Measuring Mental Illness Stigma

by Bruce G. Link, Lawrence H. Yang, Jo C. Phelan, and Pamela Y. Collins

Abstract

The effectiveness of efforts designed to address mental illness stigma will rest on our ability to understand stigma processes, the factors that produce and sustain such processes, and the mechanisms that lead from stigmatization to harmful consequences. Critical to such an understanding is our capacity to observe and measure the essential components of stigma processes. This article is designed to assist researchers in selecting or creating measures that can address critical research questions regarding stigma. Our conceptualization of stigma processes leads us to consider components of labeling, stereotyping, cognitive separating, emotional reactions, status loss, and discrimination. We review 123 empirical articles published between January 1995 and June 2003 that have sought to assess mental illness stigma and use these articles to provide a profile of current measurement in this area. From the articles we identify commonly used and promising measures and describe those measures in more detail so that readers can decide whether the described measures might be appropriate for their studies. We end by identifying gaps in stigma measurement in terms of concepts covered and populations assessed.

Keywords: Stigma, discrimination, public attitudes, measurement, social rejection.


The Surgeon General's 1999 report on mental health notes that there is a strong consensus that "our society no longer can afford to view mental health as separate and unequal to general health" (p. vii, Executive Summary, U.S. Department of Health and Human Services 1999) and that stigma "deprives people of their dignity and interferes with their full participation in society" (p. 6, 1999). From the vantage point of the Surgeon General's report, if we are to improve mental health and mental health care, we must address stigma. In keeping with this idea, a diverse set of initiatives have emerged in the United States and around the world focused on efforts to combat stigma and discrimination. But the effectiveness of any such efforts will rest on our ability to understand stigma processes, the factors that produce and sustain such processes, and the mechanisms that lead from stigmatization to harmful consequences.

Essential to the scientific understanding of stigma is our capacity to observe and measure it. The central purpose of this article is to assist researchers studying the stigma of mental illness in selecting and creating the measures they need to conduct their work. Additionally, we seek to stimulate future research in neglected areas by identifying gaps in stigma measurement that need further attention. Specifically, we review 123 articles focused on the stigma of mental illness that were published between January 1995 and June 2003 and examine the measures employed in them. This review will inform readers about the breadth of methods employed in studying stigma, the kinds of study populations that have been assessed, the geographic locations of the study sites, and the coverage of stigma concepts that the investigations have achieved. The next step in our presentation is a detailed description of promising and frequently used measures identified in the review. Specifically, we describe the origin and content of selected measures, discuss evidence for reliability and validity, indicate particular strengths or shortcomings, and provide key citations regarding their use. Following this review of measures, we address three additional topics critical to the measurement of stigma: (1) the use of vignettes, (2) the use of behavioral measures in experimental and nonexperimental studies, and (3) qualitative assessment as a substitute for or complement to quantitative measures.
Conceptualizing Stigma

As a starting point for our review, we return to the stigma concept and in particular to three theoretical frameworks relevant to that concept. We do so as a basis for judging whether current measurement corresponds closely to theoretical concepts and covers the full range of stigma-relevant conceptualization. The conceptualizations we return to were all developed by considering not just mental illnesses but multiple circumstances in which stigma arises. In reviewing these conceptual frameworks, the reader should be aware that there is debate about the definition and the utility of the stigma concept and that none of the conceptualizations should be viewed as definitive. At the same time, we believe that close attention to these and other (Fine and Asch 1988; Crocker et al. 1998; Sayce 2003) conceptual considerations will help refine and expand measurement in this area.

Goffman. Goffman (1963) is widely cited for his insightful exposition of the stigma concept, and any current researcher planning to pose and answer research questions regarding stigma is well advised to return to it. The influence of Goffman’s work is evident in the fact that researchers draw on it with such frequency when seeking definitions of stigma. Although different definitions can be taken from Goffman, a very common one is that of an “attribute that is deeply discrediting” and that reduces the bearer “from a whole and usual person to a tainted, discounted one” ([Goffman 1963], p. 3). Elsewhere, Goffman defines stigma as the relationship between an “attribute and a stereotype” (p. 4). Many of Goffman’s ideas have been carried forward to more recent work on the conceptualization and measurement of stigma. His distinctions between the “discredited” and the “discreditable” are evident in subsequent conceptualizations of the “visibility” dimension of stigma (Jones et al. 1984) and in attention paid to “secrecy” as a potential coping mechanism for people whose stigmatizing circumstance is discreditable (Conrad and Schneider and Conrad 1980; Link et al. 1989). Similarly, Goffman’s concepts of the “own” (those similarly stigmatized) and the “wise” (those who know about and accept the stigma) are evident in efforts to measure withdrawal as a stigma coping mechanism (Link et al. 1989).

Jones and Colleagues’ Dimensions of Stigma. Following on Goffman’s insights, a second conceptual framework was developed by Jones et al. (1984) in their book Social Stigma: The Psychology of Marked Relationships. Jones et al. use the term “mark” as a descriptor that encompasses the range of conditions considered deviant by a society that might initiate the stigmatizing process. Stigma takes place when the mark links the identified person via attributional processes to undesirable characteristics that discredit him or her in the eyes of others.

Jones et al. proceed to identify six dimensions of stigma. Concealability indicates how obvious or detectable the characteristic is to others. Concealability varies depending on the nature of the stigmatizing mark such that those who are able to conceal their condition (e.g., people with mental illness) often do so. Course indicates whether the stigmatizing condition is reversible or, with irreversible conditions tending to elicit more negative attitudes from others. Disruptiveness indicates the extent to which a mark strains or obstructs interpersonal interactions. For example, interaction with people with mental illness is sometimes experienced as disruptive by others because of a fear of unexpected behavior by individuals with mental disorders. Aesthetics reflects what is attractive or pleasing to one’s perceptions; when related to stigma, this dimension concerns the extent to which a mark elicits an instinctive and affective reaction of disgust. Origin refers to how the condition came into being. In particular, perceived responsibility for the condition carries great influence in whether others will respond with unfavorable views and/or punishment toward the identified offender. The final dimension, peril, refers to feelings of danger or threat that the mark induces in others. Threat in this sense can either refer to a fear of actual physical danger (e.g., from a communicable disease such as leprosy) or exposure to uncomfortable feelings of vulnerability (e.g., uneasiness or guilt resulting from watching a disabled person negotiate a flight of stairs).

