Measurement of Bias

Bethany A. Teachman and Robyn Mallett


Understanding the assessment of obesity stigma requires familiarity with the measurement of bias against marginalized groups more broadly. In this chapter, we outline the terminology used to describe bias, offer a historical perspective on the measurement of bias with particular emphasis on the role of indirect and automatic measures (reflections of bias that are involuntary or outside conscious awareness), and then describe the primary approaches that have been used to measure stigma of obesity.

Measurement of stigma derives in part from its defining features. Goffman (1963, p. 3) defined social stigma as any aspect of an individual that is deeply discrediting and thereby allows others to discount that individual as “tainted.” Jones and colleagues elaborated by specifying six dimensions on which an individual could be discredited (Jones, Farina, Hastorf, Markus, Miller & Scott, 1984): 1) concealability—whether one can hide a stigma from others; 2) course—the way that a stigma changes over time; 3) disruptiveness—how much the stigma interferes with social interactions; 4) aesthetic qualities—the extent to which the stigma makes an individual repellent or upsetting to others; 5) origin—who is responsible for the stigma or how it was acquired; and 6) peril—the type and degree of danger that the stigma poses for others.

Measurement of weight stigma has been influenced by each of these features: 1) obesity is not concealable; 2) weight often fluctuates over time, so obese people may view their status as temporary (Quinn & Crocker, 1998); 3) weight frequently plays a role in social interactions (e.g., Harris, 1990); 4) fat is often deemed aesthetically unpleasing or even repellent (see Miller,
5) despite strong evidence for the role of biological and environmental factors, people are often held personally responsible for being overweight (e.g., Quinn & Crocker, 1999); and 6) people are judged negatively for even socializing with obese persons (e.g., Hebl & Mannix, 2003), suggesting some degree of associated peril.

Because stigma has the potential to render the possessor as less than “a whole and usual person” (Goffman, 1963, p. 3), a great deal of research has focused on bias, or the way that thoughts, feelings, and behaviors may be altered because of a stigmatizing mark. Research on biased cognitions focuses on stereotypes; research on bias in affect or emotional reactions focuses on prejudice; and research on bias in behavior focuses on discrimination (Fiske, 1998).

Historical Perspective on Bias Measurement

As the nature of prejudice changed, so did its measurement. In the early 1900’s, prejudice was blatant and measurement was direct. Katz and Braly (1933) investigated the content of stereotypes by simply presenting individuals with dozens of adjectives and asking them to indicate the extent to which each adjective was descriptive of members of various races. Adorno (Adorno, Frenkel-Brunswik, Levinson & Sanford, 1950) was one of the first to investigate the nature of prejudiced individuals. Drawing from psychoanalytic theory, he proposed that the authoritarian personality (characterized by excessive conformity and submission to authority) was at the heart of prejudice, and he specified several correlates of prejudice, including religiosity and cognitive rigidity. Bogardus (1933) created one of the first measures of discrimination in the form of the social distance scale, which required individuals to select the forms of contact that s/he would be comfortable having with members of various races, occupations, and religions (e.g., ranging from “I’d exclude them from my country,” to consort with them “As close kin by marriage”).
Early bias research typically focused on race-related bias. The Civil Rights movement and subsequent social change in the 1960’s modified social norms and curtailed the expression of blatant prejudice. Although bias and hostility toward individuals from different social groups diminished in overt forms, it remained in subtle, more covert forms. For example, although many Whites claimed they were not biased against Blacks, they resisted school integration. This made clear the need for investigation of the discrepancy between overt and covert or subtler forms of bias. This was accomplished initially by changing the content of the questions to disguise the true meaning of the items. Specifically, subtle scales measure three components: 1) denial of continued discrimination; 2) antagonism toward the target group’s demands; and 3) lack of support for policies designed to help the target group (Swim, Aiken, Hall & Hunter, 1995). For example, rather than asking Whites how much they liked Blacks—a fairly overt measure of bias—the questions asked about values or policy preferences (McConahay, 1983).

