The Nature and Implications of Implicit Weight Bias

Kaaren Watts* a,b and Jacquelyn Cranney c

aPsychosocial Research Group, Prince of Wales Hospital, Level 3, Dickinson Building, Randwick NSW 2031, Sydney, Australia; bPrince of Wales Clinical School, University of New South Wales, Sydney NSW 2052, Australia; cSchool of Psychology, University of New South Wales, Sydney NSW 2052, Australia

Abstract: Although overweight is undesirable from both public and personal health perspectives, the focus of this paper is on exploring the nature of fat prejudice within a social cognition framework. Fat prejudice refers to the tendency to form judgments about people on the basis of excessive body weight. Body size has been described as one of the few personal attributes considered an acceptable target of prejudice, despite the fact that targets of fat prejudice experience significant psychological distress. Fat prejudice is likely to become an increasingly common psychosocial problem in light of the obesity epidemic that is currently affecting many countries. The current paper reviews findings from nineteen experimental studies of implicit anti-fat attitudes; these studies have used either the implicit association test or the affective priming task. The empirical data highlight that implicit anti-fat attitudes are widely held and relatively universal. Robust implicit anti-fat bias is evident among many groups including university students, members of the general public, health professionals, and among those who are themselves overweight or obese. The current data suggest that, similar to findings with other attitudinal objects, the relationship between implicit and explicit measures of anti-fat attitudes is complex. The possibility of changing implicit anti-fat attitudes, either by modifying the underlying associative structures or by altering the pattern of activation, is discussed. Avenues for future research are offered, keeping in mind the challenge of formulating appropriate public health messages whilst also challenging weight bias, and promoting acceptance of diversity in body size.

Keywords: Implicit attitudes, weight bias, attitude change, implicit association test, affective priming.

THE NATURE AND IMPLICATIONS OF IMPLICIT WEIGHT BIAS

Weight bias, also known as fat prejudice or anti-fat bias, refers to the tendency to form judgments about people on the basis of excessive body weight [1]. Body size has been described as one of the few personal attributes that many proponents still see as an acceptable target of prejudice [2]. This is disturbing because it is well documented that the targets of fat prejudice experience significant psychological distress [3]. The widespread stigmatisation of overweight, and the discrimination of obese individuals in many domains of living including employment [4] and health care, highlights the seriousness of either implied or overtly expressed negative attitudes toward overweight or obese individuals [5]. Although it is strongly acknowledged that excessive body weight is undesirable from both public and personal health perspectives, the focus of this paper is on exploring the nature of fat prejudice within a social cognition framework.

Fat prejudice is both similar and dissimilar to racial prejudice. First, like racial prejudice, anti-fat attitudes are widely held [6], and these impact on overweight individuals in a negative manner (e.g., [4]). Unlike racial prejudice, however, in which the targets often comprise a minority ethnic group, overweight and obese individuals comprise a substantial and increasing proportion of the global population. The World Health Organisation estimates that currently one billion people are overweight, and of these, 300 million people are classified as obese [7]. National survey data collected in the United States shows that in 2003-2004, 34% of adults aged 20 to 74 years were obese and 17% to 19% of children (aged 6-11 years) and adolescents (12-19 years) were overweight [8]. Since 1960, when the first National Health and Nutrition Examination Survey was implemented, the prevalence of obesity among American adults (20-74 years) has more than doubled [9]. Similar upward trends have been observed in the Asia-Pacific region, including Australia [10]. For example, in 2004 to 2005, 54% of all Australian adults aged 18 years and over were either overweight or obese, representing a substantial increase from 45% of adults in 1995. The rate of overweight adults has increased from 33% in 1995 to 36% in 2004-2005, while the rate of obesity in adults has increased from 13% to 18% over the same period [11]. These figures highlight that the potential targets of anti-fat attitudes represent an increasing proportion of the world’s population. Hence, fat prejudice is likely to become an increasingly common psychosocial problem in light of the obesity epidemic that is currently affecting most countries [12]. Alternatively, it could be argued that as the number of overweight and obese individuals increase, obesity will be ‘normalised’ and thus be seen as more acceptable. Latner and Stunkard’s [13] finding that attitudes towards obese people were worse in 2001 than they were in 1961, however, mitigates against that argument.

In a series of early studies, Crandall [14] highlighted the similarities and differences between fat prejudice and sym-
bolic racism. Symbolic racism (also known as “modern racism”) is distinguished from traditional racism in that the former is indirect and subtle whereas the latter is direct and overtly expressed [15]. In addition, symbolic racism tends to be rooted in abstract moral beliefs (e.g., belief in a just world) and conservative values reflecting concepts such as discipline, self-control, and self-reliance [16]. Crandall [14] developed a questionnaire consisting of three scales designed to assess beliefs about the controllability of weight (Willpower), evaluation of overweight people (Dislike), and one’s concerns about weight and the self-relevance of overweight (Fear of Fat). As expected, aspects of anti-fat attitudes, specifically the willpower and dislike components, correlated positively with modern racism and with a set of conservative values characterised by a tendency to assume that a person’s current predicament or circumstances are controllable. The primary difference between anti-fat attitudes and modern racism was that there appeared to be less normative pressure to suppress anti-fat attitudes than there was to inhibit negative racial attitudes. A further difference was that unlike the tendency for some racial groups to make comparisons with other social groups along dimensions that favour the members of their own group, thus enhancing self-esteem (called in-group bias; [17]), there was no evidence for a corresponding in-group favoritism among overweight individuals [14]. This absence of in-group favoritism among overweight or obese individuals has been demonstrated in other studies [18].

An important factor that may underlie the pervasiveness of fat prejudice is the widespread belief that, unlike skin colour, body weight and shape is malleable and controllable [19]. The myth that one’s body weight and shape can be readily modified is also widely promoted by the mass media [20]. This is in spite of biological research indicating that genetic and physiological factors are influential in determining body weight [21]. A recent study testing an attribution-value model of prejudice found that the simultaneous presence of greater negative cultural evaluation of fatness and attributions of responsibility for one’s weight predicted anti-fat prejudice, particularly in countries rating highly on a measure of individualism, including the United States and Australia [22]. Hence, overweight individuals may not only be held personally responsible for their body shape, but they may experience additional negative pressure through the sociocultural perception that body weight is easily controlled and modified.

**IMPLIED ANTI-FAT ATTITUDES**

In the last decade, attention has turned toward examination of implicit attitudes in the body image domain, particularly implicit attitudes toward body weight and shape, and how these relate to explicit attitudes. Implicit and explicit attitudes are best understood in terms of the mental processes that underlie them. Implicit attitudes refer to “automatic affective reactions” that occur in the instance that particular memorial associations are activated automatically upon encountering a relevant stimulus ([23], p. 693). Because implicit attitudes are automatic, they do not require conscious cognitive effort or an intention to evaluate in order to be brought to mind [24]. The most important distinction between implicit and explicit attitudes is that the former can be activated irrespective of whether the individual considers the evaluations to be true [23]. Hence, automatically activated evaluations are not necessarily congruent with one’s personally endorsed beliefs [25]. In contrast, explicit attitudes are evaluative judgments that are dependent upon effortful processes involving reflection and conscious assessment of the accuracy or inaccuracy of the evaluations of an object. Explicit attitudes, therefore, are conceptualised as personal beliefs that are endorsed as true [23].