Link and Phelan’s Components of Stigma. Link and Phelan (2001) developed a conceptualization in response to criticisms that the stigma concept locates the “problem” in the individual and tends to focus on cognitive processing of information rather than on the discrimination and exclusion that a stigmatized person experiences (Sayce 1998). Link and Phelan construct a definition that links component concepts under a broad umbrella concept they call stigma:

In our conceptualization, stigma exists when the following interrelated components converge. In the first component, people distinguish and label human differences. In the second, dominant cultural beliefs link labeled persons to undesirable characteristics—to negative stereotypes. In the third, labeled persons are placed in distinct categories so as to accomplish some degree of separation of “us” from “them.” In
the fourth, labeled persons experience status loss and discrimination that lead to unequal outcomes. Stigmatization is entirely contingent on access to social, economic and political power that allows the identification of differentness, the construction of stereotypes, the separation of labeled persons into distinct categories and the full execution of disapproval, rejection, exclusion and discrimination. Thus we apply the term stigma when elements of labeling, stereotyping, separation, status loss and discrimination co-occur in a power situation that allows them to unfold. (Link and Phelan 2001, p. 367)

Like Jones et al.'s dimensions, Link and Phelan's stigma components are useful in identifying the domain of content that stigma measures should include. Although a more detailed explication of these components is available elsewhere (Link and Phelan 2001), we attend to them briefly here. In addition, we expand the conceptualization to include a component for emotional responses. It should be noted that each component is a matter of degree—labels can be more or less socially salient, they can link to many stereotypes or just a few, the linking can be relatively strong or relatively weak, people can be more or less set apart, emotional responses can vary from overwhelming to relatively trivial, and the extent of status loss and discrimination can be relatively great or relatively small. Thus, stigma is a “matter of degree” and will vary across stigmatizing circumstances such as schizophrenia, depression, obesity, HIV status, short stature, diabetes, and cancer.

Labeling. The vast majority of human differences, (e.g., finger length, preferred vegetables) are not considered to be socially relevant. However, some differences, such as skin color and sexual preferences, are currently awarded a high degree of social salience. Both the selection of salient characteristics and the creation of labels for them are social achievements that need to be understood as essential components of stigma.

Stereotyping. In this component, the labeled difference is linked to undesirable characteristics either in the minds of other persons or the labeled person him- or herself. For example, it may be assumed that a person who has been hospitalized for mental illness represents a violence risk.

Separating. A third aspect of the stigma process occurs when social labels connote a separation of “us” from “them.” For example, members of certain ethnic or national groups (Morone 1997), people with mental illness, or people with a different sexual orientation may be considered fundamentally different kinds of people from “us.” When this separation is particularly thorough, members of a stigmatized group may accept stereotypes about themselves and view themselves as fundamentally different from and inferior to other people.

Emotional reactions. Underrepresented in the Link and Phelan (2001) formulations of stigma are the emotional responses it entails. We believe that this underrepresentation needs to be corrected, because emotional responses are critical to understanding the behavior of both stigmatizers and people who are recipients of stigmatizing reactions.

From the vantage point of a stigmatizer, the identification of human differences, the linking of those differences to undesirable attributes, and the separation of the identified person into a separate “them” group is likely to be associated with emotions of anger, irritation, anxiety, pity, and fear. The emotions engendered are likely to be important for at least two reasons. First, an emotional response is something that can be detected by the person who is stigmatized, thereby providing an important statement about a stigmatizer’s response to them as a person. For example, a person who feels some combination of pity and anxiety in the presence of a person with mental illness might modulate his or her voice, speaking softly and in an unnaturally calm tone, signaling to the person with mental illness that he or she is being approached from a standpoint of difference. Second, emotional responses may shape subsequent behavior toward the stigmatized person or group through processes identified by attribution theory (Weiner 1986). In this formulation, attributions about why a negative event occurred, especially the stigmatized person’s perceived control over the cause, has great influence over how others will think of and behave toward that person.

From the vantage point of the person who is stigmatized, emotions of embarrassment, shame, fear, alienation, or anger are possible. Thomas Scheff (1998) has, for example, argued that the emotion of shame is central to stigma and that shaming processes can have powerful and hurtful consequences for stigmatized persons. For all of these reasons, we believe that emotional responses and reactions need to be included in the broad conceptualization of stigma.

Status loss and discrimination. When people are labeled, set apart, and linked to undesirable characteristics, a rationale is constructed for devaluing, rejecting, and excluding them. This occurs in several ways. The most obvious is individual discrimination—for example, when a person rejects a job application or refuses to rent an apartment to a person with a mental illness. However, there are also more subtle mechanisms through which labeling and stereotyping lead to negative outcomes. One of these is structural discrimination, in which institutional practices work to the disadvantage of stigmatized groups, even in the absence of purposeful discrimination by individuals. For example, schizophrenia receives low levels of funding for research and treatment relative to other illnesses, and treatment facilities for schizophrenia tend to
be located in isolated settings or disadvantaged neighborhoods. Finally, once cultural stereotypes are in place, they can also have negative consequences that operate through the stigmatized person him- or herself via processes specified in modified labeling theory (Link et al. 1989), stereotype threat (Steele 1997), and stigma consciousness (Pinel 1999). Thus, there are a variety of ways—some working through nonlabeled individuals, some working through labeled individuals, and some working through societal institutions, some direct and obvious and others not—through which labeling, stereotyping, and separating result in poor life outcomes for stigmatized persons.

Dependence of stigma on power. A unique feature of Link and Phelan’s (2001) conceptualization is the idea that stigma is entirely dependent on social, economic, and political power. Groups with less power (e.g., psychiatric patients) may label, stereotype, and cognitively separate themselves from groups with more power (e.g., psychiatrists). But in these cases, stigma as Link and Phelan define it does not exist, because the potentially stigmatizing groups do not have the social, cultural, economic, and political power to imbue their cognitions with serious discriminatory consequences. Without a reference to power differences, stigma becomes a much broader and less useful concept that might be applied to lawyers, politicians, Wall Street traders, and white people.

Characterization of Current Measurement

In this article we evaluate current measurement practices in research on mental illness stigma based on a search (MEDLINE and PSYCHLIT) completed in July 2003. In both search engines, the search word “stigma” included the key terms of “prejudice,” “stereotyping,” “public opinion,” “attitude to health,” or “attitude.” Likewise, the search words “mental disorder” included the key terms of “diagnosis,” “drug therapy,” “psychology,” “education,” “epidemiology,” “etiology,” “genetics,” “therapy,” and “history of.” The search spanned from January 1995 to June 2003. In MEDLINE, 523 articles were identified with both search terms in the title or abstract; PSYCHLIT yielded an additional 161 articles. To be included in the data base, an article had (1) to be about an identifiable aspect of stigma that pertained specifically to psychiatric disorders, (2) to be an empirical study or literature review, and (3) to have an English text and abstract. A total of 95 empirical studies and 13 literature reviews were identified. Examining the reference lists of identified articles and conferencing with researchers in the field yielded an additional 14 empirical studies and 1 literature review. Thus, in total, our review included 109 empirical studies and 14 literature reviews. We do not claim that this review is exhaustive; one particular weakness of our sampling strategy is that we may have missed qualitative studies published in books. Nonetheless, the data set represents a broad assessment of current stigma measures in use.

Methodology Employed in the Reviewed Studies. Table 1 shows the number and overall percentage of each of several design types employed in the 109 studies: (1) survey (nonexperimental), (2) experiment, (3) qualitative, and (4) literature review. Surveys and experiments were further classified as “with vignette” or “without vignette,” and qualitative studies were further classified as “with content analysis” or “with interviews or participant observation.” As the table shows, most of the research involved nonexperimental survey research (60.1%), with another sizable component involving survey research with a vignette component (7.3%). Qualitative research in this review of journals was relatively rare (13.8%), and nonvignette experiments or quasi-experiments were almost nonexistent (0.8%).

