Demonstrations of Implicit Anti-Fat Bias: The Impact of Providing Causal Information and Evoking Empathy

Bethany A. Teachman
University of Virginia

Kathrine D. Gapinski, Kelly D. Brownell, and Melissa Rawlins
Yale University

Subathra Jeyaram
Oxford University

Three studies investigated implicit biases, and their modifiability, against overweight persons. In Study 1 (N = 144), the authors demonstrated strong implicit anti-fat attitudes and stereotypes using the Implicit Association Test, despite no explicit anti-fat bias. When participants were informed that obesity is caused predominantly by overeating and lack of exercise, higher implicit bias relative to controls was produced; informing participants that obesity is mainly due to genetic factors did not result in lower bias. In Studies 2A (N = 90) and 2B (N = 63), participants read stories of discrimination against obese persons to evoke empathy. This did not lead to lower bias compared with controls but did produce diminished implicit bias among overweight participants, suggesting an in-group bias.

Key words: obesity, anti-fat, bias, stereotype, attitude, implicit
ponents of the attitude response, which is why they can be variably related.

Little is known about the pervasiveness of implicit anti-fat biases in the general population or about the relationship among implicit and explicit anti-fat attitudes and stereotypes. Even less is known about changing these implicit attitudes in the service of reducing bias, although there is now mounting evidence that some implicit biases can be modified (at least temporarily; e.g., race bias; Dasgupta & Greenwald, 2001). This article presents the results of three studies that evaluated implicit and explicit anti-fat biases among general and student populations and investigated two theoretically derived approaches to changing attitudes. The literature suggests that anti-fat bias is especially strong because being overweight is deemed blameworthy (e.g., Crandall, 1994). Further, the fact that weight is thought to be controllable is likely to reduce empathy toward obese persons. Consequently, we hypothesized that manipulations that reduced blame (by manipulating the controllability of obesity) or increased empathy could shift anti-fat attitudes. On the basis of societal messages that being thin is beautiful and being fat is abhorrent, we expected that the general population would demonstrate strong implicit biases against obese persons. However, given social norms to appear tolerant, we expected little explicit report of anti-fat bias.

Study 1

Study 1 was designed to investigate implicit attitudes and stereotypes toward obese persons in the general population, to determine how these attitudes relate to explicit bias, and to evaluate whether information about the causes of obesity would influence the expression of implicit and explicit biases.

Evaluating perceived controllability as a bias-reduction strategy derives from work by Crandall and colleagues, who suggested that anti-fat attitudes result from holding individuals responsible for their condition (Crandall, 1994; Crandall & Biernat, 1990). Several studies have indicated that reducing the perceived control overweight persons are judged to have over their weight may decrease explicit anti-fat bias (e.g., DeJong, 1980, 1993). For example, Crandall (1994) was able to change the common opinion that overweight people lack willpower by educating participants about the contributions of genetics and physiology in obesity (thereby reducing blame). It is not known whether implicit bias can be changed. Given this evidence that causal attributions can influence bias and research suggesting that the general public is misinformed about the true causes of obesity, Study 1 examined whether providing information about the causes of obesity could affect both implicit and explicit anti-fat biases. We hypothesized that bias would be lower when participants were informed that the primary cause of obesity was genetics (i.e., a cause outside of the person’s control) and higher when the principal causes were indicated as overeating and lack of exercise (i.e., causes perceived to lie within the person’s control).

Method

Participants

A large tent was set up at a Connecticut beach, and participants who approached the tent or were walking by were invited to take part in a short survey about their attitudes toward different groups in exchange for either $1, a lottery ticket, or a beverage. The refusal rate was less than 5% after people approached the tent. This locale was used because it is an environment where weight and shape are salient, increasing the accessibility of anti-fat attitudes. Further, other studies conducted at the beach using implicit attitude measurement found comparable effects for attitudes measured in the laboratory on political preferences and sexual orientation (B. A. Nosek, personal communication, August 5, 2001). Individuals who approached the tent were screened to determine that they were over 18 years old and could read English fluently, and then they completed the study either alone or in a group of up to three people. The sample (N = 144) was approximately equal for gender (54% female), had a mean age of 35 years (SD = 13.99; range = 18–78 years), was predominantly Caucasian (89%), and had a mean body mass index (BMI) of 26.37 (SD = 4.71) and 24.44 (SD = 4.61) for men and women, respectively.

Materials

Cause of obesity primes. Participants were assigned to one of three prime conditions. Either they received no prime (no-prime condition, N = 48) or they were asked to read a “recently published news article” that reported the results of a research study indicating that the primary cause of obesity was genetics (genetics condition, N = 48) or the primary causes of obesity were overeating and lack of exercise (behavior condition, N = 48). This design therefore manipulated whether obesity was perceived to be predominantly within or outside personal control. To make the manipulation believable, we designed the prime to indicate that 80% of the cause of obesity could be explained by either genetics or the behavioral factors, and the remaining 20% was due to the other factor (i.e., genetics or behavior, whichever was not stated as the primary cause).