Given the intentional nature of explicit attitudes, why do implicit attitudes matter, particularly if implicit attitudes are automatically activated evaluations that can be overridden by effortful processing? First, we often rely on automatically activated implicit attitudes toward an object as the “default mode” for our explicit evaluative judgments. In other words, it is argued that a negative automatic evaluation of an object or person will often be used as the basis for judging whether a proposition such as “I dislike X” is valid or not [23, p. 694]. Thus, implicit attitudes are important precursors and shapers of subsequent information processing, including explicit attitudes. Second, in line with the notion that attitudes are functional in that they serve to direct attention as well as approach or avoidance behaviours [26], automatically activated attitudes have been shown to predict behaviour, particularly nonverbal, indirect behaviours that are beyond conscious control [15, 27]. Importantly, such nonverbal behaviours are a key factor in effective communication including the formation of positive first impressions and in a person’s perception of rapport in a dyadic interaction [28]. Third, whilst implicit attitudes have been previously defined as relatively stable memorial representations that are resistant to change [24], there is accumulating evidence that automatic affective reactions are malleable, as they can be influenced by the context in which a stimulus is encountered [23], and patterns of activation underlying implicit attitudes can be modified by exposure to counter-exemplars of a negative automatic evaluation [29, 30]. The primary objectives of the current paper are to provide a systematic review of the evidence from experimental studies conducted over the past decade of implicit anti-fat attitudes and stereotypes, to discuss the findings in light of the meaning and importance of these attitudes, particularly as they relate (or not) to explicit attitudes, and finally to suggest avenues for future research.

**INDIRECT MEASURES OF IMPLICIT ANTI-FAT ATTITUDES**

Because implicit attitudes are automatically brought to mind upon exposure to a relevant stimulus or cue, they are minimally influenced by social desirability or demand characteristics [31, 32]. Accordingly, implicit attitudes are best measured by indirect tasks that do not alert the participant to the construct being examined [33]. In contrast, explicit attitudes are consciously mediated evaluative judgments. Hence, explicit attitudes are typically assessed using direct, self-report measures. Two indirect measures that have been used to examine implicit attitudes toward body weight and shape include the implicit association test and the affective priming task.
Implicit Association Test

The implicit association test (IAT; [34]) has dominated research on implicit attitudes in the social cognitive literature [35], partly because of its capacity to produce statistically robust effects, and its ready adaptation to a number of different psychological phenomena within and beyond social psychology [36]. The IAT was designed to provide a measure of the relative strength of association between attributes (e.g., “pleasant”) and concept categories (e.g., “Blacks”). In the two critical blocks of the “race” IAT, participants are required to categorise exemplars of the concepts “Black” and “White” (e.g., faces or first names) and exemplars of the evaluative attributes “pleasant” and “unpleasant” (e.g., “gift”, “disaster”) according to combined concept and attribute categories (e.g., Black/pleasant and White/unpleasant) using one response key for each combination (e.g., left key for “Black” or “pleasant” and right key for “White” or “unpleasant”). Importantly, if the concept categories “Black” and “White” are differentially associated with the attributes “pleasant” and “unpleasant”, then the participant should find one of the combined tasks to be easier than the other [34]. For example, a Caucasian participant may more readily associate “White” with “pleasant” than “Black” with “pleasant”.

The IAT effect, therefore, comprises the difference in mean response latencies between attitudinally congruent and incongruent trials, and in turn, provides an index of the relative strength of association between the pairs of concepts and attributes.

The authors acknowledge that despite the popularity of the IAT and the proliferation of studies that have used it [35], there is an ongoing debate about whether the IAT provides an uncontaminated measure of implicit attitudes [30, 37-39]. A comprehensive review of the IAT is outside the scope of this paper. Proponents, however, continue to argue against such criticisms [40], and there is accumulating evidence for the predictive validity of the IAT [41].

Affective Priming Task

The affective priming task (APT; [42]) is an alternative implicit attitudinal measure which has recently been applied in the body image domain [43]. The APT was designed to measure automatic activation of affect toward an attitude object (the “prime”). The APT provides a powerful indirect test of automatically activated (implicit) attitudes because (a) participants respond to the target words, not the primes, (b) exposure to the priming stimuli is brief, and (c) the amount of processing time is manipulated by varying the interval between the onset of the prime and the target (called the stimulus onset asynchrony, SOA; [42, 44]).

In the standard APT, two stimuli are presented in quick succession, a prime followed by a target. The participant is not required to respond to the prime, which is simply briefly displayed and then replaced by the target word. The participant decides as quickly as possible whether the target word is “good” or “bad” by pressing one of two response keys. The dependent variable is response latency to the target word. Two key features of the prime and target pairs are manipulated to test automatic evaluation. First, the emotional match (valence congruence) of each prime and target pair is varied such that half of the pairs have the same valence (e.g., positive prime, positive target), and the remaining pairs have different valence (e.g., negative prime, positive target). These comprise the “congruent” and “incongruent” conditions, respectively. The key assumption is that brief presentation of a prime automatically activates an evaluation, as well as meaning, because semantic and affective concepts are interconnected in memory in an associative network [45]. For example, when a negative prime is presented (e.g., fat), negative affect is likely to be automatically brought to mind. Hence, participants are typically faster to judge the valence of a negative target word (e.g., awful) when it is preceded by a negative prime (e.g., fat) than when the same target word is preceded by a positive prime (e.g., slimness). This is the congruence effect; the situation whereby when the prime and target word have the same valence (congruent), participants are much faster to evaluate the target word than when the prime and target have dissimilar valence (incongruent). Second, the amount of processing time that is available, and the level of activation of negative or positive evaluation that is operating when the target word is presented, is manipulated by varying the SOA. Half of the trials have a short delay (e.g., 300 ms) and the other half have a long delay (e.g., 1000 ms). At the short SOA, participants do not have time to deliberately retrieve their attitudes [44], and fast automatic processes predominate [42, 46]. At the long SOA, it is assumed that implicit attitudes will not be evident either because (a) automatic reactions decay rapidly such that they do not impact upon response latencies to the target words or (b) that participants have sufficient time to consciously suppress their automatic affective reactions. Typically, affectively congruent prime-target pairs (e.g., “fat” – “awful”) lead to shorter response latencies to the target word than do affectively incongruent prime-target pairs (e.g., “slimness” – “awful”) at the short SOA, but not at the long SOA [42, 47, 48]. This is what we term the “classic criterion” for automatic evaluation (as indexed by the interaction between SOA and valence congruence).

More recently, however, Fazio [49] has shifted his original position (e.g., [42]) and argued that the presence of the congruence effect at the short SOA is all that is needed to demonstrate automatic evaluation. This is what we term the “parsimonious criterion” for automatic evaluation. That is, it is not necessary to demonstrate the interaction as is argued with the classic criterion. Moreover, because the stigmatisation of overweight may be viewed by proponents as a socially sanctioned form of prejudice [50], the classic criterion may not hold for body-related primes (e.g., primes representative of the concepts “fat” and “thin”). This is because at the long SOA, despite having sufficient time to initiate controlled processing, individuals may not be sufficiently motivated to suppress or to challenge their automatically activated anti-fat bias. Automatic activation has been demonstrated according to the parsimonious criterion by several researchers [46, 51, 52] and in our own research with both body-related words and images ([53], see Fig. (1); [43]). A variation of the affective priming task involves the participant making a lexical judgment in response to the target word rather than a “good” or “bad” evaluation.

METHOD

A literature search was conducted within the electronic databases of MEDLINE and PSYCINFO using the following
keywords individually and in combination: implicit, automatic, antifat, anti-fat, fat, weight, attitudes, prejudice, stereotypes, and bias. The search was limited to articles published in English during the period 1998 to 2008. The reference lists of two major peer-reviewed journals within the field, Body Image: An International Journal of Research and the International Journal of Eating Disorders, were also examined. Eighteen relevant articles were identified and are included in the current review. One further relevant data set by the authors (currently under review) was also included, giving a total of 19 studies.

**REVIEW OF EMPIRICAL STUDIES**

**Studies Using the IAT**

The IAT has been utilised in numerous studies to examine weight bias including anti-fat attitudes and anti-fat stereotypes. Stereotypical anti-fat associations refer to the pairing of negative traits (e.g., lazy, stupid) with overweight. To examine anti-fat attitudes, participants are usually required to categorise words representing “fatness” (e.g., chunky, overweight) and “thinness” (slender, slim), according to either congruent category labels (“fat/bad” and “thin/good”) or incongruent category labels (“fat/good” and “thin/bad”). Several studies have also utilised images representative of “fat people” and “thin people” as the target stimuli (e.g., [27, 54-57]). Implicit anti-fat stereotypes are typically tested by having participants categorise words representative of “fat people” and “thin people” according to either stereotype-consistent category labels (“fat/lazy” and “thin/motivated”) or stereotype-inconsistent category labels (“fat/motivated”) and (“thin/lazy”). A summary of the findings of studies that have utilized the IAT to test implicit weight bias is presented in Table 1 (note that the studies are numbered according to the order in which they are first presented in the text).