Study Population in Reviewed Studies. Studies were coded as to the type of research participant included in them. They were deemed general population studies if they employed samples of college students and/or surveys of the general population of adults. Children and adolescents were subjects 17 years old or younger, whereas professional groups consisted of either health providers (e.g., mental health professionals, general practitioners, medical students) or other professional groups who interact with people with mental illness in an official capacity (e.g., police officers, case managers). People with psychiatric illness were individuals who had been diagnosed with mental illness or who exhibited psychiatric symptoms but who have not been formally diagnosed, and families of people with mental illness were individuals related by blood or marriage to people with psychiatric illnesses.

Of the 109 empirical studies, the most common type of study population was some segment of the general population (n = 51; 46.8%), followed by people with mental illnesses (n = 24; 22.0%), professional groups (n = 20; 18.3%), and family members of people with mental illnesses (n = 11; 10.1%). Of interest, only 4 studies (3.7%) assessed stigma in samples of children and adolescents—a striking fact given the relevance of this group to understanding the development of attitudes, beliefs, and behaviors.
Table 1. Summary of methodologies

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey (nonexperimental)</td>
<td>One or more standardized questions were asked of each respondent. Responses could be either forced choice or open-ended. There was no manipulation of the independent variable by the investigator.</td>
<td>74</td>
<td>60.1</td>
</tr>
<tr>
<td>With vignette</td>
<td>Survey (nonexperimental) with use of a vignette depicting someone with mental illness.</td>
<td>9</td>
<td>7.3</td>
</tr>
<tr>
<td>Without vignette</td>
<td>Survey (nonexperimental) without use of a vignette.</td>
<td>65</td>
<td>52.8</td>
</tr>
<tr>
<td>Experiment</td>
<td>The investigators explicitly manipulate an independent variable and randomly assign the manipulated variable to study groups.</td>
<td>20</td>
<td>16.2</td>
</tr>
<tr>
<td>With vignette</td>
<td>Experiment with use of a vignette depicting someone with mental illness.</td>
<td>19</td>
<td>15.4</td>
</tr>
<tr>
<td>Without vignette</td>
<td>Experiment without use of a vignette.</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Qualitative</td>
<td>The investigators used any one of a broad band of qualitative methodologies (e.g., participant observation, focus groups, life history interviews) in the study's design and data analysis.</td>
<td>17</td>
<td>13.8</td>
</tr>
<tr>
<td>With content analysis</td>
<td>Qualitative study with the use of source material such as newspapers, books, or television programs as the primary data for analysis.</td>
<td>3</td>
<td>2.4</td>
</tr>
<tr>
<td>With interviews or participant observation</td>
<td>Qualitative study with the use of interview or participant observation as the main study design.</td>
<td>14</td>
<td>11.4</td>
</tr>
<tr>
<td>Literature review</td>
<td>The article was not an empirical study but a summary of stigma research.</td>
<td>14</td>
<td>11.4</td>
</tr>
</tbody>
</table>

Note.—Total n = 123; percentages do not add up to 100 percent because of use of more than one methodology by studies.

Location of Reviewed Studies. Studies were also coded to reflect the country in which they were conducted or, in the case of literature reviews, written. The most common location was North America (United States or Canada; n = 62 or 50.4%). The next most frequent locations were Europe (United Kingdom, Germany, Greece, Ireland, Italy, Austria, and Sweden; n = 31 or 25.2%), followed by Asia (Hong Kong, Singapore, Japan, China, and India; n = 12 or 9.7%), and Eurasia (Australia and New Zealand; n = 12 or 9.7%). Relatively few studies or literature reviews were carried out in the Middle East (Israel and Turkey; n = 4 or 3.2%) or Africa (Ethiopia; n = 2 or 1.6%). One reason for this could be that we restricted our review to English language journals. Even so, given the relevance of cross-cultural research to understanding stigma processes, information from Asia and especially Africa seems to be dramatically underrepresented.

Stigma Components Assessed in Reviewed Studies. The studies were categorized by the stigma components identified by Link and Phelan (2001). Behavior was coded if the study introduced the actual behaviors indicative of mental illness (e.g., a vignette of a person who heard voices and had delusional thinking) as a stimulus to gauge the subjects’ attitudes or behaviors. Note that the construct of behavior is not a component of stigma according to Link and Phelan, but rather a stimulus that may elicit stigma processes such as labeling and stereotyping. The stigma components of labeling, stereotyping, cognitive separating, emotional reactions, status loss/discrimination (expectations), status loss/discrimination (experiences), structural discrimination, and behavioral responses to stigma (described above) are briefly defined in table 2.

Studies were identified as having measured a stigma component if one or more of the questions employed addressed it. As table 2 shows, with one exception, each of the stigma components was assessed by more than 10 percent of the studies. We find this instructive, as it underscores the need for a multicomponent conceptualization of stigma. The single exception is “structural discrimination,” which was coded as present in only two studies. This may be because issues of structural discrimination are discussed outside of the framework of stigma and therefore did not make it into our review. Nevertheless, we see the underrepresentation of this aspect as a dramatic shortcoming in the lit-
Table 2. Summary of stigma components

<table>
<thead>
<tr>
<th>Construct/stigma component</th>
<th>Definition</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior</td>
<td>The study introduces the actual behaviors indicative of the presence of mental illness as a stimulus.</td>
<td>28</td>
<td>25.7</td>
</tr>
<tr>
<td>Labeling</td>
<td>The study includes the assigning of social significance to particular characteristics as a variable of study.</td>
<td>20</td>
<td>18.3</td>
</tr>
<tr>
<td>Stereotyping</td>
<td>The study incorporates how labeled differences are linked with negative attributes.</td>
<td>68</td>
<td>62.4</td>
</tr>
<tr>
<td>Cognitive separating</td>
<td>The study measures when social labels imply a fundamental difference (&quot;them&quot;) compared with those without the label (&quot;us&quot;).</td>
<td>18</td>
<td>16.5</td>
</tr>
<tr>
<td>Emotional reactions</td>
<td>The study measures either the affective reactions of the stigmatizer toward people with mental illness or the emotional response of the stigmatized people themselves.</td>
<td>27</td>
<td>24.8</td>
</tr>
<tr>
<td>Status loss/discrimination (expectations)</td>
<td>The study includes expectations or beliefs of how persons with mental illness are reduced in social status or face discriminatory treatment from others.</td>
<td>64</td>
<td>58.7</td>
</tr>
<tr>
<td>Status loss/discrimination (experiences)</td>
<td>The study measures actual experiences of how persons with mental illness are reduced in social status or face discriminatory treatment from others.</td>
<td>15</td>
<td>13.8</td>
</tr>
<tr>
<td>Structural discrimination</td>
<td>The study assesses how institutional practices disadvantage persons with mental illness.</td>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>Behavioral responses to stigma</td>
<td>The study measures how individuals with mental illness act in response to societal discrimination, such as utilizing coping or avoidance strategies.</td>
<td>17</td>
<td>15.6</td>
</tr>
</tbody>
</table>

Note.—Total n = 109 (excluding literature reviews); percentages do not add up to 100% because of use of more than one stigma construct by investigations.

