Explicit and Implicit Stigma Against Individuals With Mental Illness

Andrea Stier and Stephen P. Hinshaw

University of California, Berkeley

Australian Psychologist 42(2) 2007, 106-117
Abstract

Stigma against mental illness has devastating consequences for individuals with mental illness and their families. Empirical findings and qualitative evidence indicate that stigma against mental illness remains rampant in many nations and cultures, constituting a significant barrier to successful treatment, reducing key life opportunities, and predicting poor outcomes over and above the effects of mental illness per se. In this article, we define stigma, examine relevant theoretical perspectives, summarize evidence regarding the pervasive negative impact of stigma on individuals with mental illness, and discuss underlying mechanisms. We focus in particular on assessment issues, highlighting the need for transcending explicit attitudinal measures of stigma, which are susceptible to social desirability concerns and are likely to underestimate true levels of stigma, to include unconscious/implicit indicators and direct behavioral appraisals. A primary goal is to facilitate means of accurately measuring stigma against mental illness as an important step toward reducing its pernicious effects.
The stigmatization of mental illness is currently considered to be one of the most important issues facing the mental health field (Crisp, 2000). An enormous number of individuals are affected by mental illness worldwide: It has been estimated that 1 in 5 persons will suffer from a mental illness each year, with about 6% showing forms that indicate high levels of severity (Kessler, Chiu, Demler, & Walters, 2005; World Health Organization, 2001a). Although individuals with mental illness suffer from a wide range of negative effects and impairments related to the disorder itself, these outcomes are exacerbated by societal stigmatization of their illness. In fact, harsh stigmatization of mental illness occurs across nations and cultures around the world, creating significant barriers to personal development and receipt of treatment (Tsang, Tam, Chan, & Cheung, 2003; World Health Organization, 2001b).

Stigma leads to poorer individual and family functioning: High percentages of individuals with mental disorders avoid seeking treatment, even though public awareness of problems related to mental illness has increased (Jorm, Christensen, & Griffiths, 2006). Individuals are frequently deterred by the potential stigmatization associated with a diagnosis, and suffer lack of access to responsive care even when treatment is sought because of the limited funding for treatment that is a direct result of mental illness stigma (Kessler et al., 1996; Regier et al., 1993). Negative effects of stigmatization are evident even when pre-existing symptomatology or functioning is controlled, meaning that stigma’s effects add to those emanating from mental illness itself (see reviews in Hinshaw, 2006; Markowitz, 1998). Stigma is real, with devastating consequences.

Because so many are affected and because the impact on individuals and families is potentially debilitating, mental illness is internationally recognized as one of the largest
and most destructive threats to health and positive adjustment (Corrigan & Miller, 2004; Murray & Lopez, 1996; World Health Organization, 2001a). Although public awareness of mental disorder, motivation for policy change, and attempts to reduce associated stigma have increased significantly over the past few decades, stigma against mental illness remains a significant barrier to positive outcomes across cultures and nations, related to the threat value of mental symptoms, intolerance for diversity, and inaccurate conceptions of mental disorder (Hinshaw & Cicchetti, 2000; Guimon, Fisher, & Sartorius, 1999). Furthermore, and discouragingly, even though knowledge about the nature and causes of mental illness has increased over the past decades in Westernized nations (Jorm et al., 2006) and the public believes that stigma against mental illness is decreasing (Angermeyer & Matschinger, 2005), stigmatization of the most severe forms of mental illness has shown a corresponding increase, rather than decrease, largely because of associations made by the general public (and promoted by the media) between mental illness and violence or danger (Link, Phelan, Bresnahan, Stueve, & Pescosolido, 1999; Phelan, Link, Stueve, & Pescosolido, 2000). It is therefore extremely important to understand the stigma against mental illness, elucidate the reasons for its existence, and reduce its negative effects.

*The Nature and Impact of Stigma*

Before beginning our review and analysis, we briefly distinguish several related constructs (see Hinshaw, 2006, for further development). Stereotypes are characterizations of social groups that may be based on some kernel of truth but are made in rigid, global fashion. Prejudice connotes castigating attitudes toward groups that are made without full knowledge of the group members in question. Discrimination entails
differential treatment of one group relative to another, often involving restrictions of
rights or life opportunities for group members. Thus, stereotypes are cognitive
phenomena, prejudice is affective in nature, and discrimination pertains to behavioral
acts. Stigmatization encompasses all of these processes but also involves the tendency
toward attributing any and all negative attributes of the “outgroup” member to his or her
membership in the castigated group, fueling a vicious cycle of societal rebuff and
personal internalization of the rejecting messages. Stigma processes have been the
subject of considerable interest with respect to empirical work and related theorizing in
social psychology over the past few decades (e.g., Bodenhausen & Macrae, 1998;
Crocker, Major, & Steele, 1998; Devine, 1989; Fiske, 1998; Hewstone, Ruben, & Willis,
2002; Link & Phelan, 2001; Major & O’Brien, 2005), with most work focusing on stigma
and prejudice directed toward racial minorities.