In all the IAT studies reviewed, individuals demonstrated robust implicit anti-fat attitudes such that they were faster to respond to congruent trials (“fat/bad” and “thin/good”) than to incongruent trials (“fat/good” and “thin/bad”). Implicit anti-fat stereotypes were also evident in all studies have tested this phenomenon. Specifically, individuals are faster to categorise words (and images) representative of fat people and thin people according to negative traits (e.g., lazy, slow) and positive traits (e.g., motivated, energetic) respectively, compared to the reverse pattern (“fat” with a positive trait and “thin” with a negative trait). In all of the studies that examined both implicit anti-fat attitudes and stereotypes, the implicit measures correlated to a moderate degree (ranging from $r = .33$ to $r = .52$, [6, 18, 58, 59]). This highlights that implicit anti-fat attitudes and anti-fat stereotypes are related but not identical constructs, and this distinction warrants more attention in future research, particularly in relation to the predictive power of each of these constructs for discriminatory behaviours. The findings for the evaluative IAT (good/bad) and the stereotype IAT (lazy/motivated) suggest that individuals upon briefly encountering exemplars of “overweight” and “thin” automatically associate these stimuli with negative and positive attributes (or traits), respectively.

Implicit anti-fat bias is a widespread and robust phenomenon [6, 61]. Anti-fat attitudes are evident among the general population [6], among fitness and health professionals who work with obese clients [61, 62], and among overweight and obese individuals [63, 64]. The latter finding confirms that overweight, as well as individuals of average weight, hold negative implicit attitudes toward body size. As stated earlier, overweight persons may not have a preference for other overweight individuals and may even favour members of the out-group (i.e., slim individuals). For example, Rudman et al. [63] used the IAT to study the implicit attitudes of four social groups including Jewish Americans,
Table 1. Summary of Experimental Studies Examining Implicit Weight Bias Using the Implicit Association Test

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>[54]</td>
<td>Ahern et al. (2008) United Kingdom</td>
<td>95 female university students</td>
<td>Explicit</td>
<td>Participants completed measures of thin internalisation, body dissatisfaction, and dietary restraint.</td>
<td>On average, participants demonstrated an implicit negative bias toward underweight images and a positive bias toward average weight images.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit</td>
<td>Target categories: images of underweight females and images of average weight females. Attribute categories (words): positive, negative.</td>
<td></td>
</tr>
<tr>
<td>[55]</td>
<td>Ahern &amp; Hetherington (2006) United Kingdom</td>
<td>86 female university and high school students Age range: 16 to 25</td>
<td>Explicit</td>
<td>Participants completed measures of thin internalisation, body dissatisfaction, and dietary restraint.</td>
<td>Participants exhibited a strong implicit negative bias toward fat images and a positive bias toward thin images.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit</td>
<td>Target categories: images of women representing the categories “fat” and “thin”. Attribute categories (words): positive, negative.</td>
<td></td>
</tr>
<tr>
<td>[66]</td>
<td>Brochu &amp; Morrison (2007) Canada</td>
<td>76 university students (61 females, 15 males) Mean age = 20.11 years 89% Caucasian</td>
<td>Explicit</td>
<td>Attribute ratings of and behavioural intentions toward average weight or overweight targets. Anti-fat Attitudes Scale (AFAS; Morrison &amp; O’Connor, 1999).</td>
<td>Participants demonstrated both explicit and implicit anti-fat prejudice; males had more explicit negativity toward overweight targets. Implicit anti-fat attitudes did not predict behavioural intentions, except toward overweight male targets.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit</td>
<td>Target categories: average-weight and overweight persons (images). Attribute categories: positive and negative attributes (words).</td>
<td></td>
</tr>
<tr>
<td>[58]</td>
<td>Chambliss et al. (2004) U.S.A.</td>
<td>246 university students majoring in exercise science Mean age = 23.2 years 77% Caucasian</td>
<td>Explicit</td>
<td>Ratings of beliefs about fat people and thin people (semantic differential scale). Anti-fat Attitudes Test (AFAT; Crandall, 1994).</td>
<td>Strong implicit anti-fat bias was evident; being Caucasian, female, and having a lower BMI correlated with higher bias scores on the good/bad IAT. Moderate explicit anti-fat attitudes were evident especially weight blame, and the “lazy” stereotype. Stronger explicit anti-fat attitudes were associated with beliefs about greater personal control of obesity, being an undergraduate male, and not having obese friends.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit (Two IATs)</td>
<td>Target categories (words): fat people, thin people (IAT 1/2). Attribute categories (words): bad, good (IAT 1); lazy, motivated (IAT 2).</td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Contd….

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Participants completed components of the Eating Disorders Questionnaire (Mitchell, Hatsukami, Eckert, &amp; Pyle, 1985) and a single item measuring weight attitudes “How much do you believe being fat is bad?” (1 = completely disagree; 5 = completely agree). Implicit (Five IATs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 1 (weight attitude IAT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: light, heavy Attribute categories: good, bad.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 2 (Weight identity IAT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: self, other Attribute categories: light, heavy.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 3 (Gender attitude IAT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: female, male Attribute categories: good, bad.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 4 (Gender identity IAT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: self, other Attribute categories: female, male.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 5 (Self attitude IAT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: self, other Attribute categories: good, bad.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Both females and males demonstrated negative implicit attitudes toward overweight as indexed by responses on the weight attitude IAT. Participants also expressed explicit negative attitudes toward overweight. There were no gender differences in implicit or explicit weight attitudes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[56]</td>
<td>McConnell et al. (2008) U.S.A.</td>
<td>Study 1 133 university students Study 4 47 university students</td>
<td>Explicit</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Participants were either presented with an image of an overweight white male “Bob” or they saw an image of the same white male that had been morphed such that he did not appear overweight. (Study 1) Each image was accompanied by either positive or negative feedback (manipulated between subjects) about Bob’s behavioural characteristics. (Study 4) Each image was accompanied by either positive or neutral (ambiguous) verbal behavioural statements about Bob (manipulated between subjects).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: (Study 1) images of “Bob” and other white males (“not Bob”) who were either overweight or not overweight. (Study 2) words (Bob or not-Bob names e.g., “Ben”). Attribute categories (words): positive and negative adjectives.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In both studies, participants exposed to an image of an overweight “Bob” in the initial trials had more negative implicit attitudes than participants exposed to an image of a not-overweight “Bob”. In Study 1, participants exposed to positive feedback about Bob had more positive explicit attitudes toward Bob than those who received negative feedback. In Study 4, explicit attitudes toward Bob were more positive among participants who were assigned to receive positive verbal statements about him than among participants who received ambiguous verbal statements about Bob.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref.</td>
<td>Study</td>
<td>Sample</td>
<td>Measures</td>
<td>Findings</td>
<td>Explicit and Implicit Measures</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>[59]</td>
<td>O’Brien, Hunter &amp; Banks (2007) New Zealand</td>
<td>344 university students (180 physical education [PE] students, 164 psychology students) Mean age = 20 years 67% female</td>
<td>Explicit&lt;br&gt;Anti-fat Attitudes Questionnaire [14].&lt;br&gt;Several separate items assessed the importance of physical abilities. The Body Esteem Scale (Franzoi &amp; Shields, 1984) assessed feelings toward body parts and physical abilities.&lt;br&gt;&lt;br&gt;Implicit (Three IATs)&lt;br&gt;Target categories (words): fat people, thin people (IAT 1, 2, 3).&lt;br&gt;Attribute categories (words): good, bad (IAT 1); motivated/lazy (IAT 2); Smart/stupid (IAT 3).</td>
<td>Strong implicit anti-fat attitudes were evident, and these were most pronounced for PE students in their third year of training. All students reported explicit anti-fat attitudes.</td>
<td>The three IATs were moderately related with the correlations ranging from $r = .37$ to $r = .48$ (all $p$ values &lt; .001).&lt;br&gt;The implicit and explicit measures of anti-fat attitudes correlated significantly (range $r = .23$ to $r = .31$, all $p$ values &lt; .05) among the PE students only.</td>
</tr>
<tr>
<td>[65]</td>
<td>O’Brien, Hunter, Halberstadt et al. (2007) New Zealand</td>
<td>Study 1&lt;br&gt;227 university students (61.7% female) Mean age = 19.98 years&lt;br&gt;&lt;br&gt;Study 2&lt;br&gt;134 university students (73.9% female) Mean age = 20 years</td>
<td>Explicit&lt;br&gt;The Anti-fat Attitudes Questionnaire [14].&lt;br&gt;The Physical Appearance Comparison Scale (PACS; Thompson, Heinberg, &amp; Tantleff, 1991) assessed the tendency to make physical-related comparisons with others.&lt;br&gt;The appearance orientation and evaluation scales of the Multidimensional Body-Self Relations Questionnaire- Appearance Scales (MBSRQ-AS; Cash, 2000) assessed participants’ body image.&lt;br&gt;&lt;br&gt;Implicit (Study 2 only)&lt;br&gt;Target categories (words): fat people, thin people. Attribute categories (words): good, bad.</td>
<td>There were no gender differences in implicit anti-fat attitudes. A higher level of investment in one’s physical appearance, a greater tendency to make physical comparisons with others, and a lower BMI were associated with stronger implicit anti-fat attitudes.</td>
<td>Higher implicit anti-fat attitudes correlated with the explicit belief that being overweight is largely due to a lack of personal control ($r = .20, p &lt; .01$).</td>
</tr>
<tr>
<td>[81]</td>
<td>Robertson &amp; Vohora (2008) United Kingdom</td>
<td>57 fitness professionals and 56 regular exercisers (university students)</td>
<td>Explicit&lt;br&gt;Participants rated their beliefs about “thin people” and “fat people” using the attributes “good” and “bad” and “lazy” and motivated” on a semantic differential scale.&lt;br&gt;&lt;br&gt;Implicit (Two IATs)&lt;br&gt;Target categories (words): fat people, thin people (IAT 1/2).&lt;br&gt;Attribute categories (words): bad, good (IAT 1); lazy, motivated (IAT 2).</td>
<td>Significant anti-fat bias was evident for fitness professionals and regular exercises on all of the implicit and explicit measures of good-bad and motivated-lazy.</td>
<td>The two implicit measures were moderately correlated for both fitness professionals ($r = .38, p &lt; .05$) and for regular exercisers ($r = .41, p &lt; .01$).&lt;br&gt;The implicit and explicit measures of weight bias were not significantly correlated.</td>
</tr>
</tbody>
</table>
Table 1. Contd....