The literature on stigma, as the processes involved are likely major contributors to unequal outcomes for people with mental illnesses (see Corrigan et al., this issue, for an in-depth discussion of structural discrimination). The table is also revealing in that, as expected, measurement of stereotyping (62.4%) and expectations of status loss and discrimination (58.7%) are much more common than are experiences of status loss and discrimination (13.8%) and behavioral responses (15.6%) to stigma. Given that the conceptualizations of Link and Phelan (2001) underemphasized emotional responses, we were pleasantly surprised to find that nearly a quarter of the studies assessed this component nevertheless.

Methodology of Studies Grouped by Stigma Component Assessed. Tables 3 and 4 are provided to help readers find examples that show how stigma measures have been used in the literature. The numbers in the tables refer to citations in the numbered reference list. In table 3, articles are cross-classified by the methodology used in the study and by the Link and Phelan stigma component that is assessed. The table allows interested readers to find articles that have assessed aspects of stigma that particularly interest them or that use methodologies they are considering for future studies. For example, an article included in the cell for “Survey—vignette”/“Stereotyping” used a nonexperimental survey with a vignette to assess whether and to what extent labeled differences are linked to negative characteristics.

Study Population Grouped by Stigma Component. Table 4 classifies the articles by study population and stigma component. For example, an article included in the cell for “Families of people with mental illness”/“status loss/discrimination (experiences)” discusses a study conducted with families of individuals with mental illness that assesses actual experiences of status loss or discriminatory treatment.

Some Orienting Questions To Be Asked When Selecting Measures

Among the questions that need to be asked when selecting appropriate measures are the following:
### Table 3. Methodology grouped by stigma component

<table>
<thead>
<tr>
<th>Stigma component</th>
<th>Method</th>
<th>Survey—vignette</th>
<th>Survey—no vignette</th>
<th>Experimental—vignette</th>
<th>Experimental—no vignette</th>
<th>Qualitative—content analysis</th>
<th>Qualitative—interviews or participant observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior</td>
<td>2, 7, 52, 86, 95, 101, 111</td>
<td>67, 104, 114</td>
<td>3, 5, 6, 14, 20, 21, 23, 28, 49, 51, 53, 54, 57, 66, 77, 80, 99, 100</td>
<td>86, 101, 104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labeling</td>
<td>86, 101, 111</td>
<td>42, 78</td>
<td>1, 5, 6, 14, 21, 23, 49, 53, 57, 77, 100</td>
<td>98</td>
<td>32, 70, 86, 101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive separating</td>
<td>36, 86</td>
<td>25, 47, 61, 68, 72, 74, 76, 87, 92, 102, 104, 114, 120, 121</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stereotyping</td>
<td>2, 12, 36, 52, 86, 95, 101, 111</td>
<td>10, 18, 19, 22, 25, 26, 29, 31, 38, 42, 47, 61, 62, 67, 68, 71, 72, 74, 76, 78, 83–85, 88, 89, 91, 92, 94, 96, 102–104, 109, 110, 116, 120, 121, 123</td>
<td>1, 5, 6, 14, 20, 23, 28, 53, 54, 57, 77, 80, 99, 100</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status loss/ discrimination (experiences)</td>
<td>101</td>
<td>11, 33, 34, 59, 62, 63, 65, 75, 85, 113, 115, 122</td>
<td>11, 70, 101, 113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral responses to stigma</td>
<td>12</td>
<td>11, 13, 15, 40, 58, 59, 79, 82, 85, 88, 103, 106, 115, 122</td>
<td>11, 32, 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note.—Numbers represent identification numbers for stigma articles found in reference section from 1995 to 2003.
### Table 4. Study populations grouped by stigma component

<table>
<thead>
<tr>
<th>Stigma component</th>
<th>Study Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Children and adolescents</td>
</tr>
<tr>
<td>Behavior</td>
<td>86, 101</td>
</tr>
<tr>
<td>Labeling</td>
<td>1, 86, 101</td>
</tr>
<tr>
<td>Cognitive separating</td>
<td>74, 86</td>
</tr>
<tr>
<td>Stereotyping</td>
<td>1, 74, 86, 101</td>
</tr>
<tr>
<td>Emotional reactions</td>
<td>3, 8, 10, 26, 28, 31, 51, 78, 89, 108, 110, 114, 121</td>
</tr>
<tr>
<td>Status loss/discrimination (expectations)</td>
<td>74, 86</td>
</tr>
<tr>
<td>Status loss/discrimination (experiences)</td>
<td>101</td>
</tr>
<tr>
<td>Behavioral responses to stigma</td>
<td>12</td>
</tr>
</tbody>
</table>

**Note.**—Numbers represent identification numbers for stigma articles found in reference section from 1995 to 2003.

1. What is the research question regarding stigma, and what are the variables one must measure to answer the question posed?
2. Is an appropriate measure currently available? If not, can an existing measure be modified to suit research needs?
3. Is a candidate measure appropriate for the population I wish to study? Is the measure likely to be suitable to the social circumstances, culture, age group, and so forth of the population? Are the words and phrases used to refer to or describe people with mental illnesses appropriately sensi-
tive and respectful? If not, can the measure be modified to make it appropriate?

4. Is the candidate measure suitable to the methods to be used? Can it be administered by phone, by paper and pencil, within the context of a vignette, in an experimental context, and so on?

5. What is the evidence regarding the measure’s reliability and validity? Is the measure likely to be valid for its intended use? What is the evidence concerning contamination of the measure by biases such as social desirability bias?

6. How feasible is the measurement task in comparison to what might be required for other measures? How much time will its administration require, and how much will it cost?

The following review of commonly used measures will help investigators begin to answer these questions. The review is organized by considering measures applicable to (1) adult members of the general public, (2) children from the general public, (3) consumers of mental health services, and (4) family members of consumers. Following descriptions of measures that fall into these categories, we also include separate sections focused on (1) the use of vignettes, (2) measures used in laboratory experiments, and (3) qualitative approaches in assessing the stigma of mental illness.

Measures Applicable To Adult Members of the General Population

Social Distance. One of the most commonly used measures we encountered was that of social distance. Measures of social distance seek to assess a respondent’s willingness to interact with a target person in different types of relationships. Measures frequently include items that differ in the closeness of the association a respondent is asked to accept or decline.