In short, stigma refers to a visible or concealable mark that is considered by the
majority of a given social group to indicate deviance or immorality. It also signifies the
social judgment and discrimination the majority places on outgroup members who
possess such a mark, as well as the great potential for internalization and shame on the
part of the individuals who are devalued and castigated (Goffman, 1963). In the domain
of mental illness, stigma centers on both the behavioral presentations characteristic of
mental illness and the label itself, featuring the prejudicial attitudes and discriminatory
practices that accompany society’s responses to a diagnosis.

Stigma is expressed toward a wide variety of domains and conditions, including
race, gender, and sexual orientation. In general, stigmatized individuals have been found
to face a variety of social and emotional consequences, including social withdrawal, loss

of productivity, lowered self-esteem, and increased levels of negative affect (Heatherton, Kleck, Hebl, & Hull, 2000). The applicability of broad theories of stigma to the mental illness domain remains an important topic (Hinshaw, 2006).

Stigma against mental illness is a crucial phenomenon, because it has persisted even as tolerance for other stigmatized groups has gradually grown. Indeed, individuals with mental illness are members of one of the few groups whose castigation remains socially acceptable. Use of derogatory language concerning such individuals is ubiquitous; for example, “retard,” “psycho,” and “crazy” are common slurs across cultures that both children and adults often use offhand (Hinshaw & Cicchetti, 2000). In addition, the media are a continuing source of inaccurate and unfavorable depictions of individuals with mental illness. Despite recent tendencies toward at least some improvements in media portrayals, the vast majority of relevant depictions emphasize violent and bizarre behavior and social incompetence (Wahl, 1992, 1995; see review in Hinshaw, 2006).

We realize that mental illness can certainly have seriously negative impact on productivity. In fact, it is conceivable that stigma could have some “realistic” basis. However, many individuals with mental illnesses, particularly those with access to treatment, have long periods of time when they are in good mental health. Basing hiring practices on potential future episodes of mental illness is discriminatory, analogous to employers’ avoidance of hiring women because female employees might become pregnant and be less productive. Consider also prior attitudes toward hiring persons with physical disabilities: Without accommodations, it could have been contended that their productivity was inherently low. In addition, because stigma itself contributes to poor
mental health outcomes—even controlling for baseline measures of mental disorder and impairment—it is important to reduce the conception that mental illness leads automatically to lowered productivity. Unfortunately, castigation and derogation remain pervasive in most if not all societies.

Previous research on stigma and mental illness has historically focused on two major topics: (a) bias, prejudice, and discrimination against individuals with mental illness by members of society, and (b) the internalization of such perspectives by individuals with mental illness—sometimes termed self-stigma. In terms of the former, a host of research indicates that persons with mental illness are perceived to be dangerous, violent, incompetent, and a drain on societal resources (e.g., Corrigan & Cooper, 2005; Corrigan, Edwards, Green, Diwan, & Penn, 2001; Wahl, 1999). Such views have been persistent across history and remain in place in multiple nations and cultures across the world (Hinshaw, 2006). Not only are public attitudes likely to be highly derogatory, but persons with mental illness face significant discrimination when their diagnosis is known to others. These consequences can be severe, including (a) difficulties finding or keeping a job or living quarters (e.g., Corrigan & Kleinlen, 2005; Page, 1995), (b) restrictions on fundamental social rights (e.g., voting, having child custody, driving; see Corrigan, Markowitz, & Watson, 2004), and (c) lack of access to continuing treatment because of the current lack of parity in health care—in other words, coverage for mental illness is inferior to that for physical illnesses (e.g. Thornicroft & Betts, 2002; Sartorius, 1998; U.S. General Accounting Office, 2000). Research on social distancing, as measured by stated attitudes regarding willingness to come into varying degrees of contact with individuals who have a mental illness, suggests that people try to avoid individuals with
mental illness across a number of situations, even those that require little direct contact (Corrigan et al., 2001). In short, evidence for the stigmatization of individuals with mental illness by the general public is wide-ranging.

With respect to self-stigma, research in the past 15 years strongly suggests that not all stigmatized individuals or groups necessarily show reductions in self-esteem (Crocker & Major, 1989). In other words, there may be a host of self-protective mechanisms utilized by “marked” individuals that serve to mitigate internalization of shame and development of low self worth. Nonetheless, individuals with mental illness may be at particular risk for highly negative personal impacts of stigma, for at least two major reasons.