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>[63]</td>
<td>Rudman et al. (2002)</td>
<td>302 university students who belonged to one of four minority groups (48 Jewish Americans, 89 Asian Americans, 53 overweight people, 112 people from low socioeconomic backgrounds)</td>
<td>Implicit (Four IATs)</td>
<td>Attributes categories (words): pleasant and unpleasant words.</td>
<td>Performance on the implicit and explicit measure of in-group bias were uncorrelated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IAT 3 (Sizeism IAT): words representative of overweight people and slim people.</td>
<td>IAT 4 (Classism IAT): words representative of poor people and rich people.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Explicit</td>
<td>Participants indicated their feelings toward each of the minority groups on a feeling thermometer. They also rated the social status of the four minority groups.</td>
<td></td>
</tr>
<tr>
<td>[61]</td>
<td>Schwartz et al. (2003)</td>
<td>389 health professionals specializing in obesity (64% obesity researchers; 9% clinicians, 24% both)</td>
<td>Implicit (Three IATs)</td>
<td>Target categories (words): fat people, thin people (IAT 1-4).</td>
<td>Significant implicit and explicit anti-fat bias was evident. Individual characteristics that were associated with lower implicit anti-fat bias were being male, being older, being happier, having a higher BMI, having friends who are obese, and reporting an understanding the experience of obesity.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Attribute categories (words): Bad, good (IAT 1); Lazy, motivated (IAT 2); Stupid, smart (IAT 3); Worthless, valuable (IAT 4).</td>
<td>Explicit</td>
<td>The correlations between the implicit measures and the explicit semantic differential scales were not reported. A higher degree of self-reported understanding of obesity and having more obese friends was significantly associated with lower bias on the stupid-smart IAT ($r = -0.20$, $p &lt; 0.05$ and $r = -0.19$, $p &lt; 0.05$, respectively).</td>
</tr>
<tr>
<td>[64]</td>
<td>Schwartz et al. (2006)</td>
<td>4,283 individuals recruited online (83% females, 17% males) Mean age = 34.6 years 85% Caucasian</td>
<td>Implicit (Three IATs)</td>
<td>Target categories (words): fat people, thin people. Attribute categories (words): Bad/good (IAT 1); Lazy/Motivated (IAT 2); Lazy/Anxious (IAT 3).</td>
<td>Implicit and explicit anti-fat bias and stereotypes were evident, and these attitudes were weaker among individuals with a higher BMI. A significant proportion of individuals were willing to make extreme trade-offs in order to avoid being obese, and this was more pronounced among individuals with a lower BMI.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Explicit – Yes</td>
<td>Three items assessed explicit attitudes and stereotypes about fat people and thin people. Participants also responded to items relating to personal trade-offs that they would make to avoid being obese.</td>
<td>The correlations between the implicit and explicit measures were not reported.</td>
</tr>
<tr>
<td>[62]</td>
<td>Teachman &amp; Brownell (2001)</td>
<td>84 health professionals who treat obese individuals (71% male), mean age = 48 years</td>
<td>Explicit</td>
<td>Ratings of feelings about thin people and fat people (good/bad, motivated/lazy) on semantic differential scales.</td>
<td>Health professionals had strong implicit negative attitudes toward overweight individuals. They did not explicitly evaluate overweight people as “bad”, but did weakly endorse thin people as “motivated”.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit (IAT)</td>
<td>Target category labels: fat people and thin people (IAT 1, 2). Attribute category labels: bad/good, lazy/motivated.</td>
<td>The good-bad implicit and explicit measures did not correlate. The lazy-motivated implicit and explicit measures were correlated ($r = 0.25, p &lt; 0.05$).</td>
</tr>
</tbody>
</table>
Table 1. Contd.:

