becker, E. S., & Rinck, M. (2013). Approach and Avoidance of Emotional Faces in Happy and Sad Mood. *Cognitive Therapy and Research*, *37*, 1-6.
* Wagar, B. M. (2006). The good and the bad of affective guidance: Insights from the influence of automatic racial prejudice on complex decisions under uncertainty. Unpublished dissertation.

* Wallaert, M., Ward, A. & Mann, T. (2011). Explicit control of implicit responses: Simple directives can alter IAT performance. *Social Psychology*, *41*, 152-157.
* Walliser, J., Tulk, S., Hertz, N., Issler, E., & Wiese, E. (2015). Effects of Perspective Taking on Implicit Attitudes and Performance in Economic Games. *Social Robotics (ICSR 2015)*, *9388*, 684-693.
* Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two Brief Interventions to Mitigate a "Chilly Climate" Transform Women's Experience, Relationships, and Achievement in Engineering. *Journal of Educational Psychology*, *107*, 468-485.
* Watamura, E., Wakebe, T., Fujio, M., Itoh, Y., & Karasawa, K. (2014). The automatic activation of retributive motive when determining punishment. *Psychological Studies*, *59*, 236-240.

Wei, Y., & Higgins, J. P. T. (2013). Estimating within-study covariances in multivariate meta-analysis with multiple outcomes. *Statistics in Medicine, 32,* 1191–1205.

* Wennekers, A. M., Holland, R. W., Wigboldus, D. H. J., & Van Knippenberg, A. (2012). First see, then nod: The role of temporal contiguity in embodied evaluative conditioning of social attitudes. *Social Psychological and Personality Science*, *3*, 455-461.

Wheeler, M. E., & Fiske, S. T. (2005). Controlling racial prejudice: Social-cognitive goals affect amygdala and stereotype activation. *Psychological Science, 16,* 56–63.

White, I. R., Barrett, J. K., Jackson, D., & Higgins, J. P. T. (2012). Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods, 3,* 111–125.

Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social issues*, *25*, 41-78.

* Wiers, C. E., Ludwig, V. U., Gladwin, T. E., Park, S. Q., Heinz, A., Wiers, R. W., Rinck, M., Lindenmeyer, J., Walter, H., & Bermpohl, F. (2015). Effects of cognitive bias modification training on neural signatures of alcohol approach tendencies in male alcohol-dependent patients. *Addiction Biology*, *20*, 990-999.
* Wiers, C. E., Stelzel, C., Gladwin, T. E., Park, S. Q., Pawelczack, S., Gawron, C. K., Stuke, H., Heinz, A., Wiers, R. W., Rinck, M., Lindenmeyer, J., Walter, H., & Bermpohl, F. (2015). Effects of cognitive bias modification training on neural alcohol cue reactivity in alcohol dependence. *The American Journal of Psychiatry*, *172*, 335-343.
* Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science*, *22*, 490-497.
* Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*, 279-287.
* Williamson, W. P., Bishop, J., & Hood, R. W. J. (2014). Religious fundamentalism and racial prejudice: A comparison of implicit and explicit approaches. *Mental Health, Religion & Culture*, *17*, 847-859.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126.

* Wirth, J. H, Sacco, D. F, Hugenberg, K. & Williams, K. D. (2010). Eye gaze as relational evaluation: Averted eye gaze leads to feelings of ostracism and relational devaluation. *Personality and Social Psychology Bulletin*, *36*, 869-882.

* Wittenbrink, B., Judd, C. M. & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815-827.
* Wolff, W., Schindler, S., & Brand, R. (2015). The Effect of Implicitly Incentivized Faking on Explicit and Implicit Measures of Doping Attitude: When Athletes Want to Pretend and Even More Negative Attitude to Doping. *PLoS ONE*, *10*, e0118507.
* Wong, C., Harris, J., & Gallate, J. (2012). Evidence for a social function of the anterior temporal lobes: Low-frequency rTMS reduces implicit gender stereotypes. *Social Neuroscience*, *7*, 90-104.
* Woodcock, A. & Montieth, M. J. (2013). Forging links with the self to combat implicit bias. *Group Processes & Intergroup Relations*, *16*, 445-461.
* Woodcock, A. & Montieth, M. J. (2013). Forging links with the self to combat implicit bias. *Group Processes & Intergroup Relations*, *16*, 445-464.
* Woud, M. L., Hutschemaekers, M. H., Rinck, M., & Becker, E. S. (2015). The manipulation of alcohol-related interpretation biases by means of Cognitive Bias Modification-Interpretation (CBM-I). *Journal of Behavior Therapy and Experimental Psychiatry*, *49*, 61-68.
* Wyer, N. A. (2010). Salient egalitarian norms moderate activation of out-group approach and avoidance. *Group Processes & Intergroup Relations*, *13*, 151-165.
* Yang, J., Shi, Y., Luo, Y. L. L., Shi, J., & Cai, H. (2014). The Brief Implicit Association Test is valid: Experimental evidence. *Social Cognition*, *32*, 449-465.
* Yogeeswaran, K., Adelman, L., Parker, M. T., & Dasgupta, N. (2014). In the eyes of the beholder: National identification predicts differential reactions to ethnic identity expressions. *Cultural Diversity and Ethnic Minority Psychology*, *20*, 362-369.
* Yogeeswaran, K., Dasgupta, N. & Gomez, C. (2012). A new American dilemma? The effect of ethnic identification and public service on the national inclusion of ethnic minorities. *European Journal of Social Psychology*, *42*, 691-705.
* Yogeeswaran, K., Dasgupta, N., Adelman, L., Eccleston, A., Parker, M. T. (2011). To be or not to be (ethnic): Public vs. private expressions of ethnic identification differentially impact national inclusion of White and non-White groups. *Journal of Experimental Social Psychology*, *47*, 908?914.
* Yoshida, E., Peach, J. M., Zanna, M. P., & Spencer, S. J. (2011). Not all automatic associations are created equal: How implicit normative evaluations are distinct from implicit attitudes and uniquely predict meaningful behavior. *Journal of Experimental Social Psychology*, *48*, 694-706.
* Zisserson, R.N. (2011). Reducing the impact of temptation cues on health-related goal cognitions among college student cigarette smokers. Unpublished dissertation.
* Zitelny, H., & Bar-Anan, Y. (2014). Gender and Math Experiments. Unpublished dataset.
* Zogmaister, C., Arcuri, Luciano, Castelli, L., & Smith, E. R. (2008). The impact of loyalty and equality on implicit ingroup favoritism. *Group Processes & Intergroup Relations*, *11*, 493-512.
* Zvonkovic, A. & Lucas-Thompson, R. G. (2015). Refuting the myth of the 'violent schizophrenic': Assessing an educational intervention to reduce schizophrenia stigmatization using self-report and an Implicit Association Test. *Social Work in Mental Health*, *13*, 201-215.