Implicit bias. The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is a measure that has been used to reflect implicit attitudes primarily related to social prejudice, such as gender stereotypes (e.g., Rudman, Greenwald, & McGhee, 1996) and racial evaluations (e.g., Dasgupta, McGhee, Greenwald, & Banaji, 2000), and has only recently been applied to clinical research (e.g., Teachman, Gregg, & Woody, 2001). The IAT uses reaction time to measure implicit memory-based associations without requiring conscious introspection. Processing speed is assumed to be an indirect measure of the individual’s degree of association between two concepts, and the degree of implicit association is interpreted as an index of a person’s unconscious or implicit attitude. The measure is implicit in the sense that it measures evaluations that occur outside conscious control and, at times, outside conscious awareness. The IAT instructs participants to classify words into superordinate categories (categories that are at a more general level). For example, in the practice task for the current study, participants decided whether words such as daisies, tulips, bugs, and mosquitoes belonged to the superordinate category flowers or insects. Simultaneously, they classified words associated with the descriptive categories good and bad. Participants classified words from the four categories under two different conditions. In one condition, the category labels flowers plus good versus insects plus bad were paired; in the other condition, the labels were switched, so flowers plus bad versus insects plus good were paired together. Participants generally categorized stimuli faster when the paired categories matched the way they

1 Supporting this assertion, we found higher effect sizes at the beach on our measures of implicit anti-fat bias relative to other implicit attitude measures gathered at the same site and time (e.g., political attitudes: d = .8 and attitudes toward sexual orientation: d = .9; B. A. Nosek, personal communication, August 5, 2001).
implicitly associated or evaluated those categories in memory (e.g., flowers as good and insects as bad) than when they were mismatched.

To evaluate implicit anti-fat bias, we asked participants to classify stimuli while associating fat and thin people with positive and negative attributes. Because of negative social attitudes about weight, we expected stimuli to be classified more easily when the category pairings reflected negative associations toward obesity (e.g., fat people with bad) versus category pairings that reflected positive associations (e.g., fat people with good). Implicit associations to one target category were assessed relative to a participant’s associations to the other target category (associations with fat people were measured relative to implicit associations with thin people).

Since we were interested in both implicit attitudes and stereotypes toward overweight individuals, we had participants complete two different IAT tasks. To measure attitudes, or simple valence evaluation, participants completed an IAT in which they were asked to associate the target categories fat people and thin people with the attribute categories good and bad. To measure an implicit stereotype about overweight individuals, participants completed an IAT in which they were asked to associate the target categories fat people and thin people with the attribute categories motivated and lazy.

Each IAT task consisted of two pages (the order was counterbalanced across participants). On one page, the target and attribute categories were paired on either side of a column in a way expected to match negative implicit associations with overweight (e.g., fat people with bad) heading up one side of the column, and thin people with good heading up the other side of the column; see sample IAT page in Teachman & Brownell, 2001, p. 1531). On the other page, the target and attribute categories were paired to contradict expected negative associations with overweight (e.g., thin people was paired with bad on one side and fat people was paired with good on the other side). Participants were given 20 s to classify as many words as possible on one page; then they were given 20 s to classify words on the second page, on which the category pairings were switched. The variable of interest was the difference in the number of correctly classified items under the two different category pairings (incorrectly classified items were not counted in the difference score because errors can indicate misunderstanding of the instructions or an effort to increase speed at the cost of accuracy, thus distorting the actual implicit evaluation).

Participants were asked to work as quickly and accurately as possible, and they were told to try to avoid making mistakes (i.e., misclassifying a word) but to continue without stopping should this occur. Further, they were told not to skip items. Given the novelty of the task, we asked all participants initially to complete an unrelated practice IAT task to familiarize them with the procedure. Three items for each target and attribute category were selected based on participants’ expected ease of categorization and familiar usage. These stimuli were approximately matched for length, and the ease of categorization was evaluated during pretesting.2

Explicit bias. Participants completed an established measure of stereotypes about overweight individuals. The Fat Phobia Scale (FPS; Robinson, Bacon, & O’Reilly, 1993) is a 50-item, semantic differential scale in which participants rate their feelings about what “fat people are like” on a series of different opposing dimensions (e.g., smart vs. stupid). It parallels the IAT design in that it compares both ends of an attribute dimension (whether fat people are smart relative to stupid) but is dissimilar in that it does not evaluate attitudes toward fat persons relative to thin persons (asking only for judgments about overweight individuals).

Manipulation check. To evaluate the effectiveness of the prime manipulation about the causes of obesity, we asked participants, “What do you feel is the primary cause of obesity?” as the final question of the study. We then coded their open-ended answers to indicate whether they reported a factor that lay within or outside personal control (e.g., overeating vs. genetics) to determine whether participants remembered the news article they had read.

Procedure

Participants first completed a general demographics questionnaire and then received instructions on how to complete the IAT practice task. Participants completed one of four versions of the practice task (which was counterbalanced to control for order effects and placement of category labels on either the left or right column), and then they underwent the prime manipulation. Participants in the no-prime condition simply completed the IAT tasks, whereas participants in the genetics and behavior conditions carefully read the “recently published news article” before completing the two IAT tasks. The order of the implicit attitude and stereotype tasks was counterbalanced, and, within each IAT task, the order in which matched versus mismatched category pairings appeared was counterbalanced. Participants then completed the FPS and the manipulation check.

Results

Data were checked to confirm that all participants completed a minimum of five items on each IAT page as evidence that they understood and attended to the task (one lazy–motivated IAT score was deleted based on this criterion). Next, IAT pages with high error rates (i.e., ≥35% incorrectly classified items) were omitted because this may indicate distraction or lack of understanding. These corrections resulted in deleting 17 participants’ bad–good and/or lazy–motivated IAT scores from IAT analyses (their explicit data were retained). In addition, 1 participant’s IAT effect was more than 3 SDs above the mean. However, when data were analyzed both with and without this participant’s data, the pattern of results was the same, so we report the full sample here.