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>[6]</td>
<td>Teachman et al. (2003) U.S.A.</td>
<td><strong>Study 1</strong>&lt;br&gt;144 adults*&lt;br&gt;(54% female)&lt;br&gt;Mean age = 35 years</td>
<td><strong>Explicit</strong>&lt;br&gt;Study 1: Fat Phobia Scale (FPS; Robinson et al., 1993) which uses a semantic differential scale.&lt;br&gt;Study 2: Ratings of feelings about thin people and fat people (good/bad, valuable/worthless) on semantic differential scales.&lt;br&gt;<strong>Implicit (IAT)</strong>&lt;br&gt;Study 1: Identical to the IATs utilized by [62].&lt;br&gt;Study 2a: Five IATs as follows&lt;br&gt;(IAT 1) – target categories: fat people and thin people; attribute categories: bad/good.&lt;br&gt;(IAT 2) – identical to IAT 1 except pictorial stimuli were used to depict fat people and thin people.&lt;br&gt;(IAT 3) – target categories: fat people and thin people (pictorial stimuli); attribute categories: valuable/worthless.&lt;br&gt;(IAT 4) – target categories: overweight people and underweight people; attribute categories: bad/good.&lt;br&gt;(IAT 5) – target categories: fat and thin; attribute categories: me/not me.&lt;br&gt;Study 2b: Two IATs identical to IAT 1 and IAT 2 utilised [62].</td>
<td>Implicit anti-fat bias was demonstrated across all three studies, whilst explicit anti-fat attitudes were not evident. In Studies 2a and 2b, implicit anti-fat bias was lower among overweight individuals who were primed to feel empathy toward an obese person, but this was not the case for individuals in the same prime condition with a BMI less than 25.</td>
<td>Study 1: The good-bad IAT and the lazy-motivated IAT correlated to a moderate degree ($r = .33, p &lt; .05$)&lt;br&gt;The lazy-motivated IAT, but not the good-bad IAT, correlated with the FPS ($r = .29, p &lt; .05$).&lt;br&gt;Study 2a: Composite implicit bias scores and explicit bias scores were calculated by averaging across the five IAT tasks and three of the explicit tasks, respectively. The total implicit and explicit bias measures were moderately correlated ($r = .36, p &lt; .05$).</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Study 2a</strong>&lt;br&gt;90 female university students&lt;br&gt;Mean age = 21 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Study 2b</strong>&lt;br&gt;63 adults*&lt;br&gt;(51% female)&lt;br&gt;Mean age = 42 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>* Recruited on a beach</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[57]</td>
<td>Vartanian et al. (2005) Canada</td>
<td><strong>Study 1</strong>&lt;br&gt;56 female university students, mean age = 19.5 years&lt;br&gt;(28 restrained eaters, 28 unrestrained eaters)</td>
<td><strong>Explicit</strong>&lt;br&gt;Study 1: five items designed to assess general beliefs and attitudes toward body weight.&lt;br&gt;Study 2: Attitudes and beliefs toward body weight were rated as in Study 1. Participants also rated fatness and thinness on a semantic differential scale. Thin ideal internalisation was assessed using the Sociocultural Attitudes Toward Appearance Questionnaire (SA-TAQ); (Heinberg, Thompson, &amp; Stomer, 1995).&lt;br&gt;The Restraint Scale (RS; Herman &amp; Polivy, 1980) was used to classify participants according to their level of dietary restraint.</td>
<td>Restrained and unrestrained eaters had equally strong implicit negative attitudes toward fatness in both studies. Restrained eaters had more negative explicit attitudes toward fatness than did restrained eaters. BMI did not moderate scores on either the implicit or explicit measures.</td>
<td>The IAT measures and the explicit beliefs/attitudes (including thin internalisation) were not correlated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Study 2</strong>&lt;br&gt;53 female university students, mean age = 21 years&lt;br&gt;(22 restrained eaters, 31 unrestrained eaters)</td>
<td><strong>Implicit (IAT)</strong>&lt;br&gt;Target categories: thin words and fat words. Attribute categories: pleasant and unpleasant words.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Asian Americans, poor people, and overweight people. Overweight people differed from the other minority groups in that they tended to implicitly prefer slim people to their own group members, and they also exhibited only modest explicit preference for their in-group. This is an important finding because it suggests that overweight individuals may lack the protective function that in-group preference may provide [17]. More research is required to examine potential moderating or mediating influences on the in-group and out-group phenomenon including self-esteem and race, and tendency toward social comparison [18].

The evidence for a relationship between implicit anti-fat attitudes and individual characteristics is mixed. In three studies, higher body mass index (BMI) was associated with weaker implicit anti-fat bias, as indexed by the IAT [61, 64, 65]. Vartanian et al. [57], however, examined implicit attitudes toward fatness and thinness among restrained and unrestrained eaters and found no moderating influence of BMI. It seems that BMI may influence the magnitude of implicit anti-fat bias in some instances, but in general, automatic anti-fat attitudes exist across the weight spectrum [18]. In two studies, female gender was associated with stronger implicit anti-fat attitudes [58, 61], although several studies show no gender differences in implicit weight bias [e.g., 62, 64, 66, 67]. It is possible that the underlying associative memory networks between concepts relating to body size and valence may be more chronically accessible in females than in males because of females’ past and repeated exposure to the notion that “thin is good” and “fat is bad”. Clearly, gender influences in implicit weight bias warrants elucidation in future research. It will also be prudent in future studies to examine implicit weight and shape biases in men using stimuli representative of the male muscular ideal, as well as concepts related to “fatness” and “thinness”, given evidence that exposure to images of idealised male bodies has a significant negative impact on men’s body satisfaction [68].

Body image concerns have also been examined as potential correlates and moderators of implicit anti-fat attitudes. Ahern and Hetherington [55] instructed participants to categorise images representative of thin people and fat people according to congruent category labels (“fat/unpleasant” and “thin/pleasant”) and incongruent category labels (“fat/pleasant” and “thin/unpleasant”). Contrary to expectation, scores on a measure of thin internalisation did not correlate with the IAT index of anti-fat attitudes. In a more recent study, images of underweight females and of average weight females were employed as the target stimuli in an IAT and were paired with positive and negative attributes. On average, participants demonstrated more positive implicit attitudes toward normal weight images than ultra-thin images. Yet individuals with a higher drive for thinness tended to evaluate thin images more positively, especially if they viewed the media as an important source of information about beauty and fashion [54]. Vartanian et al. [57] used a similar IAT (with “fat” and “thin” words) to examine implicit attitudes toward fatness and thinness in restrained and unrestrained eaters. It was predicted that both groups would demonstrate implicit negative attitudes toward fatness because of exposure to pervasive cultural messages promoting the desirability of thinness, but that restrained eaters would exhibit stronger explicit negative evaluations of fatness. Both predictions were supported. Thus, the relationship between implicit anti-fat bias and individual differences, including body image concerns and BMI, appears complex and may be affected by the importance an individual attaches to the attitude [69], and by variables such as the context in which the stimuli are encountered, and attentional focus (see [29]); this will be examined more closely in the Discussion section.

The relationships between explicit and implicit measures of anti-fat attitudes vary across experiments. In several studies, negligible correlations were found between the implicit and explicit measures of anti-fat attitudes [18, 57, 63]. In other studies, the “good/bad” measures of implicit and explicit attitudes were uncorrelated, but the “lazy/motivated” IAT and explicit measure were significantly related [58, 62, 6, Study 1]. A possible explanation is that it is more socially acceptable for proponents to explicitly express a negative stereotype which associates overweight people with a behaviour, such as “laziness”, than to endorse an overweight person as “bad” [57, 62]. That is, contrary to Crandall’s [14] finding, individuals, at least in an experimental context, may tend to suppress or challenge their negative evaluation of fatness on explicit measures, as is often the case for explicit

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>[18]</td>
<td>Wang et al. (2004)</td>
<td>U.S.A.</td>
<td>Study 1</td>
<td>Explicit: None</td>
<td>Strong negative attitudes and stereotypes toward fat people were evident on all of the IATs. Participants reported that fat people were lazier and less motivated than thin people, but no other explicit anti-fat bias was evident. There were no differences on the implicit or explicit measures as a function of age, sex, or BMI.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Study 2: eight items assessed attitudes toward fat people and thin people.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Implicit: Study 1: Target categories: fat people and thin people. Attributes categories: bad-good. Study 2: Four IATs as follows</td>
<td></td>
<td>In Study 2, there were moderate to high correlations within the implicit measures ranging from ( r = .38 ) to ( .68 ), all ( p ) values &lt; .01.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Target categories: fat people and thin people. Attribute categories: good-bad (IAT 1); smart-stupid (IAT 2); lazy-motivated (IAT 3); worthless-significant (IAT 4).</td>
<td></td>
<td>The implicit and explicit measures were uncorrelated.</td>
</tr>
</tbody>
</table>

| Study 1: Contd…. | | | | | |

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>In Study 2, there were moderate to high correlations within the implicit measures ranging from ( r = .38 ) to ( .68 ), all ( p ) values &lt; .01.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>The implicit and explicit measures were uncorrelated.</td>
</tr>
</tbody>
</table>
measures of racial prejudice [e.g., 32, 70]. We now turn to an alternative measure of implicit attitudes, called the affective priming task, which has recently been applied to assess automatic affective reactions toward body-related words and images.