The concept was initially defined by one of the leaders of the Chicago School, Robert Park (1924), as “the grades and degrees of understanding and intimacy which characterize pre-social and social relations generally” (p. 339). The first social distance scale (Bogardus 1925) was used to describe social distance by race/ethnicity. The scale’s first published use in the area of mental illnesses came in the context of Cumming and Cumming’s (1957) classic effort to change public attitudes in a town in Saskatchewan, Canada. Shortly thereafter, Whatley (1959) administered an eight-item agree-disagree social-distance scale to 2,001 persons in 17 parishes (counties) in Louisiana. Much like its original use in the area of race/ethnicity, Cumming and Cumming and Whatley used the scale to assess attitudes toward people with mental illness as a group. To our knowledge, Phillips (1963) was the first to employ a social distance scale in the context of a vignette experiment designed to assess responses to a described individual. Since that time, variants of the scale have been used with great frequency in research on stigma but particularly in the context of studies employing vignettes.

Social distance scales tend to show good to excellent internal-consistency reliability ranging from 0.75 to greater than 0.90. Evidence regarding the validity of social distance scales is available mainly with regard to construct validity (Cronbach and Meehl 1955). Most studies show that older persons, persons with relatively low educational levels, and persons who have never known anyone with a mental illness are more likely to desire social distance than their younger, more educated counterparts who have had more extensive contact with people with mental illnesses. Moreover, individuals who believe that people with mental illnesses are dangerous are much more likely to desire social distance from a person who is described as having a mental illness (Link et al. 1987, 1999; Martin et al. 2000). These patterns of association would fit expectations of results if social distance scales measured what they claim to measure.

There are two main limitations to the validity of social distance scales. The first of these is social desirability bias. Over the years, public education and anti-stigma campaigns have made it clear to the public that rejecting a person simply because he or she has seen a psychiatrist or been treated in a mental hospital is unenlightened and wrong. Not wanting to appear heartless or ignorant, people might deny social distancing responses in order to appear enlightened and caring. To the extent that social desirability bias is operative, assessments of reported social distance underestimate the extent of true social distancing sentiments. Furthermore, to the extent that measures of social distancing responses are contaminated by social desirability bias, one cannot know whether predictors of reported social distance are due to willingness to report social distance or to true social distancing responses.

Another limitation arises if one seeks to infer behavioral responses from reported intentions. Although behavioral intentions like social distance items are often good predictors of behavior, situational circumstances, competing attitudes, and other such factors can intervene so as to make the association far from perfect. If the target concept is in fact a specific behavior or a set of specific behaviors, then the relevant behaviors need to be measured directly.

Semantic Differential and Related Measures. The semantic differential is a measurement technique that pro-
vides a direct assessment of stereotyping—that is, the tendency to link a label like “mental patient” with negative attributes. It was developed by Charles E. Osgood et al. (1957) as a general technique for measuring the psychological meaning that concepts have for people. Jum Nunnally (Nunnally and Kittross 1958; Nunnally 1961) and later Olmsted and Durham (1976) and Crisp et al. (2000) applied the semantic differential to the question of public conceptions of people with mental illnesses and the professionals who treat them. The semantic differential presents respondents with labels, or concepts (e.g., “mental patient”), and asks them to evaluate the extent to which those labels are associated with various characteristics. Specifically, respondents are asked to rate the concept on a number of 7-point rating scales, each bounded by a pair of polar adjectives such as “insincere—sincere” and “cold—warm.” In addition to the concepts of interest (e.g., “mental patient”), respondents rate one or more additional concepts (e.g., “average person” or “me”) using the identical response scales to provide a context or point of comparison for evaluations of the target concepts. See Nunnally (1961), Olmsted and Durham (1976), and Crisp et al. (2000) for the stimulus concepts and adjective pairs they employed. Nunnally provides instructions and sample pages from his instrument.

The semantic differential has several features to recommend it. First, it provides a direct measure of stereotyping. Regarding reliability, Olmsted and Durham (1976) found very high correlations among ratings on their 14 adjective pairs (r = 0.95 to 0.99 depending on the stimulus concept). Regarding validity, although neither Nunnally (1961) nor Olmsted and Durham (1976) reported the statistical significance of the differences, evidence for construct validity is provided by the finding that concepts with mental illness labels (e.g., neurotic man, insane woman, mental patient) were rated substantially more negatively (e.g., more worthless, dirty, dangerous, cold, unpredictable, and insincere) than those without such labels. Furthermore, Nunnally found that a “neurotic” individual was rated less negatively on the evaluative dimension but also as less potent than an “insane” individual or a “mental patient,” and Olmsted and Durham found that an “ex-mental patient” was evaluated more positively than a “mental patient.” Similar differences in ratings applied to mentally ill and non–mentally ill persons were found by Crisp et al. (2000). Finally, because the semantic differential is a measurement approach rather than a specific measure, it allows for flexibility in applying it to the concepts and evaluative dimensions of interest to the researcher. If researchers want to maintain comparability to other studies, however, they should replicate at least some of the previously used adjective pairs. We caution, though, that some of the descriptive phrases used in the original applications are dated (e.g., “mental patient,” “insane person”), and researchers should consider their appropriateness before adopting them.

A shortcoming of this measurement approach, shared by many of those we review, is its vulnerability to social-desirability bias. However, the consistent differences observed between ratings of mental illness–related concepts and other concepts suggests that social desirability bias is not extreme. This may be in part because, when respondents are asked to rate concepts such as “average woman” or “average man,” they are encouraged to make implicit comparisons to the mental illness–related concepts that reduce the tendency to rate the mental illness concepts positively. That is, respondents may experience the following thought process: “If I’m going to give ‘average woman’ a ‘6’ on ‘warm/cold,’ I can’t give ‘mental patient’ more than a ‘4.’”

**Opinions About Mental Illness.** This scale was developed in the early 1960s by Cohen and Struening (1962; Struening and Cohen 1963) and has been used extensively ever since. Cohen and Struening (1962) sought the “adequate conception and objective measurement of attitudes toward mental illness” (p. 349) through a multidimensional scale. The Opinions About Mental Illness Scale (OMI) was developed in two large psychiatric hospitals using the responses of 1,194 hospital workers. By creating new items and by including items from the Custodial Mental Illness Ideology Scale, the California F-Scale, and Nunnally’s work on popular conceptions of mental health, Cohen and Struening constructed a 70-item instrument. When items were factor-analyzed, five dimensions were identified: (1) authoritarianism—that obedience to authority is critical and that people with mental illness are an inferior class requiring coercive handling; (2) benevolence—a kindly, paternalistic view of people with mental illnesses supported by humanism and religion rather than science; (3) mental hygiene ideology—the idea that mental illness is an illness like any other and that a rational, professional approach to people with mental illnesses is crucial for adequate treatment; (4) social restrictiveness—that the activities of people with mental illnesses should be restricted in domains such as marriage, voting, childbearing, jobs, and parenting; and (5) interpersonal etiology—the idea that mental illnesses arise from interpersonal experiences, particularly the lack of a loving home environment. In a subsequent paper (Struening and Cohen 1963), the original 70 items were reduced to 51 by retaining only items with mental illness content. Evidence suggesting similar fac-
...structures among personnel of three newly sampled hospitals was obtained, and explicit scoring procedures for the five factors were presented. One of the factors, mental hygiene ideology, showed less than adequate internal consistency in the three hospitals, with coefficients ranging from 0.29 to 0.39. However, authoritarianism (0.77–0.80), benevolence (0.70–0.73), restrictiveness (0.71–0.77), and interpersonal etiology (0.65–0.66) had adequate internal consistency. It is this 51-item version of the OMI that has been and continues to be used in research today (Struening and Cohen 1963).