Manipulation Check

A logistic regression to determine whether prime condition would predict whether a factor that lay within or outside personal control was stated as the primary reason for obesity was significant: Wald’s $\chi^2(1, N = 133) = 4.86, p = .03, B = −.53, SE = .24$. Fifty-nine percent of the no-prime group reported an internal cause such as overeating (suggesting that this is the baseline response), whereas 81% of the behavior prime group indicated an internal cause, compared with only 65% of the genetics prime group. Follow-up t tests between the behavior and genetics prime groups (the two main groups of interest) indicated significant differences in the reporting of internal versus external causal factors for obesity, $t(88) = 3.53, p = .001, d = .79$, and similar differences were found between the behavior and no-prime groups, $t(86) = 4.55, p < .0001, d = .98$.

There was no significant difference between the genetics and no-prime groups, $t(89) = 1.15, p > 1.0$, which somewhat limits the conclusions that can be drawn from the genetics manipulation, but the nature of the study design required that the behavior and genetics research study manipulations be identical to make the test fair. Consequently, the overall difference across groups and the difference between the behavior and genetics prime groups were the critical factors in establishing the validity of the test. The genetics manipulation is congruent with previous research by

---

2 The materials used across the three studies are available from Bethany A. Teachman on request.
Chlouverakis (1975), who found that many people considered obesity to be due to overeating despite expert opinion about the role of genetic factors. These studies suggest that blaming attitudes toward obese persons are easier to exacerbate than to diminish.

**Bias Measures**

Because the IAT is a relative measure, evidence of an implicit effect can be interpreted as both pro-thin and anti-fat biases. IAT effects were calculated by creating difference scores where the number of items correctly classified in the mismatched condition (e.g., fat people + motivated) was subtracted from that in the matched condition (e.g., fat people + lazy). This resulted in a positive score for most participants but a negative score for individuals who classified more items when fat people was paired with positive attributes (only 4% of the sample indicated such a pro-fat/anti-thin bias). To better control for individual differences in positive attributes (only 4% of the sample indicated such a pro-fat/anti-thin bias), we inserted the difference score into the following algorithm: max/min – 1 * \( \sqrt{\max - \min} \), where maximum (max) and minimum (min), respectively, represent the category pairings where the highest versus the lowest number of items was correctly classified. This composite scoring of the IAT is based on simulations run by Nosek and Lane (1999).

To establish evidence of implicit bias, we conducted t tests to show that the IAT effects differed from zero. More items were correctly classified when fat people was paired with negative attributes, lazy: \( t(109) = 9.68, p < .0001, d = 1.85 \); bad: \( t(109) = 10.37, p < .0001, d = 1.99 \), than with positive attributes. The implicit attitude (bad–good) and stereotype (lazy–motivated) IAT measures were not significantly different from one another, \( t(93) = 1.11, p > .10, d = .23 \), and were moderately positively correlated (\( r = .33, p = .001 \)).

In contrast to the strong evidence for anti-fat/pro-thin implicit bias, the total score for the explicit FPS indicated a slight pro-fat bias. The average item score (\( M = 2.81, SD = .44 \)) was significantly different from 3 (the neutral point on the 5-point semantic differential scale), \( t(143) = 5.29, p < .0001, d = .88 \), and all six factors of the scale indicated a similar pattern.

The bad–good IAT was not significantly related to the explicit FPS (\( r = .03, p > .10 \)), but the lazy–motivated IAT showed a positive correlation with the scale (\( r = .29, p = .002 \)). Thus, the implicit and explicit stereotype measures were related, but the attitude measures were not. These variable relations are not surprising given the relative nature of the implicit measures (fat vs. thin people) compared with the explicit scale and the likelihood that explicit responses are more vulnerable to social desirability. Additionally, there were neither significant sex differences on the bias measures nor significant correlations with age or education. Interestingly, BMI for women (but not men) was negatively related to the lazy–motivated IAT (\( r = -.34, p = .008 \)), indicating that women with higher body mass tend to have weaker stereotypes of fat people as lazy. This finding suggests that an in-group bias among women may be present (e.g., higher weight persons show less negative beliefs about overweight people).

**Influence of Manipulating the Causes of Obesity on Anti-Fat Bias**

To increase reliability, we averaged the bad–good and lazy–motivated IATs, and then we used this total IAT and the total explicit scale as the dependent measures for planned contrasts comparing the three prime conditions. As expected, telling participants that obesity was primarily due to overeating and lack of exercise resulted in higher implicit bias, \( t(124) = -2.22, p = .03, d = .40 \), and a trend toward higher explicit bias, \( t(141) = -1.91, p = .06, d = .32 \), relative to the no-prime control group and marginally significantly higher implicit bias than the genetics prime condition, \( t(124) = -1.69, p = .10, d = .30 \). In contrast, telling participants that obesity was primarily due to genetics did not result in a lower bias than the no-prime condition for either the implicit or explicit measures (both \( t < .60, p > .10 \)). For easier visual inspection, Figure 1 indicates the z scores (with a meaningful zero point) and standard error bars for the implicit and explicit bias measures for each prime condition.

**Additional Explicit Stereotype Data**

To address the absence of data regarding explicit stereotypes about thin people, we asked a second sample of participants (\( N = \)
25) from the same beach location to complete the FPS, rating what thin people were like. Although this did not allow direct comparison of attitudes toward fat versus thin people (because the samples were independent), it did permit some comparison, given the parallel nature of the FPS and the Thin Phobia Scale (TPS) and the assumed similarity of the samples. As expected, participants rated thin people slightly more positively than negatively on the total scale (reflecting the 50 semantic differential adjective pairs), $t(24) = 3.94, p = .001, d = 1.61$, indicating that the average item score ($M = 2.69, SD = .40$) was significantly different from 3, the neutral point on the 5-point semantic differential scale. Importantly, average item scores did not differ across the FPS and the TPS, $t(167) = 1.12, p < .20$, suggesting no reported explicit bias toward fat people.