First, whereas stigmas such as obesity or non-Caucasian ethnicity are readily apparent to observers and cannot be hidden, *concealable stigmas* appear to produce particular anxiety and stress (see initial work by Goffman, 1963, and Jones et al., 1984; see also Quinn, Kahng, & Crocker, 2004). In their attempts to avoid detection and discrimination, individuals who possess a concealable stigma may expend a large amount of energy to hide characteristics that might identify them as belonging to the stigmatized category—and such mental efforts may backfire at personal and social levels (Smart & Wegner, 1999; Quinn et al., 2004). Because a diagnosis of mental illness can often be hidden, individuals with mental illness constitute such a group. For example, an individual with major depression and suicidal ideation might attempt to hide these thoughts from coworkers and family by attempting to suppress his thoughts or withdrawing from others’ company. Unfortunately, thought suppression often has a rebound effect, meaning that the individual is likely to become even more preoccupied
with thoughts of suicide and less able to engage effectively with other people (Wegner, 1997). Thus, attempting to conceal one’s illness may come at the cost of intimacy, friendship, support, and likeability, further increasing the very symptoms the individual is trying to hide.

Second, the very symptoms of mental illness often include internalization of blame and tendencies toward depression and self-derogation as well as identity diffusion. As a result, individuals with mental illness are particularly prone to internalize the stigma that society places on them and may overestimate the probability of facing discrimination as a result of their illness (Link, Struening, Neese-Todd, Asmussen, & Phelan, 2001; Wahl, 1999). It is also the case that individuals tend to be perceived more negatively when they believe their diagnosis is known to others, whether or not this is actually the case (Corrigan & Watson, 2002; Farina, Allen, & Saul, 1968; Farina, Gliha, Boudreau, Allen, & Sherman, 1971; Kleck & Strenta, 1980). Thus, even low levels of stigma against mental illness are likely to cause significant social and emotional problems for individuals with a mental illness diagnosis. Further compounding these problems, stigmatization of mental illness is often quite strong, as mental disorders (along with substance abuse and homelessness) are consistently revealed to be the most despised attributes a person can possess (e.g., Tringo, 1970; see review in Hinshaw, 2006). Indeed, mental illness has been likened to the vast historical tendencies to banish and castigate conditions like leprosy (Rabkin, 1972). In all, several factors conspire to render self-stigmatization particularly likely for individuals with mental disorders.

The negative impact of the stigma of mental illness unfortunately extends beyond individuals with a diagnosis to their family members and close friends. Families perceive
stigma as a major issue, not only for their mentally ill family member but for themselves as well (Wahl, 1999; Wahl & Harman, 1989). Family members feel shame, self-blame, and mistrust; they must often cope with the objective burdens related to caregiving and the subjective burdens related to societal rejection and embarrassment over the relative’s behavior patterns, with a resultant negative impact on their own mental and physical well-being (Corrigan & Miller, 2004; Greenberg, Greenley, McKee, Brown, & Griffin-Francell, 1993; Lefley, 1989; Martens & Addington, 2001; Szmukler et al., 1996). Such attitudes are not surprising, given that the predominant professional orientation for much of the 20th century was to blame nearly all forms of mental illness on faulty parenting or caregiving (Hinshaw, 2005). Stigma thus prevents many individuals from ever seeking treatment because they and their families are ashamed of the existence of mental illness and concerned that they may face significant discrimination and prejudice from neighbors, friends, and even mental health providers if their diagnosis is known (Wahl, 1999).

Compounding this issue is that stigma against mental illness has led, at a broader level, to insufficient funding for research and treatment (Sartorius, 1998). The perception exists that such individuals can and should control their illnesses and, even worse, that mental disorders are stable, permanent conditions that are unresponsive to treatment. Overall, self-stigma and family stigma are fueled by societal attitudes and practices, discriminatory policies, tendencies toward concealment and silence, and structural processes that preclude the seeking or funding of adequate treatment (see Corrigan, 2005; Hinshaw, 2006; Holmes, Corrigan, Williams, Canar, & Kubiak, 1999). Furthermore, children and the elderly, as well as persons of low socioeconomic status—who are least
able to gain treatment for themselves because of their dependency on other family members or the inadequate health care system—appear to be particularly susceptible to the negative effects of stigma (Hinshaw, 2005).