Recognizing that some individuals would attempt to disguise their true attitudes in response to changing social norms, researchers developed alternatives to the traditional self-report questionnaires. Assuming that some individuals might not realize they act in discriminatory ways, some researchers looked to aspects of interpersonal communication, including tone of voice, speech content, facial expression, and body language to assess bias (Babad, Bernier & Rosenthal, 1989), while others assessed the likelihood of helping an individual from a different social group (e.g., Gaertner & Dovidio, 1977). One ingenious measure that distinguished overt and covert attitudes was the bogus pipeline (Jones & Sigall, 1971), which misled participants to believe that their untrue responses could be detected via physiological responses to stimuli.
Indirect Measures of Bias

In the 1990’s, Devine and colleagues further clarified the difference between overt and covert forms of bias with the dissociation model, which posits that although people may harbor prejudice at an automatic (involuntary and immediate) level, they can consciously control the expression of prejudice (Devine, 1989). The model suggests that stereotypes are automatically activated when an individual sees a member of a stereotyped group, but that individuals can respond without prejudice by controlling the automatically activated stereotype. Devine’s (1989) dissociation model advanced bias measurement by challenging researchers to create measures that assessed automatic cognitions using advances in technology (i.e., computers).

One method of assessing automatic attitudes is to present individuals with faces or words that belong to their own or a different social group very rapidly (called priming), and then measure the speed of response to a subsequent word or non-word (e.g., Fazio, Jackson, Dunton & Williams, 1995; Wittenbrink, Judd & Park, 1997). For instance, when presented with words or faces related to Blacks, Whites respond more quickly to negative than positive stereotypic attributes (Dovidio, Evans & Tyler, 1986). Other research has used a lexical decision task to assess automatic bias. The lexical decision task requires participants to indicate as quickly and as accurately as possible whether an item is a real word, or a non-word. The response latency or reaction time to make this judgment is interpreted as an indication of the cognitive resources required to attend to and evaluate the item. For example, after being presented with a Black face, individuals who were biased against Blacks should take longer to identify a positive word as being a real word but should take less time to identify a negative word as a real word (e.g., Gaertner & McLaughlin, 1983).
A more recent development is the implicit association test (IAT; Greenwald, McGhee & Schwartz, 1998). The IAT is a widely used measure to reflect automatic memory-based associations, which requires participants to classify words or pictures into superordinate categories (e.g., “Fat People” vs. “Thin People”). Simultaneously, the task requires categorization of stimuli into descriptor category pairs, such as Good vs. Bad (an attitude measure) or Motivated vs. Lazy (a stereotype measure). There are both computerized and paper/pencil versions of the IAT; both assume that classification is facilitated when categories are paired so that they match a person’s automatic associations in memory. The instrument has adequate psychometric properties (see Greenwald & Nosek, 2001).

Bias researchers have also begun to consider how targets of stereotypes, prejudice and discrimination perceive the bias against them (Swim & Stangor, 1998). A variety of methods have been developed to assess a person’s experience, including in-depth interviews, focus groups, and reactions to contrived situations in a laboratory setting. Daily diaries are another unique method in that the observers are not researchers but are members of the target group of interest (Miller & Myers, 1998).

**Obesity Bias Measurement**

As evident from even this brief review, there is a rich and dynamic history to the measurement of stereotypes, prejudice and discrimination. Relative to this history, assessment of stigma of obesity is relatively new. Notwithstanding, researchers have used a variety of explicit (self-reported or endorsed), automatic (outside of conscious control or awareness), and behavioral indicators to capture the pervasive negativity toward overweight persons.

Applications of these measures need to be understood in the context of the unique features of weight stigma, relative to measurement of stigma toward other marginalized groups.
In particular, it is still comparatively acceptable to report anti-fat views (Kilbourne, 1994). In addition, the rising rates of obesity (Berkow, 1997) mean that not only is obesity a highly visible stigma, but is increasingly normative. Further, unlike many marginalized groups, there is less direct legal protection for obese individuals who have been discriminated against (McDermott, 1995) because the Americans With Disabilities Act does not identify weight as a protected characteristic (see Johnson & Wilson, 1995; Roehling, 1999), though the legal status of weight discrimination may be changing (Adamitis, 2000; Solovay, 2000). This lack of social and legal protection may help explain why the majority of explicit measures of weight stigma find that both obese and average weight people report similar levels of anti-fat views, suggesting that no protective in-group bias exists (Crandall, 1994; Wang, Brownell & Wadden, 2004). Finally, unlike race, gender and other marginalized designations, weight is seen as an attribute over which people have control (Quinn & Crocker, 1999; Weiner, Perry, & Magnusson, 1988).

Explicit/Questionnaire Measures

One of the more popular questionnaire measures of weight bias is Crandall’s (1994) Anti-fat Attitudes Test, which was originally designed to evaluate parallels between anti-fat views and symbolic (covert) racism. The 13-item scale includes three subscales: dislike of fat people (e.g., “fat people make me feel somewhat uncomfortable”), fear of fat (e.g., “I feel disgusted with myself when I gain weight”), and beliefs about the controllability of weight or willpower (e.g., “people who weigh too much could lose at least some part of their weight through a little exercise”). The measure has adequate psychometric properties, though to increase internal consistency, Quinn and Crocker (1999) added items to the Dislike and Controllability subscales.