Study 1 provided strong support for the presence of anti-fat/pro-thin implicit attitudes and stereotypes but a relative absence of explicit anti-fat bias. The manipulation of information about the cause of obesity indicates that the bias can be made more negative by promoting personally controlled attributions of cause but cannot easily be made more positive, given the difficulty of convincing the general public that obesity is largely due to genetic factors.

**Study 2A**

In Study 2A, we used a new approach to bias reduction. Rather than providing statistical information about the causes of obesity, we tried a more affective manipulation in the hopes of changing feelings about overweight persons. The rationale for trying this new approach came in part from theoretical and empirical work demonstrating that evoking empathy to a member of a stigmatized group can reduce bias and also from our Study 1, which pointed to the difficulty of changing attitudes through an information approach.

Empathy is similar to sympathy but has a stronger component of relating to another person and taking his or her perspective. Batson et al. (1997) induced more positive attitudes toward stigmatized groups, such as AIDS victims, homeless people, and convicted murderers, by evoking empathy for members of these groups. Few studies have used empathy to reduce anti-fat stigma. However, an intervention by Wiese, Wilson, Jones, and Neises (1992) with medical students included (among other strategies) a video of an obese woman talking about discrimination she faced and role-playing exercises in which students took the perspective of an obese person. These components of the intervention likely induced empathy, and this study did find reduced bias against obese persons. On the basis of these suggestive findings, we explored whether empathy induction by means of written stories of weight discrimination could lead to lower anti-fat bias. We hypothesized that bias would be lower relative to controls because the discrimination story would evoke empathy for the overweight protagonist of the story, which would then generalize to the larger group of overweight people.

Participants completed explicit and implicit measures of anti-fat bias after reading story primes that either (a) evoked empathy toward an obese person, (b) evoked empathy toward a person in a wheelchair (an alternate stigmatized group), or (c) were neutral in valence (and were not expected to evoke empathy). This study was conducted in the laboratory to provide better control, and a computerized form of the IAT was used that allowed us to evaluate a range of implicit bias measures. In addition, in Study 2A we used a different approach to assess explicit stereotypes about fat people relative to thin people to make the implicit and explicit measures more comparable, thereby addressing this weakness from Study 1.

**Method**

**Participants**

Female participants were recruited from the Yale University community, either through introductory psychology classes or from signs posted around campus. Only women were used in this study because of the need to reduce the confound of mismatched gender of participant to gender depicted in the stimuli (i.e., stories and pictures). They completed the study in exchange for $7 or course credit. The sample ($N = 90$) had a mean age of 21 years ($SD = 3.87$; range = 18–44 years); the reported ethnicity was 56% Caucasian, 21% Asian/Pacific Islander, 10% Black, 8% Hispanic, and 5% other; and the mean BMI was 21.48 ($SD = 2.63$).

**Materials**

**Story primes.** Empathy was evoked by having participants read a first-person account of the troubles experienced by a member of a stigmatized group (an approach well established in previous research; e.g., Batson, Klein, Hignerberger, & Shaw, 1995; Batson et al., 1997) and then having them think about the feelings of that individual and complete a writing exercise to further elaborate their feelings. Depending on predetermined counterbalancing, participants were assigned to one of three conditions: reading about a person who is obese, reading about a person who is in a wheelchair, or reading a neutral story about a nonstigmatized person. The wheelchair condition was included to determine the effects of evoking empathy toward a nonstigmatized person (an approach well established in previous research; e.g., Batson et al., 1997) and then generalizing to influence anti-fat/pro-thin bias. In each condition, participants read two narratives that were three to four pages in length, followed by a short writing task. The setting of one story was always a student’s first day at college, and the other was a 15-year-old’s first day at sleep-away camp (order of the story primes was counterbalanced).

In the obesity-empathy condition, the story primes gave first-person accounts of an extremely overweight woman experiencing prejudiced jokes and social rejection because of her weight, with the protagonist describing her sadness and hurt feelings. The story primes for the wheelchair-empathy condition used identical wording, except that instead of being overweight, the main character was in a wheelchair. Finally, story primes in the neutral condition used the same college and camp settings and also gave first-person accounts, but these accounts were from the perspective of a nonstigmatized woman who described activities that were neutral in valence, without referencing emotions or feelings.

**Manipulation check.** To check the effectiveness of the empathy manipulation, we asked participants to complete a short questionnaire that rated their feelings toward the main character in the story on a 7-point Likert scale for each of six characteristics: sympathetic, warm, compassionate, softhearted, tender, and moved. This empathy index was developed by Batson and colleagues and has been used to measure empathy in a number of previous studies (e.g., Batson, 1991; Batson et al., 1995).

**Implicit bias.** The computerized version of the IAT is the more common form of the test and follows logic identical to that of the paper-and-pencil version already discussed. Participants were asked to classify stimuli

---

3 Demographic information was not gathered for this second sample; however, given the identical location and recruitment procedures, the samples were expected to be comparable.
into categories as quickly and accurately as possible to index the strength of implicit associations between two paired categories (relative to a different pairing). The dependent variable was the difference in average response latency across all classification trials for one category pairing versus the other pairing.