Overall, a growing number of voices are recognizing the impact of stigma on individuals with mental illness, their family members, and communities and societies as a whole (Corrigan & Kleinlen, 2005; Hinshaw, 2006; Link & Phelan, 2001). Stigma toward mental illness has been rampant throughout history, suggesting universal, or even naturally selected, “exclusion modules” toward persons with mental disorders and the mental illness label itself (Kurzban & Leary, 2001; Link & Phelan, 2001). Indeed, stigmatization is not limited to Western societies but rather exists cross-culturally (Guimon et al., 1999). Although space does not permit a listing of the reasons for, or functions of, stigmatization for the social majority, the act of castigating others—particularly those who are believed to be irrational, out-of-control, and irresponsible—leads to increases in perceivers’ self-esteem, justifies societal inequalities that promote mental illness stigma and other forms of discrimination, and preserves a sense of order in response to the threats to social decorum and personal well-being posed by mental illness (see Solomon, Greenberg, & Pyszczynski, 2001; Fein & Spencer, 1997; Stangor & Crandall, 2000). Stigmatization is now recognized as perhaps the central issue facing all who are attempting to understand, prevent, and treat mental illness (Sartorius, 1998; U.S. Department of Health and Human Services, 1999).

**Mechanisms Underlying the Stigmatization of Mental Illness**

Stigmatization of individuals with mental illness has significant impact in modern society, though the reasons for stigma in this domain have generally not been rigorously
examined. Stigmatization of individuals with mental illness may result from many of the same processes that appear to underlie stigma against other groups: a relative lack of social contact, a tendency to categorize individuals unlike oneself into a different group and derogate that group to enhance the power of one’s ingroup, evolutionarily adaptive processes, self-esteem enhancement, and/or avoidance of situations or persons who increase mortality salience (Heatherton et al., 2000; Hinshaw, 2006). There are also indications, however, that phenomena related to mental illness stigma might look quite different from stigma processes related to other conditions.

For example, attribution theory, which suggests that stigmatized conditions or traits perceived as less controllable receive sympathy and support (Weiner, Perry, & Rasmussen, 1988), appears to apply in a different way to mental illness stigma than would be expected. A biological/genetic explanation for mental illness yields a perception of less controllability but also greater stability (and even permanence), which appears, paradoxically, to lead to increased stigma. In fact, individuals who interact with partners whose mental illness was said to have a biological cause (medical model), compared to individuals whose partners’ illness was said to have social cause (psychosocial model), expressed more positive attitudes in the form of decreased blame but more negative behavior in the form of increased punishment for mistakes (Mehta & Farina, 1997). This and related findings (e.g., Read & Harre, 2001; Read & Law, 1999) suggest that not only do attitudes and behaviors about mental illness have imperfect correspondence but that certain interventions leading to improved attitudes about mental illness may, ironically, increase punitive behaviors toward individuals with mental illness (Haslam, 2000; Hinshaw, 2006).
Conceptions of the severity of mental illness also appear to influence stigmatization. The psychiatric label given to individuals with mental illness and the length of time they have been ill are cues to mental illness severity, acting as a heuristic for degree of social disability (Mueller et al., 2006). Increased conception of social disability increases the degree to which individuals are stigmatized (Gaebel, Zäske, & Baumann, 2006). Several studies have shown that desire for social distance increases as labels become more severe, with attitudes generally more favorable, for example, toward depression than schizophrenia (Angermeyer & Matschinger, 2003). Nonetheless, stigma exists even against relatively mild forms of mental illness (Hinshaw, 2006), and an important step toward designing effective interventions against mental illness stigma in general is identifying which processes appear to be driving the stigma.

As noted earlier, several factors have been identified as possible mechanisms underlying stigma across key domains of interest; these include social cognition, contact, self-esteem enhancement, evolutionarily adaptive cognition, and cultural support (Dovidio et al., 2000). These factors play a role in stigma against mental illness, with some of them likely to have strong application. For example, terror management theory suggests that (a) increases in mortality salience increases stereotyping (Solomon et al., 1991) and (b) individuals or situations that increase mortality salience will be avoided. People with mental illness, particularly those with extreme behavioral deviance or physical dishevelment, may be more likely to increase such salience than, for example, individuals of a different race. If this is the case, mortality salience may mediate the strong relationship between mental illness status and consequent discrimination.
In addition to reducing mortality threats, using stereotypes can bolster self-esteem after esteem has been threatened by negative feedback (Fein & Spencer, 1997). Such self-esteem enhancement may be particularly relevant in the case of mental illness, as motivation to control prejudice against mental illness is typically low. In other words, given the continuing use of derogatory terms regarding mental illness, as well as the salience of the negative stereotypes surrounding this topic, such stereotypic views may be expressed with relatively little guilt or perceived social costs in order to promote a boost in self-esteem. Although high self-esteem is often seen as a positive attribute, individuals with high self-esteem are particularly likely to seek competency feedback via downward comparisons against outgroup members when they receive negative feedback that is ego-threatening (Vohs & Heatherton, 2001). Similarly, individuals with unstable high self-esteem (self esteem that fluctuates according to situational factors) are likely to derogate others after they receive an ego threat, in order to restore their self-esteem (Kernis, Cornell, Sun, & Berry, 1993). In short, blows to perceivers’ self-esteem are likely to increase derogation and punishment of individuals with mental illness, particularly by individuals initially high in self-esteem.