A number of other attitude questionnaires have been used to assess stigma of obesity, but unfortunately, most of the measures have only been used in a small number of studies, and there
is limited psychometric information (e.g., on reliability and validity) available for many of the scales. Allison, Basile, and Yuker (1991) developed companion scales, the 8-item Beliefs About Obese Persons Scale and 20-item Attitudes Toward Obese Persons Scale. One advantage of these scales is that normative data is available for college students and for members of the National Association to Advance Fat Acceptance. Bagley, Conklin, Isherwood, Pechiulis, and Watson (1989) developed the Attitudes Toward Obese Adult Patients with a particular focus on assessing nurse’s attitudes toward their overweight patients, and Price and colleagues developed a series of measures for health professionals and school officials (e.g., Price, Desmond, Ruppert, & Stelzer, 1989). Bray (1972) developed the 47-item Bray Obesity Attitude Scale, which was later revised by Sims (1979) to test attitudes across different ethnic and racial groups. Other obesity attitude measures include sentence completion measures (Canning & Mayer, 1966), ratings of line drawings, and rankings of sketches of various physiques (e.g., Richardson, Goodman, Hastorf, & Dornbusch, 1961).

Some researchers have focused on adjective ratings to reflect stereotypes about “fat people” (e.g., Staffieri, 1967). Often, vignettes or some kind of personal descriptions are presented, and then participants are asked to rate the presented character on various attributes (see Counts, Jones, Frame, Jarvie, & Strauss, 1986; Harris, Harris, & Bochner, 1982). The focus in this case is on an individual overweight person. To evaluate stereotypes about overweight people as a group, researchers have sometimes used semantic differential scales, frequently evaluating attitudes toward fat people relative to thin people. For instance, Teachman, Gapinski, Brownell, Rawlins, and Jeyaram (2003) asked participants to rate their feelings about “fat people” and about “thin people” on negative and positive attributes, and then calculated a difference score between the items. In contrast, the Fat Phobia Scale (Robinson, Bacon, &
O’Reilly, 1993), one of the more rigorously developed instruments, has been used to assess group-level weight bias in a non-relative way. This 50-item, semantic differential scale asks participants to rate their feelings about what “fat people are like” on a series of different opposing dimensions (e.g., smart vs. stupid). This and other non-relative measures (e.g., Harris, Walters, & Waschull, 1991) simplify interpretation, but leave open the question of what constitutes the baseline evaluation.

Few studies have measured stigma of obesity by directly asking overweight persons about their experiences of discrimination, although it appears that participants can readily recall instances of perceived prejudice and discrimination. For instance, 96% of overweight adolescents reported experiences of stigma in qualitative clinical interviews (Neumark-Sztainer, Story, & Faibisch, 1998).

**Implicit Measures**

Most research on anti-fat biases has used explicit measures of attitudes and stereotypes, but new evidence suggests that anti-fat bias can be activated without conscious intention (and perhaps outside of awareness), and can even differ in important ways from explicit views. This incongruence may arise because automatic responses to marginalized groups can occur outside of conscious control or awareness, or because individuals are motivated to deny these responses, perhaps to appear or to actually be fair-minded.

Bessenoff and Sherman (2000) used a lexical decision task to demonstrate that implicit anti-fat evaluations predict how far participants choose to sit from an overweight woman, whereas explicit attitudes do not. In their study, participants were presented with pictures of fat and thin women, and then evaluated fat-stereotypical, thin-stereotypical, and stereotype-irrelevant words. Not surprisingly, the authors found greater implicit activation of negative
evaluations to fat compared to thin women. Researchers have also recently applied the IAT to measure implicit anti-fat biases, based on the idea that classification of words and pictures is faster when categories are paired so that associations with obesity reflect the person’s automatic evaluations. Teachman and Brownell (2001) found strong implicit bias even among health professionals who specialize in obesity treatment and who did not explicitly report negative attitudes. In addition, Grover, Keel, and Mitchell (2003) found that implicit anti-fat attitudes were ubiquitous across both average weight and overweight women and men. Similarly, Teachman et al. (2003) found strong evidence of implicit anti-fat/pro-thin stereotypes of overweight people as lazy, stupid, and worthless among both the general population and college students. Finally, Geier, Schwartz, and Brownell (2003) demonstrated implicit anti-fat biases following presentation of "before and after" diet advertisements.