In the computer IAT, picture and word stimuli appeared in random order in the center of a PC screen from any of the four categories being associated (e.g., fat people, thin people, good, bad), while two of the category labels were always paired visibly near the top of one side of the screen (such as fat people paired with bad) and the other two category labels were paired on the other side of the screen (in this case, thin people paired with good). After participants categorized stimuli into their respective sides of the screen on the basis of the paired category labels, the category pairings were switched (i.e., switching the example above, fat people was paired with good and thin people was paired with bad). Again, participants were asked to categorize stimuli from the four categories into the new category pairings (see Greenwald et al., 1998, for a more detailed description). Participants completed an unrelated practice IAT task to initially familiarize themselves with the procedure and a series of 20 practice trials classifying the stimuli for each new category pairing before the 36 experimental classification trials were conducted. Stimuli included words to depict each category label, along with pictures of overweight and underweight women.

Participants completed two sets of IAT tasks, with the order of the IAT tasks and category pairings randomized within each task (there was a break between the two sets to reduce fatigue and to reinforce the empathy manipulation). In the first set of IAT tasks, the strength of associations between three pairs of category labels was evaluated. Two parallel attitude tasks were included, comparing associations toward fat people and thin people as good versus bad. In one case, word stimuli were used to depict all four categories (referred to as Fat/Bad Words IAT), and in the other case, pictures were used to depict fat and thin people (referred to as Fat/Bad Pictures IAT). The third task in this set looked at stereotypes of fat people and thin people (depicted using pictorial stimuli) as valuable versus worthless (depicted using word stimuli; referred to as Fat/Worthless IAT).

The second set of IAT tasks looked at another bad–good attitude measure but varied the category labels, using the terms overweight people and underweight people (referred to as Overweight/Bad IAT). In the final task, implicit evaluations of one’s own weight were determined on the basis of associations between the categories fat and thin with me and not me (referred to as Others/Fat IAT).

Explicit bias. Participants were asked to rate their feelings, on a 7-point semantic differential scale, from 1 (bad/negative) to 7 (good/positive), describing how they feel fat people and thin people are. Attitudes toward normal weight people were also evaluated to allow a different relative comparison to fat and thin people. Using analogous scales, participants were asked to rate how worthless to valuable they felt fat people and thin people were. In addition, participants used comparable semantic differential scales to rate how underweight or overweight they felt themselves to be, as well as how underweight or overweight they felt others to be on average. These explicit rating scales mirrored the relative associations assessed by the implicit association tests, so in each case a difference score between attitudes toward fat versus thin people, and the self versus others, was obtained.

Procedure

Participants were told that the purpose of the study was to investigate attitudes and feelings in novel situations and that they would be reading two true stories. Following informed consent, the participants then read the first story (either obesity empathy, wheelchair empathy, or neutral, depending on their condition) and completed the associated writing exercise and manipulation check. Participants then completed the first set of IAT tasks. To reinforce the empathy manipulation, participants then read the second story prime, with the associated writing exercise and manipulation check, followed by the second set of IAT tasks and a demographics questionnaire. The placement of the explicit bias measure was counterbalanced so that half of the participants completed it before the first set of IAT tasks and half completed it at the end of the study with the demographics questionnaire. Finally, participants were fully debriefed and thanked for their time.

Results

Prior to conducting the planned analyses, IAT response latencies less than 300 ms or greater than 3,000 ms were counted as errors and recoded as 300 or 3,000 ms, respectively. In addition, data were examined for error rates (i.e., percentage of stimuli classified incorrectly) on the critical IAT blocks. All participants’ total error rates were less than 25%, though 1 participant had one IAT task score removed because of an error rate on this task greater than 40%. Additionally, 1 participant’s IAT data were excluded because this individual’s average response time was over 1,400 ms on the IAT tasks. This slow responding suggests strategic, rather than implicit, responding. Further details on these data management procedures are provided in Greenwald et al. (1998). Finally, the split-half reliability of the IAT data was calculated across IAT tasks (average r = .67; range r = .52–.83), suggesting good psychometric properties relative to other reaction time measures.

Manipulation Check

To establish the effectiveness of the manipulation, we compared the total score on Batson’s empathy index (e.g., Batson, 1991) across the three prime conditions using an analysis of variance (ANOVA). As expected, results indicated a significant difference across conditions, $F(2, 86) = 24.54, p < .0001, f = .39$, and follow-up tests confirmed higher reported feelings of empathy toward the main character in the obesity condition, $t(56) = 6.27, p < .0001, d = 1.68$, and the wheelchair condition, $t(58) = 6.57, p < .0001, d = 1.73$, relative to the neutral story condition, and no difference between the obesity and wheelchair conditions, $t(58) = .48, p > .10$. (The effect size $f$ was described in Rosenthal and Rosnow, 1991, and is commonly used for ANOVAs to index the magnitude of an effect independent of sample size. As recommended by Cohen, 1988, a magnitude between .1 and .25 reflects a small effect, between .25 and .4 reflects a medium effect, and above .4 reflects a large effect.) On the 7-point Likert scale (with 7 indicating high empathy), the average item scores for the conditions were as follows: obesity ($M = 5.53, SD = .87$), wheelchair ($M = 5.64, SD = 0.93$), and neutral ($M = 4.30, SD = 0.61$).