The OMI has several important strengths that have undoubtedly contributed to its extensive use for such a long time and in so many different cultural settings. One such strength becomes apparent when its items are closely scrutinized and compared with items in other scales in this area of research. The items tend to have a poignancy and complexity aimed at supplying a stimulus that affects the respondent and provides something potent to react to. For example, one item reads, “Even though patients in mental hospitals behave in funny ways, it is wrong to laugh about them,” and another, “All patients in mental hospitals should be prevented from having children by a painless operation.”

A second advantage of the OMI is its breadth of coverage of salient issues. It includes items assessing many of the components that Jones et al. and Link and Phelan identify. The linking of labels (mental hospital patient) to stereotypes is prominent, for example, in items such as “People who are mentally ill let their emotions control them: normal people think things out” and “People who were once patients in mental hospitals are no more dangerous than the average citizen.” The notion of separation into “us” and “them” is evident in items such as “A heart patient has just one thing wrong with him, while a mentally ill person is completely different from other people” and “There is something about mental patients that makes it easy to tell them from normal people.” In addition, status loss (“To become a patient in a mental hospital is to become a failure in life”) and inclination to discriminate (“Anyone who is in a hospital for mental illness should not be allowed to vote”) are also prominent in the OMI, as are items reflecting Jones et al.’s (1984) dimensions of origins, course, peril, and disruptiveness.

A third, as yet unrealized, advantage of this measure is its long history and thus the possibility of assessing changes in attitudes over time. Although it was never part of a nationwide study, the OMI has been used in so many populations for so long that useful comparisons could be made that would allow an opportunity to assess whether and to what extent attitudes have changed. A disadvantage of the OMI is that new issues have arisen since it was developed. The social policy of deinstitutionalization and the increased salience of genetic factors in the etiology of mental illnesses are examples of domains that might be represented in a new formulation of the OMI.

Community Attitudes Toward the Mentally Ill. Partly because the OMI did not cover issues of deinstitutionalization and the community treatment of people with mental illness, Taylor et al. (1979) and Taylor and Dear (1981) created the Community Attitudes Toward the Mentally Ill (CAMI). The measure they developed used the OMI as a conceptual basis, seeking to regenerate three of five OMI factors (authoritarianism, benevolence, and social restrictiveness) and to create a fourth factor assessing community mental health ideology. The scale includes 40 items, for each one of the proposed factors. The four a priori scales ranged in internal consistency reliability from 0.68 for authoritarianism to 0.88 for community mental health ideology, with benevolence (0.76) and social restrictiveness (0.80) lying in between. Taylor and Dear also show that the scales correlate highly with each other and as expected with demographic variables such as age, gender, occupational status, and household income. The major strength of the CAMI is its exploration of attitudes toward community mental health treatment facilities. Deinstitutionalization represents a major new development in the care and management of people with mental illnesses, and it is very important to have this area included in attitudinal assessments. The wording of the 40 items in the CAMI and the response format is available in Taylor and Dear (1981). Citations to the use of the measure should also include an earlier paper by Taylor et al. (1979), although this article does not include the item wording.

Attributional Measures. Measurement focused on a subject’s emotional reactions (e.g., pity, anger), a subject’s behavioral intentions, and the perceived controllability of a stigmatizing condition stems from attribution theory (Weiner 1986). According to attribution theory, the target’s perceived responsibility for the stigmatizing circumstance predicts either anger and punishing actions (if believed to be controllable) or pity and helping behaviors toward the target (if believed to be uncontrollable). Causes that are seen as changeable over time (unstable) generate conceptions that recovery from the condition is possible, whereas causes that are seen as unchanging (stable) elicit beliefs that the condition is immutable.

Corrigan (2000) provided a conceptual review of two social cognitive models based on Weiner’s work. He first expanded upon Weiner et al.’s (1988) approach by suggesting specific ways in which attribution theory’s causal pathway of cognition-emotion-behavior could be inte-
integrated with anti-stigma change strategies. Second, Corrigan proposed that signals of mental illness (e.g., "that person muttering to himself is crazy") yield stereotypes ("crazy people are unpredictable") that lead to behavioral reactions, including discrimination ("I am not going to let that person talk to me"). In his formulation, stereotypes of people with mental illness are cognitive knowledge structures that mediate behavioral reactions.

Weiner et al.'s (1988) original attribution measure included eight questions about ten illnesses. These eight questions consisted of: (1) three questions using nine-point scales to assess the responsibility, blame, and changeability of each illness; and; (2) five questions about the subject's liking, pity, anger, charitable donations, and personal assistance toward each of the ten conditions. Three separate indexes—controllability, positive emotions, and a helping variable—were created by adding individual item scores together, and internal consistency was reported as good for each of these indexes (see Weiner et al. 1988 for details).

The Attribution Questionnaire (AQ) was developed by Corrigan (2003) to measure key constructs defined in his social cognitive models. He used Weiner et al.'s (1988) measure and 11 questions from Reisenzein (1986) that measured controllability, sympathy, anger, and helping behavior. Corrigan's AQ consists of 21 items measuring six constructs, using a nine-point response scale (1 = not at all, 9 = very much). In Corrigan's study, the AQ was administered following a vignette to assess personal responsibility (three items, alpha = 0.70), pity (three items, alpha = 0.74), anger (three items, alpha = 0.89), fear (four items, alpha = 0.96), helping/avoidant behavior (four items, alpha = 0.88), and coercion-segregation (four items, alpha = 0.89).

The literature shows some evidence for construct validity of the attribution questionnaires. Weiner et al.'s measure (1988) showed the expected finding that subjects rated causes of five mental-behavioral stigmas (Vietnam War syndrome, AIDS, child abuse, drug abuse, and obesity) as more controllable than five physically based stigmas (Alzheimer's disease, blindness, cancer, heart disease, and paraplegia). Subjects also endorsed less liking, pity, and assistance and more anger toward the mental-behavioral stigmas than the physically based ones.

Corrigan's (2003) measures also show evidence of construct validity by correlating with relevant concepts in an anticipated manner. For example, Corrigan has shown that perceived controllability is related to avoidance, withholding help, and the endorsement of support for coercive treatment (Corrigan 2003). In another study, Corrigan et al. (1999) found that subjects who rated mental illness as less blameworthy were more likely to perform real-world helping behaviors (e.g., to sign a petition protesting unfair depiction of mental illness in the media). Finally, Corrigan et al. (this issue) conducted an experiment that showed that contact with a person with mental illness led to significant changes in perceived personal responsibility and dangerousness and that these benefits remained at follow-up 1 week later. Hence, it appears that subscales measured by Corrigan's AQ relate to each other and to a highly related construct (i.e., previous exposure to people with mental illness) in ways that attribution theory would lead us to expect.