Although a number of mechanisms have been identified as underlying stigma in other domains, relatively little has been done with respect to mental illness stigma per se. Because these mechanisms are likely to affect the efficacy of interventions aimed at reducing stigma, it is critical to identify the processes driving stigma in this domain, including perceptions of the threat value of disturbed behaviors, automatic cognitions, emotional responses, attributions made regarding the behavior patterns in question, and
the like. Utilizing measures to address the processes that accompany stigma will greatly aid the attempt to comprehend the underpinnings of stigmatization.

Explicit vs. Implicit Measurement of Stigma

Although the foregoing picture is indeed discouraging, deeper levels of stigma may well exist, which resist simple evaluation and about which almost nothing is known in relation to mental illness. Specifically, research to date on the processing of social information—including, but not limited to, stigmatization—has indicated that individuals process such information on both an explicit level (i.e., consciously, controllably, and reflectively) and an implicit level (i.e., subconsciously, automatically, and intuitively). Key references on this topic include Bargh, Chaiken, Govender, and Pratto (1992), Devine (1989), Greenwald and Banaji (1995), and Greenwald and Farnham (2000).

Past work on prejudice and stereotyping suggests that self-reported, explicit measures of bias, prejudice, and stigma are problematic. In particular, explicit measures are subject to social desirability and often correlate poorly with alternative measures of stigma that focus on less consciously exhibited attitudes or on behavioral discrimination (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Greenwald & Banaji, 1995). In important research performed two decades ago, Link and colleagues found that typical questionnaire measures of attitudes toward mental illness were prone to socially desirable responding, such that participants reported more benign attitudes in response to portrayals of mental illness under the typical response formats used for such questionnaires than on questionnaires that assessed attitudes as attitudes as acted upon or more deeply held attitudes (Link & Cullen, 1983).
Tellingly, research on stigma in other domains has suggested that explicit measures of attitudes and behavior often fail to accurately assess underlying biases (e.g., Dovidio et al., 1997; Greenwald & Banaji, 1995). Given that it is often no longer socially acceptable to express prejudice overtly, individuals may have learned to avoid displaying such biases openly. In other words, even individuals who hold deeply-seated negative beliefs may present accepting attitudes on explicit measures. More subtle forms of prejudice have arisen as prejudice has become less acceptable. That is, overt racism appears to have been replaced, in many instances, by alternate forms. One is termed symbolic or “modern” racism—where negative beliefs are concealed on overt measures and instead revealed through alternative, less socially unacceptable beliefs such as favoring the Protestant work ethic or meritocracy. Another is called “aversive racism,” in which an individual consciously committed to non-stigmatizing views may still hold overlearned, unconscious, negative attitudes that “leak” when there are no cues related to egalitarian expression or when a factor other than race, such as ambiguous past job performance, is available to use as an “excuse” for discriminatory behavior such as racially-profiled hiring practices (Gaertner & Dovidio, 1986; McConahay, 1986). Despite the fact that derogating mental illness may be more common and generally acceptable than expressing racial bias, individuals are likely to still be motivated by positive self-presentation biases to present themselves in a positive light and appear more tolerant and caring than the average member of society (e.g. Rosnow, Goodstadt, Suls, & Gitter, 1973) and may show similar underreporting of mental illness stigma on traditional explicit attitudinal measures.