4 Pictures were used in some tasks to explore the generalizability of the implicit anti-fat findings and to encourage participants to focus on overweight and underweight people rather than simply on overweight or underweight as a concept. All pictures were full-body, realistic computer-animated models, developed from the Landsend.com Web site, where they are used to model clothing. Models were created that varied in hair and skin color, but all were dressed in identical black jump suits. For each variation of hair and skin model, matching overweight and underweight counterparts were developed so that the only difference in appearance was body size.
Bias Measures

We used t tests to evaluate anti-fat/pro-thin bias on the IATs, on the basis of whether the relative response time was significantly different from zero. Main effects for all five IAT tasks confirmed our hypothesis of implicit bias: Fat/Bad Pictures IAT, t(88) = 2.45, p = .02, d = .52, Fat/Bad Words IAT, t(88) = 4.40, p < .0001, d = .94, Overweight/Bad IAT, t(87) = 3.38, p = .001, d = .72. Fat/Worthless, t(88) = 3.15, p = .002, d = .67, and Others/Fat IAT, t(88) = 6.55, p < .0001, d = 1.40.

We conducted t tests to determine whether the explicit bias difference scores were significantly different from zero. Despite the evidence of strong implicit anti-fat/pro-thin bias, there was no evidence of explicit fat–thin bias on the Bad/Good scale, t(88) = 1.43, p > .10, d = .30, on the Worthless/Valuable scale, t(87) = 1.25, p > .10, d = .27, or on the explicit evaluations of others as fat relative to the self, t(87) = .43, p > .10, d = .09. Interestingly, the only evidence of explicit bias emerged when attitudes toward fat people as good versus bad were compared relative to attitudes toward normal weight people, t(87) = 7.14, p < .0001, d = 1.53, with means indicating more positive evaluations of normal weight people (M = 5.01, SD = 1.21) than either thin people (M = 4.18, SD = 1.12) or fat people (M = 4.01, SD = 1.04), with 4 indicating the neutral point on the scale.

When collapsing across conditions and looking at the full sample, a pattern of small positive relationships between the various measures of implicit bias was found (though correlations ranged from r = .04 to r = .39, indicating wide variability), with the expected significant positive relations among the three implicit measures of attitude (Fat/Bad Words and Pictures IATs, r = .24, p = .03; Overweight/Bad IAT with the Fat/Bad Words IAT, r = .30, p = .005, and with the Fat/Bad Pictures IAT, r = .37, p = .0004). To look at the relationship between implicit and explicit anti-fat/pro-thin biases, we created composite measures of implicit and explicit biases to simplify analyses given the large number of variables and to increase power and reliability of the measures. The five IAT tasks were averaged to create the total implicit bias measure, and the three comparable explicit difference scores were averaged to create the total explicit bias measure (the scale comparing evaluations of fat people relative to normal weight people was excluded given the incongruent relative comparison to the other tasks). There was a moderate positive correlation between the total implicit and explicit bias measures (r = .36, p = .0004). Interestingly, this relationship was driven by the strong correlation between the implicit and self-reported biases in the obesity condition (r = .52, p = .004), as the implicit and explicit measures were not significantly related in the other conditions (wheelchair: r = .28, p > 1.0; neutral: r = .11, p > 1.0). BMI was strongly related to explicit ratings of the self as fat relative to others (r = .55, p < .0001) but not to implicit ratings (r = -.14, p > 1.0).

Influence of Manipulating Empathy on Anti-fat Bias

ANOVA conducted for the total bias measures (with prime condition as a three-level between-subjects variable) indicated no significant differences on either the total implicit bias measure, F(2, 86) = 1.36, p > .10, or the total explicit bias measure, F(2, 86) = .41, p > .10. However, when looking at the bias as a function of weight status, the empathy manipulation had been effective for overweight participants (based on the standard cutoff of a BMI over 25), with lower bias expressed when empathy was evoked. Specifically, a significant interaction was found across the prime conditions, with BMI as a two-level between-subjects variable (BMI > or < 25), on the total implicit bias measure, F(2, 86) = 6.63, p = .002, f = .40. This interaction is depicted in Figure 2, with the means and standard error bars for each weight group and prime condition indicated. Follow-up tests looking at each weight group separately indicated that there was no difference across empathy conditions for normal and underweight participants, but there was a difference for overweight participants, F(2, 7) = 4.05, p = .07, f = 1.08. The interaction with BMI for the total explicit bias measure was not significant, F(2, 86) = 1.03, p > .10, f = .16, though the pattern of results for the two BMI groups across empathy conditions was similar for the implicit and explicit bias measures.

Study 2B

Study 2B pulled data from a larger study of implicit anti-fat bias (see Teachman, Gapinski, & Brownell, 2002) to replicate the findings from Study 2A that hearing about discrimination against obese persons has a differential impact on implicit anti-fat/pro-thin bias depending on weight status. Given the small number of overweight women in Study 2A (typical of a college-age female population), a sample from the general population was used to include a larger proportion of overweight participants and a mixed-gender sample.

Method

Participants

Participants were recruited from the same location and in the same manner as in Study 1. This sample (N = 63) was approximately equal for gender (51% female), the mean age was 42 years (SD = 16.51, range = 18–80 years), and the sample was predominantly Caucasian (79%). BMI was 25.94 (SD = 6.25) for men and 25.64 (SD = 3.64) for women.

Materials and Procedure

Story prime. In this study, we used a novel approach to evoke empathy. Participants read a story of severe discrimination that was modified from a true news story (with identifying information changed for ethical reasons), which described an accomplished, obese young woman who was sent to “fat camp,” where she died after being verbally abused and forced to exercise excessively in the hot sun.

5 The views of self and others as fat were included with the attitude and stereotype measures given the expected consistency between views of one’s own identity and views of out-groups.