Studies Using the APT and the Lexical Decision Task

A series of studies were conducted with undergraduate females, using the APT to examine automatic attitudes toward body-related stimuli [43, 53]. A summary of the results of these studies is presented in Table 2 (note that the studies are numbered according to the order in which they are first presented in the text). In the studies by Watts and colleagues [43, 53] body-related words or images were the priming stimuli, and positive and negative nouns or adjectives comprised the target stimuli. In three of the experiments, the primes were individually selected in an initial evaluation task. Each participant was presented with the set of potential primes (one at a time) and was required to judge quickly whether each word was “good” or “bad”. Then the negative and positive primes with the fastest response latencies were automatically inserted as the primes in the APT. In the fourth experiment the primes were normatively selected. In a pilot study, graduate psychology students judged whether each prime was “good” or “bad”. Then the primes that were most consistently rated as “good” or “bad” were selected as the normative primes for the APT. It was predicted that responses to congruent pairs (e.g., “chubby” followed by “bad”) would be faster than responses to incongruent pairs (e.g., “slender” followed by “bad”). Further, it was predicted that females with elevated concerns about weight, shape, or dieting would demonstrate stronger implicit attitudes toward body-related words and images. Automatic attitudes toward both individually- and normatively- selected body-related words and images were evident. In all four experiments, participants were faster to respond to congruent pairs relative to incongruent pairs [43, 53]. There was, however, minimal evidence for moderation of automaticity by individual differences in concerns about weight, shape, or dieting [43, 53]. The absence of moderating effects is consistent with the pervasive implicit negative attitudes toward overweight obtained in several of the IAT studies [e.g., 18, 55, 57] although this contrasts with the positive correlation between the concepts “fatness” and “thinness” and negative attitudes toward overweight [71]. It is likely, therefore, that by the time individuals reach adulthood, they have developed well-rehearsed and complex associative networks in memory between the concepts “fatness” and “thinness” and negative and positive affective nodes, respectively. The current data suggest that these associative memory structures are readily activated, with minimal effort, by salient environmental cues.

In two experiments [43, 53], the processing advantage for congruent pairs persisted at the long delay. It appears that anti-fat attitudes were activated at the short delay and participants may not have been motivated to suppress or to challenge these responses at the long delay. This issue, however, warrants further investigation because it was not consistent across all four experiments. It should also be noted that in several IAT studies demonstrating implicit anti-fat bias, endorsement of “fat is bad” and “thin is good” was not evident on all of the explicit self-report measures [18, 62], suggesting that socially desirable responding may have occurred.

To date, one experiment only has examined the relationship between implicit weight bias and a behavioural measure [27]. In that study, participants were briefly presented with images of overweight and thin females, together with neutral images. The primes were followed by either a positive or negative word that was stereotypical of overweight or thinness, or not stereotypical of weight, or a nonword. Participants were required to judge as quickly as possible whether the target was a word or a nonword. Consistent with all of the other studies reviewed, participants demonstrated significantly more negative attitudes toward overweight than toward thin individuals. Implicit, but not explicit attitudes correlated with the behavioural measure such that participants with stronger implicit negative attitudes toward fat primes compared to thin primes elected to sit farther away from an overweight person. This finding is consistent with studies of racial prejudice which have demonstrated that the activation of implicit negative racial attitudes among Caucasian participants predicts their negative spontaneous behaviours (e.g., nonverbal behaviours) toward African American confederates, including reduced eye contact [15] and decreased smiling [31].

DISCUSSION

Prevalence and Automaticity of Anti-Fat Attitudes

The empirical data highlight that implicit anti-fat attitudes and beliefs are widely held and relatively universal. Robust implicit anti-fat bias is evident among many groups including university students, members of the general public, health professionals who work with obese clients, and among those who are themselves overweight or obese. In the IAT and APT studies reviewed, the duration of exposure to the experimental weight-related stimuli was brief such that participants did not have time to plan and execute intentional responses. This suggests that automatic affective processing occurs upon brief exposure to both words and images representative of body size and shape. The data are consistent with the theoretical conceptualization of implicit attitudes as patterns of activation that occur within associative memory networks in response to exposure to relevant stimuli [23]. In Western society the cultural norm for female beauty and attractiveness includes extreme slenderness [2], and from an early age children are aware of the negative connotations of being overweight [71]. It is likely, therefore, that by the time individuals reach adulthood, they have developed well-rehearsed and complex associative networks in memory between the concepts “fatness” and “thinness” and negative and positive affective nodes, respectively. The current data suggest that these associative memory structures are readily activated, with minimal effort, by salient environmental cues.

Relationship Between Implicit and Explicit Measures of Anti-Fat Attitudes and Stereotypes

A contentious issue in the social cognitive literature is whether implicit and explicit measures represent different aspects (unconscious, and conscious respectively) of a unitary attitudinal construct, or whether they tap into independent constructs [30]. If implicit and explicit measures are indexing a unitary construct, then they would be expected to correlate. If the measures tap into different constructs, then it is unlikely that they will correlate substantially. Wilson et al.’s [24] model of dual attitudes postulates that it is possible to hold two different attitudes toward the same object (implicit and explicit), which differ qualitatively from one another. Implicit attitudes are posited to be stable memorial
Table 2. Summary of Experimental Studies Examining Implicit Anti-Fat Bias Using the Affective Priming Task or the Lexical Decision Task

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
</table>
| [27] | Bessenoff & Sherman (2000) U.S.A. | 127 university students (64 females, 63 males) | Explicit  
Participants were selected such that they were either low or high on a measure of fat prejudice as assessed by the Dislike subscale of the Anti-fat Questionnaire (Crandall, 1994).  
Implicit (Lexical Decision Task)  
Primes: Images of fat and thin females, and neutral images (e.g., mug, chair).  
Targets: Positive and negative fat-stereotypic (e.g., kind, insecure), thin-stereotypic (e.g., confident, selfish), and weight stereotypic-relevant words (e.g., violent, musical).  
Each prime was presented briefly (15 ms) and was followed by a positive, negative, or neutral word, or a non-word letter string. Participants judged whether the target was a word or a non-word as quickly as possible.  
Behavioural Measure  
The distance that participants chose to sit from an overweight person. | There was evidence for implicit anti-fat bias, but not pro-thin bias. That is, participants were faster to respond to negatively valenced traits that had been preceded by fat primes than thin primes, relative to neutral primes, but responses to positive traits were not facilitated when they were preceded by thin rather than fat primes (relative to neutral primes). Participants low and high on explicit prejudice did not differ in their automatic associations to the thin or fat primes. | The anti-fat component of implicit attitudes was positively correlated with explicit dislike of fat people ($r = .19, p < .05$). Individuals who reported greater dislike of fat people responded more quickly to negative traits than to positive traits when these were preceded by a fat prime. Implicit, but not explicit, attitudes correlated with the behavioural measure. Participants with greater negative than positive activation for fat primes compared to thin primes elected to sit farther away from an overweight person. |
| [43] | Watts et al. (2008) Australia | Study 1  
87 female university students  
Mean age = 18.91 years  
Study 2  
72 female university students  
Mean age = 19.28 years | Implicit (APT)  
Primes: body-related images (e.g., depicting an overweight female or a slender female), and nonbody-related images (e.g., depicting a crime scene or a parade).  
Target words: nonbody-related nouns (e.g., disaster, holiday).  
Implicit (APT)  
Primes: body-related images (e.g., depicting an overweight female or a slender female), and neutral primes (four colored squares in muted pastel shades).  
Target words: nonbody-related adjectives (e.g., awful, splendid).  
Explicit  
In both studies, participants completed self-report measures of appearance schematicity, body dissatisfaction, thin internalisation, and dietary restraint. | In each study, participants were faster to respond to affectively congruent pairs (e.g., an image of an overweight female paired with “disaster” and an image of a slender female paired with “holiday”) relative to incongruent pairs (e.g., an image of an overweight female paired with “disaster” and a slender female paired with “disaster”). | The self-report measures of appearance schematicity, thin internalisation, body dissatisfaction, and dietary restraint did not correlate reliably with the implicit measure in either study. |
representations that are automatically activated. In contrast, explicit attitudes require conscious effortful processing and motivation in order to be retrieved. This model predicts little or no correlation between implicit and explicit measures because of the qualitative differences between them [24].