*Implicit Measures of Stigma*
A significant advance in research on stigma has been the development of measures that assess covertly expressed or implicit attitudes, defined as those that exist without the conscious knowledge of the respondent. Individuals are posited to have significantly less control over their responses on these measures, suggesting that implicit measures can more accurately assess underlying attitudes, particularly when these are socially unacceptable. One such measure that has been used successfully is the Implicit Association Test, or IAT (Greenwald & Banaji, 1995). On the IAT, respondents pair concepts with stimulus groups of interest; the outcome measure is reaction time, with shorter latencies believed to be an indication of stronger automatic association of the concept with the stimulus group. As an illustration, we use the example of the obesity IAT, which assesses implicit bias against obese individuals relative to thin people. The IAT contains four stimulus categories: fat people, thin people, positive adjectives, and negative adjectives. The respondent first classifies the items in each set of stimuli according to two relevant category options: She must decide whether each adjective is “good” or “bad” and each person is “fat” or “thin.” Respondents are then asked to classify each item according to pairs of options that cross-list the face and adjective categories, such as “fat OR good” and “thin OR bad”. The process is then repeated with the reverse pairing (i.e. “fat OR bad” and “thin OR good”). Implicit bias against individuals with obesity is revealed by reduced latency when classifying these individuals as “fat OR bad” than “fat OR good”. In fact, many individuals exhibit bias against those with obesity on this measure even when they show no bias against individuals with obesity on explicit attitude measures (Teachman, Gapinski, Brownell, Rawlins, & Jeyaram, 2003). Because respondents are asked to respond as quickly as possible and
because the variable of interest is reaction time, it is theorized that implicit measures tap response biases that occur below individuals’ usual level of awareness.

Support for this view comes from the finding that these implicitly assessed, unconscious biases show predictive validity to other indices of bias and discrimination, such as behavioral measures, often with a different pattern of correlations than those obtained from explicit scales (e.g., Hofmann, Gawronski, Gshwendner, Le, & Schmitt, 2005). For example, despite decreases in explicit racial bias, implicit measures show continued bias against non-Caucasian racial groups (e.g., Baron & Banaji, 2006).

Similarly, research on the stigma of obesity shows that low levels of explicit bias against obese individuals frequently co-exist with marked implicit biases and behavioral rejection in the same respondents (Teachman et al., 2003). Importantly, Teachman and colleagues noted that implicit attitudes predict prejudicial and discriminatory behavior toward individuals with obesity, whereas explicit attitudinal measures are uncorrelated with such behavior patterns.

It should be noted that the relationship between implicit measures of bias and behavior may be complex, with explicit measures better predicting intentional behaviors that are under conscious control, such as friendliness (e.g., Karpinski, Steinman, & Hilton, 2005; Shelton, Richeson, Salvatore, & Trawalter, 2005), whereas implicit measures are superior predictors of more automatic behavior, particularly when motivation to control prejudice is low (e.g., Bessenoff & Sherman, 2000; Neumann, Hulsenbeck, & Seibt, 2004; Ziegert & Hanges, 2005). Even in cases where implicit and explicit measures both indicate negative bias, they may represent distinct attitudinal components and differentially predict behavior toward the target (Teachman & Woody,
Thus, implicit measures appear to tap important processes that exist below the level of consciously controlled responses, predicting actual discrimination better than explicit measures.

The low correlations typically found between explicit and implicit measures prompted a recent meta-analysis of the relationship between the IAT and self-report measures, which noted a small effect size of $\rho = .24$ (Hofmann et al., 2005). This modest level of association strongly suggests that in key stigma domains, such as racial attitudes, explicit measures are often insufficient to assess actual levels of stigma—or at least that they need to be complemented by less direct assessment methods to reveal the full picture of bias. Determining whether explicit measures are similarly underestimating actual levels of stigma against mental illness is unknown but essential to understand.