Another instrument that assesses causal attributions is the Revised Causal Dimension Scale (CDSII, McAuley et al. 1992; CDS, Russell 1982). The CDSII has a locus of causality scale, which determines whether the cause is located within or is external to the person making the attribution, and a stability dimension, which refers to whether the cause is fixed or changeable over time. Three items with anchoring statements, rated on a nine-point Likert scale, compose each of the four attributional domains in the CDSII: Locus of causality (internality), stability, external control, and personal control. Alpha for each subscale ranged from 0.60 to 0.92 across four studies, with the following averages: locus of causality, 0.67; stability, 0.67; personal control, 0.79; and external control, 0.82. When used to measure stigma, the CDSII has been modified to assess an observer's attributions about a stigmatized individual (Boisvert and Faust 1999). A copy of the AQ can be found in Corrigan (2003), while the CDSII can be found in McAuley et al. (1992).

Key Informant Questionnaire and EMIC. The structured Key Informant Questionnaire (Wig et al. 1980; Alem et al. 1999) assesses the way mental disorders are perceived in the community, how people react to disorders, and whom people turn to for treatment when faced with a mental disorder. It was administered to key informants (people whose perceptions were based on a wide range of problems in the community and whose views would be important in decision making) in India, Sudan, and the Philippines. Seven vignettes depicting mental retardation, epilepsy, acute psychosis, mania, depressive psychosis, schizophrenia, major depression, and neuromuscular depression were presented to respondents. Subjects were asked to rate the gravity (seriousness or harmfulness), prognosis, marriage prospects (likelihood that the condition would impair chances of marriage), ability to live at home, and ability to work/study of the individual in each vignette on a three-point (e.g., not serious—most serious) scale.

The pattern of findings in the two studies that have used this measure provides evidence for the construct

---

validity of the vignette section and scales. A weakness of this measure is the absence of reliability data for the vignette questions. However, one notable strength of these vignettes and attached questions is their apparent adaptability to different cultures; “skeleton” vignettes describing basic features of each illness can be fleshed out to accommodate different cultural settings using local language and expressions.

A similar instrument called the EMIC (Chowdhury et al. 2000) was developed and used in India. The EMIC uses vignettes and a scale measuring perceived stigma based on Goffman’s (1963) conception of “spoiled identity.” Internal consistency of the EMIC stigma scale was reported as 0.66 and 0.76 when administered to the general public and health care professionals, respectively. The original Key Informant Questionnaire can be found in Wig et al. (1980), and the EMIC is presented in Chowdhury et al. (2000).

**Emotional Reaction To Mental Illness Scale.**
Angermeyer and Matschinger (1996) developed a scale to measure emotional reactions toward people with mental illnesses. In two representative surveys of adults in Germany, the investigators administered two vignette descriptions (one describing schizophrenia and the other depression). In the first survey (1990), the measure consisted of 18 five-point Likert scale items, with each item assessing a single emotional response. Factor analysis yielded three dimensions: (1) aggressive emotions (e.g., anger, irritation); (2) prosocial reactions (desire to help, sympathy); and (3) feelings of anxiety (uneasiness, fear). The final version of the instrument included the four items that loaded highest on each factor. This instrument’s key strengths are its assessment of affective experiences of the stigmatizer, which have previously been underassessed; its demonstrated reliability; and its validity in demonstrating a predicted pattern of relationships with the construct of previous contact with mentally ill people. A copy of the original measure can be found in Angermeyer and Matschinger (1996).

**Perceived Devaluation-Discrimination—General Public.**
Link (1987) constructed a perceived devaluation-discrimination measure to test hypotheses associated with the “modified labeling theory.” The measure assesses a respondent’s perception of what most other people believe—a key feature of modified labeling theory. Link (1987; Link et al. 1989, 1991, 1997, 2001) developed a 12-item perceived devaluation-discrimination measure that asks respondents the extent to which they agree or disagree with statements indicating that most people devalue current or former psychiatric patients by seeing them as failures, as less intelligent than other persons, or as individuals whose opinions need not be taken seriously. The scale also includes items that assess perceived discrimination by most people in jobs, friendships, and romantic relationships. The items were originally administered in a paper-and-pencil questionnaire in a six-point Likert format.

The scale has been used mainly among people in treatment for mental illnesses but can be administered to members of the general public. In fact, information from the general public is crucial to testing modified labeling theory. The theory predicts that perceived devaluation-discrimination should have no impact on social or psychological functioning in people who have never been officially labeled with mental illness. Absent the label, the individual may believe that others will reject psychiatric patients, but this belief has no personal relevance. Consistent with this prediction, when the scale is administered to persons who have never been officially labeled, there are no significant associations between scale scores and earnings, unemployment, social network ties, or psychological demoralization (Link 1987).

**Measures Applicable To Children in the General Population**

Our data base included four papers based on child or adolescent research participants, one of which developed measures tailored specifically to children. Adler and Wahl (1998) assessed third graders’ conceptions of people with mental illness with two methods. First, children were shown pictures of a man labeled as “a mentally ill person,” “a physically disabled person,” or “a regular grown-up.” The children were asked to tell stories about the man, and the stories were rated on a number of known negative attributions about people with mental illnesses (e.g., dangerous, unpredictable). Second, children were asked a series of questions about what the man would be likely to do, such as “dress nicely,” “have a lot of friends,” or “yell a lot.” While no reliability data were reported, some evidence for the validity of the coded-story approach is available in the expected finding of more globally negative statements about the person with mental illness. Concerning the direct questions, results did not differ for the mentally ill man, the physically disabled man, and the regular grown-up. This may indicate that third graders have a general idea that people with mental illness are not as good as other people but have not developed specific negative stereotypes. The 18 direct questions can be found in Adler and Wahl (1998). Readers interested in children’s views of mental illness are also directed to Wahl’s (2002) review of this literature, which describes additional measures adapted for use with children.
Measures Applicable To Patients/Consumers

Mental Health Consumers’ Experience of Stigma. The most comprehensive measure of mental health consumers’ experience of stigma is Wahl’s (1999) self-administered questionnaire. The questionnaire includes nine statements concerning stigma experiences such as worrying that others will view one unfavorably, being treated as less competent, being advised to lower one’s expectations, and hearing others say unfavorable things about oneself. There are 12 items concerning discrimination experiences such as being denied a job, a passport, educational opportunities, housing, or health insurance coverage when consumer status was revealed, as well as avoiding indicating consumer status on written applications for fear of discrimination. Items are presented as statements that are responded to as occurring “never,” “seldom,” “sometimes,” “often,” or “very often.” Factor analysis indicated that, of four factors accounting for 50 percent of the variance, the first two factors represented stigma and discrimination, respectively (Mann and Wahl 2003). Evidence for validity of the measure is provided by Zaveri et al. (2003), who reported that (using an appropriately modified version of the measure), people with epilepsy report experiences of stigma and discrimination, but to a lesser extent than people with mental illnesses.