The current data suggest that the relationship between implicit and explicit measures of anti-fat attitudes is complex. Strong implicit negative evaluation of fatness is correlated with explicit, self-report measures of anti-fat bias in some studies [27, 65, 66, 6; Study 2a] but not in others [e.g., 6, 57, 62], although this varies according to the individual characteristics of the participants [e.g., 59]. Several studies have demonstrated that performance on the IAT consistent with implicit anti-fat stereotypes correlates positively with explicit measures of the negative stereotype that "overweight individuals are lazy" [6, 58, 62], but once again, this is not always the case [18].

A recent meta-analysis examining the correlation between the IAT and explicit measures confirms the complexity of this issue [72]. Contrary to the hypothesis that implicit and explicit attitudes are independent, the authors found that on average, the IAT and explicit measures correlate positively and to a moderate degree ($r = .24$). Interestingly, in contrast to the social desirability argument put forward by previous researchers [34], the correlations were not influenced by the degree of social desirability surrounding each topic of evaluation. In general correlations between the IAT and explicit measures were higher when there was greater opportunity for spontaneous (rather than deliberative) responses on the explicit self-report measures [72]. This finding fits with the notion that implicit measures primarily reflect automatic associations, whereas explicit self-reports depend on the effortful retrieval of information from memory [23, 24].

Gawronski and Bodenhausen [23] posit that automatic affective reactions (implicit attitudes) are generally used as the "default" basis for explicit evaluations (p. 694), in which case the two types of attitudes would be expected to correlate. Correlations between implicit and explicit measures are expected to decrease, however, as a function of cognitive elaboration. Elaboration refers to the degree of active, effortful thought devoted to an attitude object [23]. That is, if an individual has the opportunity to actively consider information about an attitude object in addition to their implicit attitude toward that object, then there is a possibility that this additional information will be inconsistent with the individual’s implicit evaluation. In this case, the individual’s implicit attitude toward the object is unlikely to be used as the default basis for their explicit judgments [23]. Hence, explicit self-reported attitudes are likely to reflect automatic associations to a greater extent when people do not have the motivation, cognitive capacity or time to retrieve additional information from memory that may compete with or invalidate individuals’ implicit associations [72]. The meta-analysis of Hofmann et al. [72] also revealed that correlations between implicit and explicit measures were higher when there was greater conceptual overlap between them; that is, when the automatically activated attitudes were directly relevant to the judgments or evaluations required on the explicit measures.

The correlation between scores on the implicit measures of anti-fat bias and explicit measures of anti-fat attitudes in the studies reviewed here range from $r = -.20$ to $r = .36$ [58, 62, 65, 66]. It is unclear whether Hofmann et al.’s [72] two variables, spontaneity and conceptual overlap, were influential in the current studies. A fruitful focus for future research would be to investigate whether manipulation of variables such as cognitive load, spontaneity, motivational state, and the conceptual match between implicit and explicit measures influences the relationship between implicit and explicit anti-fat attitudes.

**Individual Correlates of Implicit Attitudes**

The nature of implicit anti-fat attitudes and their relationship with individual differences is also complex because

<table>
<thead>
<tr>
<th>Ref</th>
<th>Study</th>
<th>Sample</th>
<th>Measures</th>
<th>Findings</th>
<th>Explicit and Implicit Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>[53]</td>
<td>Watts &amp; Cranney (2009)</td>
<td>Australia</td>
<td>[Implicit (APT)]&lt;br&gt;- Primes: body-related words (idiographically selected e.g., fat, slenderness).&lt;br&gt;- Target words: nonbody-related nouns (e.g., awful, splendid).</td>
<td>In each study, participants were faster to respond to affectively congruent pairs of words (e.g., “fat” paired with “awful” and “slenderness” paired with “splendid”) relative to incongruent pairs of words (e.g., “fat” paired with “splendid” and “slenderness” paired with “awful”).</td>
<td>Study 1: Participants’ scores on the measures of appearance schematics, thin internalisation, body dissatisfaction, and dietary restraint did not correlate reliably with the implicit measure.</td>
</tr>
</tbody>
</table>
there is considerable heterogeneity in the empirical data. For instance, stronger implicit evaluation of “fat is bad” has been found in females compared to males [58, 61], and among physical education students compared to psychology students [59]. In some studies, but not others, higher BMI was associated with lower anti-fat bias [58, 61, 64, 65, Studies 2a and 2b]. In the aforementioned meta-analysis, approximately half of the variability across correlations could be attributed to moderator variables [72]. Yet, in other IAT studies and the affective priming studies examining weight bias, individual differences in body image concerns were not associated with the implicit measures [43, 55, 57]. In one study examining attitudes toward images of ultra-thin females, higher drive for thinness was associated with more positive attitudes toward ultra-thin models [54]. How do we make sense of this variability and lack of cohesion in the data? One explanation lies in the notion that implicit attitudes are not as inflexible and immutable as originally hypothesised, and that they are influenced by factors such as the focus of attention, counter-stereotypical information, and context.

Implicit Attitudes: Immutable Structures or Context Dependent?

There is an impressive and growing body of evidence in the social cognitive literature suggesting that automatic attitudes are more malleable than previously suggested (see [29] for review). In other words, it may be possible to change implicit attitudes either by modifying the underlying associative structures or by altering the pattern of automatic activation that is produced by manipulating contextual or situational variables [23].

Contextual cues influence the type of implicit attitudes that are expressed. It is assumed that attitudes toward particular objects or people are represented in a multi-faceted manner within an associative memory network [23, 45]. Hence, the manipulation of situational variables can change the pattern of associations that are activated in response to a particular attitude object. For example, in a study examining the influence of social context on racial prejudice, participants who viewed a video clip of African Americans in a positive stereotypic context (i.e., an outdoor barbeque) showed a significantly larger decrease in automatic negative bias toward that group relative to the baseline IAT, compared to participants who viewed a video clip of African Americans in a negative stereotypic context (i.e., a gang-related setting) [73]. This study suggests that the pattern of associations activated by a given individual (or other attitude object) can vary as a function of the context in which the individual is encountered [23].

There is evidence that the associative structures, as well as patterns of activation, underlying the expression of implicit attitudes may be changed through repeated pairings with counter-evaluative information. For example, Karpinski and Hilton [30] found that implicit prejudice against older adults was modified by repeatedly pairing the words youth and elderly (conditioned stimuli) with positive and negative words (unconditioned stimuli). That is, participants’ implicit prejudice against older people in an IAT was reduced when youth was repeatedly paired with negative words and elderly with positive words, compared to repeated pairings of youth with positive words and elderly with negative words. In another study, participants who were exposed to admired African Americans (e.g., Bill Cosby) compared to those who were exposed to disliked African Americans (e.g., OJ Simpson) or to non-racial stimuli, demonstrated less automatic prejudice toward African Americans on an IAT, and this difference persisted when tested 24 hours later [74]. These findings are promising and suggest that such techniques require further investigation to determine the usefulness of similar manipulations in the reduction of implicit anti-fat attitudes.

One important series of studies examined implicit anti-fat bias and its potential modification through manipulation of causal attributions for obesity (genetics versus lifestyle behaviours; Study 1), and by attempting to evoke empathy toward overweight individuals (Studies 2a and 2b) [6]. In Study 1, participants were allocated to one of three prime conditions. Participants either received no prime (no-prime condition) or were asked to read a “news article” summarizing research indicating either that genetics is the primary cause of obesity (genetics condition) or that the primary causes of obesity are overeating and lack of exercise (behaviour condition). The individuals in the genetics condition and behaviour condition read the news articles prior to all three groups completing two IATs (evaluative: fat people/thin people paired with good/bad and stereotypical: fat people/thin people paired with motivated/lazy), followed by completion of explicit measures of fat bias. It was expected that implicit and explicit anti-fat bias would be lower when participants were informed that the primary cause of obesity was genetics, and higher when the primary causes were indicated as overeating and lack of activity. The hypotheses were only partially supported. Implicit anti-fat bias was higher for the behaviour condition compared to the no-prime condition, but there was no effect of exposure to information about genetic causes of obesity upon either implicit or explicit anti-fat bias. That is, the manipulation to increase anti-fat bias was successful, but the manipulation to decrease anti-fat bias was not. The empathy manipulation in Studies 2a and 2b was also largely unsuccessful in reducing anti-fat bias, except among overweight individuals [6].