Despite the limitations of explicit measures, research on mental illness to date has overwhelmingly neglected any alternative methods of assessing stigma. A notable exception appeared in the groundbreaking investigation of Graves, Cassisi, and Penn (2005), who measured psychophysiological responses to labels of mental illness. Specifically, when presented with an experimentally manipulated label of schizophrenia (as compared to no diagnosis), participants showed increases in brow muscle tension and heart rate deceleration, with the latter finding thought to be an indicator of preparation for action. Importantly, these increases in physiological reactivity predicted greater preferences for expressed social distance against individuals with the schizophrenia label. This investigation suggests that exposure to the label of a serious mental illness spurs automatic responses that are likely to influence subsequent behavior negatively; it also reveals the importance of including implicit, automatic indicators of bias and stigma.
In addition, Teachman, Wilson, and Komarovskaya’s (2006) work comparing explicit and implicit stigma of mental illness is the first published study on mental illness using an implicit measure and represents a crucial first step toward a deeper understanding of attitudes toward mental illness. In their analysis of explicit and implicit attitudes toward mental illness in individuals both with and without mental illness, Teachman and colleagues found that although explicit reports of attitudes toward mental illness were neutral, they were relatively more negative than explicit reports of attitudes toward physical illness. Furthermore, IAT assessments revealed negative implicit attitudes toward mental illness: 58-78% of participants associated the concepts “bad,” “blameworthy,” and “helpless” with mental illness, and there were stronger associations between these adjectives and mental illness than physical illness. Although similar patterns of bias against mental illness relative to physical illness appeared on the explicit and implicit measures, implicit measures showed additional absolute bias against mental illness. Echoing earlier research, the explicit and implicit measures were essentially uncorrelated, suggesting that they tapped independent constructs. This provocative first look at implicit measures of the stigma of mental illness suggests their substantial promise in furthering our understanding of stigma in this domain. In future work, an implicit measure such as the IAT or Go-No Go Task (GNAT, see Nosek and Banaji, 2001) could be used to replicate these preliminary findings in culturally and socioeconomically diverse samples or assess the degree to which an individual endorses other stereotypes with respect to mental illness, such as association with violence or incompetence.
Explicit attitudinal measures currently predominate at two levels: the stigmatization of individuals with mental illness by social observers, and internalized self-stigma in individuals with mental illness. Although these measures provide a valuable initial assessment of stigmatizing attitudes, it is quite likely that much of the literature on stigma and mental illness may actually underestimate the actual, less “censored” attitudes of social observers and of persons with mental illness themselves. Furthermore, let us assume, with real hope, that derogation of individuals with mental illness becomes less socially acceptable in the future. If this is the case, pressure to hide one’s underlying opinions about individuals with mental illness is likely to increase. Social desirability factors are also likely to apply to participants in intervention studies designed to reduce stigmatization, who may feel significant social pressure to suppress underlying bias or stigma following participation in the intervention. The ramifications could yield results that are quite misleading: Explicit measures of stigma applied after an intervention aimed at reducing stigma may yield supportive evidence for positive intervention effects that do not reflect actual levels of deep attitude change or reductions in discriminatory behaviors. The use of implicit measures is therefore crucial in evaluating the efficacy of such interventions.

*Behavioral Measures of Stigma*

Although a number of studies have assessed self-reported attitudes and behavioral tendencies toward individuals with mental illness, less research thus far has focused on using direct behavioral measures to assess stigma and discrimination against mental illness. The seminal investigation of Farina and Felner (1973) revealed that employers were significantly less likely to offer jobs to individuals believed to have a mental illness.
Similarly, Page (1977, 1995) found that landlords were significantly less likely to rent an apartment to an individual who disclosed a history of mental illness and mental hospitalization, with empirical findings holding across three decades. More recently, Corrigan and colleagues found that fear of dangerousness negatively predicted helping behavior toward individuals with mental illness (Corrigan et al., 2002). In short, measures of actual behavioral responses and discrimination have great potential for supplementing appraisals of overt attitudes—and for complementing implicit measures of bias as well.

Behavioral methods from social psychology research paradigms have the potential to further advance such valuable approaches in the mental illness domain. One relevant procedure was used by Macrae and Johnston (1998), who counted how many of the experimenter’s “accidentally” dropped pencils were picked up by the participant and used this measure to assess the effects of primed helpfulness. Bessenoff and Sherman (2000) chose a direct behavioral measure of social distance—that is, how far away the participant chose to sit from an overweight partner—to assess bias. Importantly, this behavioral measure was correlated significantly with an implicit measure of bias against overweight individuals, but not with an explicit measure of the same kind of bias. Helping and behavioral social distancing paradigms may be extremely applicable to research on stigma in the mental illness domain.

Furthermore, behavioral measures are considered the hallmark of sound social psychological research, which is important given the continuing patterns of behavioral discrimination against people with mental illness. Internal, implicit attitudes and biases can fuel consequent behaviors (Dovidio, Major, & Crocker, 2000). Although attitudes (even implicit ones) and behaviors are correlated imperfectly, behavioral measures of
Stigma may well be predicted more strongly from implicit measures of bias. In addition, social distance attitude scales would be productively supplemented by actual behavioral indicators of social distancing (e.g., how far away one chooses to sit from an individual labeled with mental illness), which are less subject to desires for positive self-presentation. As was the case with respect to overweight individuals in the investigation of Bessnoff and Sherman (2000), we anticipate that such behavior toward individuals with mental illness will be worse than explicitly stated attitudes would predict. This information would be extremely valuable in assessing both the true level of stigma against mental illness and the real-life impact of such stigma on individuals with a mental illness diagnosis.