6 Interestingly, when the wheelchair and neutral conditions were compared, there was a marginally significant trend indicating lower implicit anti-fat/pro-thin bias when empathy toward a disabled person was evoked, t(58) = 1.69, p < .10, d = .75, suggesting that evoking empathy toward one marginalized group effectively minimized implicit bias to a different marginalized group.
IMPLICIT ANTI-FAT BIAS

Figure 2. Interaction between body mass index (BMI) and empathy prime conditions for implicit anti-fat/pro-thin bias in Studies 2A and 2B. A: Interaction between BMI and empathy condition (obesity–wheelchair–neutral) for implicit anti-fat/pro-thin bias. Zero point indicates no difference in response time across Implicit Association Test (IAT) conditions, whereas positive numbers reflect an anti-fat/pro-thin bias and negative numbers reflect an anti-thin/pro-fat bias. B: Interaction between BMI and empathy condition (discrimination story–no prime) for implicit anti-fat/pro-thin bias. Zero point indicates no difference in number of correctly classified items across IAT conditions, whereas positive numbers reflect an anti-fat/pro-thin bias and negative numbers would reflect an anti-thin/pro-fat bias.

Implicit bias. Study 2B also extended the evidence of implicit anti-fat bias to another stereotype domain. The lazy–motivated IAT task was included along with an IAT that paired the categories fat people and thin people with the categories stupid and smart (using stimuli like dumb and intelligent). This task was selected because explicit stereotypes about obese persons’ incompetence have been documented, but this domain has no grounding in health outcomes. The design, administration, and method of scoring of these IAT tasks were identical to those in Study 1 (further details about this sample, materials used, or procedures are available from Bethany A. Teachman on request).

Results

Once again, t test results indicated strong implicit anti-fat/pro-thin associations with lazy relative to motivated, \( t(49) = 6.39, p < .0001, d = 1.83 \), and similar support was found for the smart–stupid stereotype, \( t(49) = 5.27, p < .0001, d = 1.51 \). To look at the influence of the discrimination story manipulation (half of the participants read the story of discrimination, half did not), we conducted t tests for the implicit and explicit measures to determine whether the no-prime group was significantly different from the discrimination prime group. Similar to Study 2A, when the sample was analyzed as a whole, there was no evidence of a reduction in either implicit or explicit biases following the discrimination prime (all ts < 1.5, ps > 1.0). However, replicating the findings from Study 2A, the manipulation was effective for overweight participants (based on the standard cutoff of a BMI over 25). A significant interaction was found indicating that overweight participants showed less implicit bias if they had read the discrimination prime, but if participants’ BMIs were below 25, the manipulation had no effect: combined IAT measure, \( F(1, 40) = 6.09, p = .02, f = .40 \) (see Figure 2).

General Discussion

Across these studies, strong evidence of implicit anti-fat/pro-thin attitudes was found, regardless of the category labels used to portray overweight and underweight people, the use of pictorial or word stimuli, or the form of the implicit association test used (paper–pencil or computer-based). Implicit stereotypes of fat people (relative to thin people) as lazy versus motivated, stupid versus smart, and worthless versus valuable were demonstrated, further establishing the pervasiveness of the implicit bias. These stereotypes, which include attributes clearly unrelated to health, were exhibited in both the general population and among college students. The consistency of the implicit bias across the different samples and methodologies stood in contrast to the lack of explicit anti-fat/pro-thin bias. Interestingly, the only evidence of self-reported anti-fat attitudes emerged when evaluations of fat people as good versus bad were compared relative to evaluations of normal weight people. These studies support the pervasiveness of implicit bias against overweight people and the unique demonstrations of bias that can be learned from implicit measures. Given the relative nature of the measures used, the results can be interpreted as both anti-fat and pro-thin biases. On the basis of popular messages in the media that promote fat jokes and the idolization of thin models, it seems likely that both pressures are active in Western cultures.

Across the studies, a surprising pattern of results emerged from the different approaches to modifying anti-fat/pro-thin biases. Although providing information to participants that obesity is predominantly caused by behavioral factors, such as overeating and lack of exercise, led to higher bias (compared with the other groups), giving comparable information that obesity is mainly due to genetics did not result in lower implicit or explicit biases. Our attempts to evoke empathy through stories of discrimination against an overweight young woman did not produce lower bias across the whole sample, further demonstrating the durability of the anti-fat judgments. However, across two studies, the manipulation did lead to lower implicit bias for overweight participants. This may be important given that self-blame and internalizing of negative social messages are common in overweight people.
In general, variable relations were found between implicit and explicit bias measures, though across two studies in this article (and in Teachman & Brownell, 2001), there was a tendency to find small positive correlations between stereotypes of implicit and explicit beliefs about fat people being relatively lazy. It may be more socially acceptable to acknowledge beliefs that overweight people are lazy (given cultural values that obese people are responsible for their condition) than it is to directly evaluate an overweight person as bad. It is interesting to note that in Study 2A, there was a strong relationship between implicit and explicit biases when people were primed with empathy toward an obese person, compared with the other prime conditions. Reading about obesity may lead participants to think about their own views of obese persons, and this elaborative process may encourage more consistency across implicit and self-report measures (support for this link between elaboration and attitude consistency comes from Nosek, 2002).

**Perceived Controllability as a Bias Reduction Strategy**

The finding in Study 1 that giving participants information about the behavioral causes of obesity led to higher bias (than the other groups), but giving participants information about genetic contributions did not lead to lower bias is consistent with Bell and Morgan (2000). However, the finding differs from other studies that found an obese person was viewed less negatively when participants were informed that the obesity was due to a medical condition, such as a thyroid problem (e.g., DeJong, 1980, 1993; Rodin, Price, Sanchez, & McElligot, 1989; Weiner, Perry, & Magnusson, 1988). One possible explanation is that our research study manipulation was more credible when the behavioral causes, rather than the genetic causes, were cited, even though the manipulation check indicated overall differences across groups. The absence of a difference between the no-prime control group and the genetics prime group on the manipulation check supports this idea, and Chlouverakis (1975) has also found that many people consider obesity to be caused by overeating, despite expert opinion about the role of genetic factors. Consequently, participants' strong prior beliefs may have led them to discount the research study. Future work that uses a more extensive manipulation emphasizing the role of genetics and biology in weight control could address this issue.