One possible explanation for the resistance of anti-fat bias in these studies is that the participants were not highly motivated to change their bias. There is evidence that highly motivated individuals can modify the automatic operation of stereotypes and prejudice [29]. For example, individuals may automatically inhibit negative stereotypes and activate positive ones if doing so would bolster their self-image [e.g., 7]. This appears counter-intuitive, given that traditional definitions of automaticity suggest that automatic processes are minimally influenced by attention, perceiver strategies, and motivational states [46]. However, Blair [29] suggests that this strict definition requires modification. That is, just because a process is automatic and requires minimal attention or motivation in order to proceed, does not mean that the perceiver’s attention (and presumably internal motivational state) cannot influence that process (e.g., Logan cited in [29]). In Teachman et al.’s study [6] only overweight participants showed a lowering of implicit anti-fat bias when exposed to discriminatory information about an overweight individual that was designed to evoke empathy. It is possible, therefore, that overweight persons were more motivated to inhibit their negative implicit fat bias than average weight.
individuals, in order to protect the threat that the discriminatory material posed to their self image.

The finding that the perceivers focus of attention can influence the type of attitudes that are automatically activated upon encountering a stimulus [29] also demonstrates that automatic attitudes are malleable constructs [73]. For example, Wittenbrink et al. [76] examined how focusing upon evaluation versus meaning can influence participants’ automatic evaluations of African Americans. All participants completed a priming task in which the primes were the category labels “Black” and “White” and the target items were either adjectives (stereotypic of African Americans, stereotypic of White Americans, or nonstereotypic), or nonwords. The focus of attention was manipulated such that half of the participants were required to judge whether the target item was a word or a nonword (conceptual judgment), and the other half were asked to judge whether the target item was good or bad (evaluative judgment). As predicted, the judgment task influenced the type of attitudes that were automatically activated. Higher levels of generalised automatic prejudice (i.e., associations between Black-negative and White-positive with adjectives unrelated to either category) were produced by participants who focused upon evaluation than participants who focused on meaning. This finding supports the notion that attitudes are represented in memory in a multifaceted, heterogeneous manner such that subtle variations in attential focus or contextual cues, when an attitude object is encountered, can activate different automatic evaluations [23, 76]. Whilst the focus of attention has been examined as a moderator of automatic gender and racial stereotypes [76, 77], it has not yet been explored in relation to anti-fat attitudes. This is an important area for future research because altering one’s focus of attention when interacting with an overweight or obese individual may be one avenue for minimizing the impact of implicit negative anti-fat attitudes that may arise.

Further Avenues for Future Research

Implicit attitudes predict behaviour, particularly those that are nondeliberative and outside of conscious control [e.g., nonverbal behaviours; 15, 31]. To date, only one study has examined the relationship between an implicit measure of weight bias and behaviour [27]. In view of the widespread stigmatisation of obesity [5] and the negative psychological impact of such stigma [3], the potential relationship between implicit weight bias and other behavioural measures, including indices of fat discrimination, warrants more research attention. For example, a future study could involve priming implicit negative anti-fat attitudes in the presence of either an overweight (condition 1) or average-weight (condition 2) female experimenter. The experimenter could then conduct a face-to-face health/fitness survey with each primed participant during which she records the frequency of the participants’ direct eye contact and amount of smiling. To provide an explicit measure of fat prejudice, the participant could complete a form evaluating the experiment but with several items embedded within this, asking for his or her opinion of the experimenter (e.g., communication skills and professionalism). In this way, the relationship between spontaneous nonverbal behaviours and both explicit and implicit anti-fat attitudes could be determined.

One limitation of the studies which have examined implicit anti-fat bias is that most (cf. [27]) have utilised “thin” versus “overweight” stimuli and category labels, and have then inferred anti-fat bias from the pattern of response latencies to the congruent and incongruent trials. However, it is not possible to conclude with certainty whether the pattern of response latencies indicates anti-fat bias, pro-thin bias, or both [6, 62]. Only one study to date has examined implicit attitudes toward ultra-thin images, compared to normal weight images, and interestingly, the majority of women demonstrated implicit negative attitudes toward the underweight models [54]. Future studies need to include an appropriate control category of words or images that can function as a baseline against which response latencies toward the congruent and incongruent categories can be compared, so that stronger inferences about a true “anti-fat bias” can be drawn. Hence, future research could include three sets of primes or categories in the APT that vary along the single dimension of body size (thin, average-sized, overweight), with the average-sized (normal weight) category functioning as the baseline category.

Currently, implicit anti-fat attitudes have been investigated primarily in non-clinical samples. Vitousek and Hollon [78] postulate that individuals with an eating disorder process information relevant to body shape, food, and dieting differently to those without an eating disorder. They argue that individuals with an eating disorder have highly developed body-related self schemas that fuse information about weight and shape with their concept of self. In view of their intense preoccupation with weight, shape, and dieting, these individuals are likely to demonstrate strong implicit negative bias toward fatness and positive bias toward thinness. These implicit attitudes may also drive some of their maladaptive cognitions (e.g., black and white thinking, cognitive distortions, and cognitive biases). Investigation of implicit attitudes in eating disordered samples may, therefore, assist in elucidating the mechanisms of action underlying cognitive behaviour therapy for eating disorders [79], and may also provide an index of potential changes in body dissatisfaction in response to treatment [80].

We know that implicit attitudes are predictive of behaviour and are likely to form the basis for explicit judgments [15, 23, 27, 31], so an important question for future research is whether implicit anti-fat attitudes can be changed. There is mounting evidence that implicit attitudes can be modified, but this is largely unexplored in the domain of body image and implicit fat prejudice. Although there is promising evidence in the social cognitive literature that implicit attitudes can be modified, it remains unclear whether such changes endure over time and under what conditions they may produce corresponding changes in explicit attitudes [80, 81]. Future research needs to examine both short-term and longer-term changes in implicit attitudes in the domain of implicit weight bias.

CONCLUSION

Studies employing implicit measures of weight bias, including the IAT and the APT, have demonstrated robust automatic activation of attitudes toward a range of body-related stimuli, including words and images. In all of the studies reviewed, individuals typically evaluated stimuli rep-
Resentative of “overweight” negatively. In contrast, stimuli depicting “thinness” were generally evaluated positively. Both patterns of responding confirm that implicit attitudes and stereotypes toward body size and shape are widely held. The relationship between implicit and explicit measures of anti-fat attitudes and how they covary with individual differences, including gender and body image concerns, is complex. In view of the robust and pervasive nature of implicit anti-fat attitudes, an important area for future research is to determine precisely how these automatic associations, once activated, influence behaviours including fat discrimination, dieting, and levels of physical activity. Moreover, the existing evidence that implicit attitudes are not as immutable as originally hypothesised, and that patterns of activation and the underlying associative structures can be modified, requires careful examination in relation to implicit weight bias. The complexity of formulating appropriate messages for combating obesity (by promoting healthy lifestyle behaviours) whilst also challenging weight bias, and promoting acceptance of diversity in body size, awaits further innovative and sustained research endeavour.

ACKNOWLEDGEMENTS

We gratefully acknowledge discussions with Professor Gail Huan and Dr Fiona Monro Aldridge who inspired some of the ideas for this paper. We are also grateful to Dr Fiona White for her comments on an early draft of this manuscript. In addition, we thank Shirley Zhang and Leigh Mellish for their valued contributions to revisions.

REFERENCES