Implications for Assessment and Intervention

Existing measures related to mental illness stigma have not yielded a “gold-standard” test for the efficacy of interventions. This is not to say that older, established explicit measures of attitudes about mental illness should be discarded; on the contrary, measures such as Corrigan’s Attributions About Mental Illness Questionnaire (Corrigan et al., 2002), various social distance scales (Link, Cullen, Frank, & Wozniak, 1987; Link et al., 1999; Link, Frank, Phelan, & Collins, 2004), and the Opinions about Mental Illness Scale (Cohen & Struening, 1962) have been used successfully in prior research, showing validity in predicting controllable behavior (e.g. Corrigan et al., 2002). In addition, measures of perceived stigma can reduce social desirability concerns by asking respondents about the “average person’s” beliefs about mental illness, rather than their own (Angermeyer & Matschinger, 2005; Chung & Wong, 2004; Griffiths, Christensen, Jorm, Evans, & Groves, 2004). However, responses on these measures are still under
conscious control and potentially influenced by individuals’ desire to appear unprejudiced.

In future investigations, explicit measures should be compared with newer implicit and behavioral measures to find the best estimate of current levels of stigma and examine the potentially telling divide between socially acceptable, overtly stated attitudes on the one hand and underlying, implicit attitudes and automatic behavior on the other. Use of multiple methods and assessments in the same investigation is a particularly important priority, toward the end of validating explicit, implicit, and behavioral strategies. Note that such research will need to be done with care: Utilizing explicit measures of mental illness stigma before the participant engages in implicit attitude testing or behavioral paradigms could prime responses that might not otherwise have been observed. Similarly, following implicit or behavioral measures with assessments of overt attitudes could taint the latter. Still, it will be important to understand the correspondence between various methods of appraising stigma. A likely outcome is that reliability and power will be gained when various methods are composited.

A multi-method analysis of current levels of mental illness stigma has important implications for the development and testing of interventions aimed at reducing such stigma. First, an accurate assessment of preexisting levels of stigma provides a standard of comparison for levels of stigma following an intervention. Second, individuals may have little conscious awareness of the processes underlying their stigmatization of individuals with mental illness, and be unable or unwilling to report their reasons for doing so. Implicit and behavioral measures thus show promise in elucidating potential mechanisms that underlie such stigma; such information about underlying processes can
be used to design more effective and efficient interventions. For example, it may be that certain types of exposure to (or contact with) individuals with mental illness are more effective than others in reducing unconscious bias, with implications for reduced discrimination; only implicit measurement strategies can uncover such mediator processes. Third, as we have emphasized earlier, implicit measures will allow for an accurate assessment of the efficacy of interventions, given their lowered susceptibility to social desirability factors.

Concerns for the Present and Future

Implicit and direct behavioral measures should allow a more accurate assessment of changes in stigma toward mental illness over the coming years. Explicit measures, which are subject to social desirability issues, may fluctuate across samples differing in sophistication, whereas implicit and behavioral measures are more likely to reveal deeper levels of bias in a more consistent manner. Crucially, accurate assessments of unconsciously expressed levels of stigma will allow us to verify whether stigma is improving as a result of interventions aimed toward reducing stigma or, alternatively, whether such interventions could potentially backfire.

Although we wholeheartedly believe that stigma against mental illness has an enormous negative impact on treatment and prognosis for individuals with mental illnesses, there is an ironic possibility that reducing stigma may increase the severity of mental illness symptomatology. For example, the acceptance of alcoholism as a “disease” may decrease perceived self-efficacy to control maladaptive drinking and increase drinking behavior. In addition, “pro-Ana” websites that support and lionize individuals with anorexia nervosa, promoting destructive behavior patterns as a “controlled lifestyle
choice” rather than a mental disorder, may increase the incidence and relapse of anorexia nervosa in young women. As stigma eventually begins to decrease, leading to increased support for individuals with mental illness, it will be important to assess both the potential benefits and potential drawbacks of such support with respect to illness severity and motivation for change.

In all, investigators and interventionists in future years would do well to pay the utmost attention to their means of appraising stigmatizing attitudes and discriminatory behaviors. Explicit measures are tried and true, and recent investigations utilizing such measures continue to show major stigmatization of mental illness (see review in Hinshaw, 2006). Yet it is quite conceivable that evaluation of less conscious and more implicit stigma processes will yield even stronger evidence of prejudice, bias, and stigma with respect to mental illness and will predict, more strongly than overt attitudes, actual behavioral rejection, distancing, and discrimination. There is far too much at stake with respect to the stigma that attends to mental illness—at the levels of individuals, family members, communities, nations, and the world population as a whole—to rely on measurement instruments that cannot ascertain the true nature of hidden, covert, and implicit biases. The future of research on stigma, and the future of endeavors to provide intervention to reduce stigmatization, relies heavily on the potential for integration of explicit, implicit, and behavioral approaches to the appraisal of stigma.
References


Crisp, R. (2000). A qualitative study of the perceptions of individuals with disabilities...


625-664.


Teachman, B., Wilson, J., & Komarovskaya, I. (2006). Implicit and explicit stigma of