An alternative explanation is that the studies cited above noted a very specific disease that caused the obesity rather than just referring to more general biological or genetic factors. This greater causal specificity may have been more credible. Further, our study asked people to make judgments toward obese persons as a group rather than to a particular obese individual (as in many of the previous studies). It may be easier to accept that a given person is blameless for his or her condition but more threatening to one's worldview to believe that obesity in general is blameless.

The fact that negative attitudes toward obese persons could be so easily exacerbated suggests that fat jokes, teasing, derogatory portrayals of obese persons in the media, and other social phenomena might intensify bias. Greenberg, Eastin, Hofshire, Lachlin, and Brownell (2001), for instance, found that overweight characters on entertainment television were imbued with many more negative characteristics than were other characters and were less likely to be shown in positive interactions (e.g., romantic relationships). Given the results of Study 1, one might argue that negative portrayals, particularly those that support blaming attributions, make anti-fat attitudes worse.

**Empathy as a Bias Reduction Strategy**

The finding across two studies, that evoking empathy did not result in lower bias in general (compared with a control group) but did diminish bias for overweight participants, is intriguing. These results suggest that it may be premature to propose that "there is simply no evidence to suggest that fat people display an in-group bias when it comes to expressing prejudice about fat people" (Crandall, 1994, p. 890). Rather, it may be that although both overweight and normal weight people endorse approximately the same amounts of anti-fat dislike in general (Crandall, 1994), there is an in-group bias such that overweight persons are more responsive to contextual factors that can mitigate anti-fat views. Reminding overweight persons about anti-fat discrimination may be one such factor that would promote in-group support, perhaps because obese people tend to view their status in the group as temporary (Quinn & Crocker, 1998). Thus, without the reminder of the shared prejudice obese persons have experienced, overweight people may be less inclined to develop the group consciousness that could lead to an in-group bias.

It is also possible that the empathy stories did not produce lower bias for the full sample because in addition to evoking empathy, the stories also reminded participants of the negative stereotypes associated with obesity. There is evidence that the strength of implicit bias toward a stigmatized out-group is increased after viewing images of an out-group member (relative to viewing the face of an in-group member; Dovidio et al., 1997; Fazio, Jackson, Dunton, & Williams, 1995). In the present study, participants read stories describing negative evaluations of an obese person. It is possible that this strengthened previously established anti-fat bias, which countered the empathic feelings produced by the manipulation. Future research might investigate the effects of priming participants with positive attitudes and stereotypes toward obese persons so that new pro-fat associations can be formed, allowing for reconstruction of the current negative biases. This approach has shown promise in the area of race bias, in which priming participants with positive exemplars of African Americans led to a reduction in bias (Dasgupta & Greenwald, 2001).

The current line of studies firmly establishes the presence of implicit anti-fat/pro-thin biases, but many open questions remain about what factors are needed to effectively modify these biases. These studies are limited by the cross-sectional design and brevity of the manipulations used relative to the pervasive anti-fat messages in our culture, making it difficult to determine whether null results occurred because of the weakness of the prime or because the intervention does not reduce bias. Also, it is possible that the location of Studies 1 and 2B (a beach) is a unique environment that may actually strengthen anti-fat attitudes, making it more difficult for our primes to be effective. The revealing clothing typically worn at the beach may intensify body focus, causing more negative feelings about fat in general. However, the similar findings for Study 2A, which was conducted in the laboratory, suggest the setting played a limited role in the results. A further limitation of
these studies is that our samples are relatively homogenous (all data were collected in Connecticut and participants were predominantly Caucasian), potentially limiting the generalizability of our results. It would be interesting to examine implicit anti-fat bias among people of different ethnic and racial backgrounds, given findings of different levels of explicit anti-fat bias across subcultures.

The strength of the implicit biases observed in these studies is alarming, as is their resistance to modification. The effect sizes found in this study are at least as strong as, and often stronger than, the effects for implicit biases found against other stigmatized groups (e.g., race bias: $d = .71$ and .88; age bias: $d = 1.42$ and .99; and gender stereotypes: $d = .72$; Nosek, Banaji, & Greenwald, 2002). Further, the relative absence of self-reported anti-fat views indicates to us the importance of alternative methods of assessing the anti-fat biases that are clearly so insidious and pervasive in our culture. These studies establish implicit anti-fat stereotypes across a range of domains that have not previously been explored, demonstrate the implicit bias in the general population for the first time, and represent some of the first attempts to alter implicit anti-fat bias. Although the strength of the implicit biases in these studies was disappointing, we remain hopeful that these implicit anti-fat biases can be effectively reduced. Change in this domain may be even more challenging than changing other stereotyped beliefs, given the relative acceptability of anti-fat attitudes (e.g., Kilbourne, 1994), meaning that not everyone is motivated to change these beliefs. This may help us understand why overweight people in our studies showed reduced implicit bias given their likely motivation to control negative beliefs against their own in-group. This finding is hopeful because of the psychological and health consequences that follow the experience of discrimination from others and self-blame. The challenge is set to continue exposing implicit biases and to identify ways to counter them to ultimately promote equitable treatment of our increasingly overweight population.

References