**Behavioral/Rejection Measures**

Rather than relying on direct evaluation of obese persons, many studies have used more indirect (but arguably more externally valid) behavioral indicators of rejection across domains from health care to housing, employment, school, and relationships. The measures range from micro-level indicators, such as proximity of a chair to an overweight confederate (Bessenoff & Sherman, 2000) or time before salespersons respond to overweight customers (Pauley, 1989), to macro-level surveys with nationally representative samples illustrating the negative relationship between weight and wages (e.g., Pagan & Davila, 1997; Register & Williams, 1990).

There have been a number of studies demonstrating discrimination against overweight persons in the workplace (see review by Roehling, 1999), including field studies like the surveys cited above and experimental studies that involve judgments about hypothetical job applicants or employees whose weight status is manipulated (e.g., Brink, 1988; Klassen, Jasper, & Harris,
Participants are typically given a resume or videotape presentation of a candidate, and weight is either embedded in the resume or is manipulated verbally or through a photo or video of the applicant (e.g., using theatrical prostheses; Pingitore et al., 1994). Job applicants are then rated on a variety of features, including desirability for hiring, personal attributes, fitness for the job, professional qualifications, perceived skills, and professional ethics. These measures indicate that obese persons routinely experience discrimination in the workplace at all stages of the employment process: they are less likely to be hired (Klesges et al., 1990; Roe & Eickwort, 1976); their wages are lower and they experience discrimination on the job (Rothblum, Brand, Miller, & Oetjen, 1990); and promotions occur less frequently (Larkin & Pines, 1979). In an intriguing variation on the job applicant paradigm, Hebl and Mannix (2003) found that average weight applicants perceived to be in a social relationship with an obese person were stigmatized themselves.

Analogous approaches have been used to measure weight stigma across other life domains. To demonstrate housing discrimination, Karris (1977) had obese and average weight college students inquire in person about apartments for rent, and showed that landlords were less willing to rent the apartment to obese potential tenants. In relationships, overweight persons are less likely to be trusted or be chosen as friends or romantic partners (DeJong & Kleck, 1986; Harris, 1990). Similar rejecting views have been assessed in the healthcare field, including negative attitudes by nurses (Maroney & Golub, 1992), medical students (Wiese, Wilson, Jones, & Neises, 1992), doctors (Klein, Najman, Kohrmann, & Munro, 1982; Teachman & Brownell, 2001), and psychotherapists (Davis-Coelho, Waltz, & Davis-Coelho, 2000). In school settings, weight bias has been measured through sociometric nominations by classmates indicating which children are liked and disliked (Cohen, Klesges, Summerville, & Meyers, 1989), which children
are desirable playmates (Jarvie, Lahey, Graziano, & Framer, 1983; Strauss, Smith, Frame, & Forehand, 1985), and which children are “mean” or “nice” (Cramer & Steinwert, 1998). These studies show that weight bias can be demonstrated in children as young as 3-years-old, and that school-based discrimination continues across developmental stages (e.g., parents are less willing to finance their overweight daughters’ college education; Crandall, 1995).

Recommendations for Future Research

Although measurement of weight bias is still a relatively young field, much has been learned about the numerous ways that overweight persons routinely experience stereotypes, prejudice, and discrimination. We conclude this chapter with suggestions for future studies on the measurement of weight bias, based on both current gaps in the obesity literature and exciting advances in stigma research:

1) Evaluating the psychometric properties of the many anti-fat questionnaires will help to establish ‘gold-standard’ measures that have strong reliability and validity.

2) Further research on measures of the experience of obese persons is needed. This includes assessing awareness and sensitivity to stigma and their consequences for behavior, thoughts and feelings, as well as methods used by overweight persons to cope with or respond to perceived bias (e.g., Miller & Myers, 1998).

3) It will be important to assess how attitudes versus beliefs, and implicit versus explicit bias measures, predict discriminatory behaviors against overweight persons.

4) As the field becomes more adept at reducing stigma of obesity, developing instruments that are sensitive to change will be critical. For example, establishing measures of interactions between obese and average weight individuals can foster recognition of covert forms of bias.
5) It will be helpful to more consistently use standard criteria for ‘overweight’ and ‘obese’ to facilitate comparison across studies.

We look forward to seeing how the measurement of weight bias will evolve, given the remarkable advances in the fields of stigma, assessment, and obesity research in recent years.
Measurement of Obesity Bias References