Dickerson et al. (2002) adapted the instrument for a sample of outpatients with schizophrenia by supplementing the term “consumers” with the terms “persons with mental illness,” “persons who have a psychiatric disorder,” and “persons who use psychiatric services,” because pilot interviews indicated that many respondents were not familiar with the use of the term “consumers” to refer to consumers of mental health services.

Measures for Consumers Associated with Modified Labeling Theory. Previously, we described Link’s (1987) perceived devaluation-discrimination scale. According to modified labeling theory, one’s perception of how most people treat a person who is officially labeled as having a mental illness becomes personally relevant when a person develops a mental illness and is officially labeled. Moreover, anticipating status loss and discrimination, a person may seek to avoid such negative outcomes by adopting coping orientations such as keeping a history of treatment a secret, educating others so as to ward off stereotypes, or withdrawing to a group of people who have been similarly labeled or who know about the label and can be trusted. While such coping orientations can protect a person from rejection, they may also carry costs. To evaluate whether people do indeed adopt such coping orientations and to assess the consequences of doing so, Link et al. (1989) developed measures of the coping orientations of secrecy, education, and withdrawal. More recently, Link et al. (2002) expanded the assessment of coping orientations to include distancing and challenging as well as stigma-related feelings, of being misunderstood by others and different and ashamed (see below). The significance of these new measures for modified labeling theory is that they provide a more fully elaborated set of empirical assessments that can be used to test that theory.

Measures of rejection and perceptions of rejection. Devaluation-discrimination is the 12-item measure whose content and scoring are described in the section on measures for the general public. In the measure’s use among consumers of mental health services, internal-consistency reliability ranges from 0.82 (Link et al. 1991) to 0.86 (Link et al. 2001). Evidence for validity comes mainly from construct validity. Theory predicts that the measure will be related to a wide variety of outcomes among people with mental illnesses, and to date evidence exists linking the measure to earned income (Link 1987), employment status (Link 1987), social support networks (Link et al. 1989; Perlick et al. 2001), demoralization (Link 1987), quality of life (Rosenfield 1997), depressive symptoms (Link et al. 1997), help seeking (Sirey et al. 2001), and self-esteem (Link et al. 2001).

Link et al. (1997) also constructed a 12-item measure (alpha = 0.80) of rejection experiences, for use among dually diagnosed persons with serious mental illness and substance abuse. The scale was a strong predictor of depressive symptoms in this longitudinal study. However, subsequent to the creation of this scale, Wahl (1999) developed a more complete set of items measuring rejection/discrimination experiences (see above), and we now recommend the use of that scale over the one developed by Link et al.

Measures of coping orientations. Secrecy assesses the extent to which participants endorse concealment as a means of avoiding rejection. An early scale constructed by Link et al. (1989) included five items (alpha = 0.71). More recently, Link et al. (2002) revised and expanded the scale to include nine items (alpha = 0.84) with new items such as, “You encourage other members of your family to keep your mental illness a secret.”

Withdrawal assesses the extent to which people endorse withdrawal or avoidance as a means of self-protection from potential rejection. The idea stems from the work of Goffman (1963) and focuses on the tendency to limit social interaction to those who know about and tend to accept one’s stigmatized circumstance. An original seven-item version developed by Link et al. (1989)
(alpha = 0.67) was recently modified and expanded to a nine-item scale (alpha = 0.70) (Link et al. 2002).

Educating measures participants' orientation to educating others as a means of reducing the possibility of rejection. The idea stems from the work of Schneider and Conrad (1980) and their ideas about "preventive telling." An earlier five-item version developed by Link et al. (1989) (alpha = 0.71) was recently revised to yield a new three-item scale (alpha = 0.67).

Challenging measures people's orientations to confronting prejudice and discrimination. Link et al. (2002) created a five-item scale (alpha = 0.72) to assess the extent to which participants are likely to point out stigmatizing behavior when it occurs, disagree with people who make stigmatizing statements, and so on.

Distancing is a newly created three-item scale (Link et al. 2002) (alpha = 0.63) assessing the extent to which people cope with stigma by cognitively distancing themselves from the stigmatized group, by indicating that their problems are very different from those of other people with mental illness and that they have little in common with such people.

Evidence for the validity of the coping orientation scales varies from scale to scale. Face validity is evident for most of the items, as they appear to assess the construct the scale purports to measure. Some evidence for the construct validity of the scales comes in a pattern of expected correlations between the various measures, as reported in Link et al. (2002). Additional evidence for the construct validity of the withdrawal scale exists in Link et al. (1989) and subsequently Perlick et al.'s (2001) findings that withdrawal is associated with a constricted social support network outside the household.

Stigma-related Feelings. A missing component in the measurement of stigma-related processes associated with modified labeling theory has been an assessment of the feelings that stigma creates in the people exposed to it. As a beginning effort to fill this gap, Link et al. (2002) introduced two new scales—feeling misunderstood and feeling different and ashamed.

Feeling misunderstood assesses the extent to which people feel that their experience of mental illness has been misunderstood by others (alpha = 0.62). Example items are as follows: "Most people cannot understand what it is like to be a patient in a mental hospital" and "Most people have no idea what it is like to experience a serious mental illness." The scale correlates with other scales created by Link et al. (2002) in expected ways: 0.34 with perceived devaluation-discrimination, 0.38 with rejection experiences, and 0.24 with withdrawal.

Feeling different and ashamed is a four-item scale (alpha = 0.70) measuring the extent to which people's experiences of mental illness and mental hospitalization make them feel set apart, different from other people, and ashamed. Once again, this scale correlates as expected with other scales: perceived devaluation-discrimination 0.48, rejection experiences 0.28, and withdrawal 0.48. In addition, feeling different and ashamed is more strongly correlated with self-esteem (0.50) and depressive symptoms (0.51) than any other stigma scale mentioned above. These correlations are consistent with the possibility that other stigma variables have their effect on self-esteem and feelings of depression through feelings of being different and ashamed.

The items in some of the measures associated with modified labeling theory (perceived devaluation-discrimination, secrecy, withdrawal, and education) are published in the American Sociological Review (Link et al. 1989). However, all of the items for all of the newly developed and revised scales are available in the spring 2002 issue of Psychiatric Rehabilitation Skills (Link et al. 2002).

Measurement Biases in Self-Report Measures of Stigma Components. As with other measurement approaches, there are potential biases associated with using self-report measures of stigma experiences or stigma feelings. For example, Major et al. (2002) have indicated that measures of neuroticism can be associated with the perception of being stigmatized and with measures of well-being, thereby bringing into question any causal link between stigma and well-being. Similarly, a person who is unemployed, isolated, or beset by low self-esteem may seek to explain his or her disadvantaged status by invoking stigma. In such a scenario, levels of measured stigma do not cause bad outcomes but are instead consequences of those outcomes. There is no single prescription for avoiding such biases. Instead, one needs to carefully consider these possibilities as threats to validity and seek ways of addressing them. For example, one might include a measure of neuroticism (or other such confound) in one's measurement protocol (Major et al. 2002), employ multiple approaches to assessing stigma rather than relying on self-report alone, or craft measurement and design features that help address potential confounds such as these (e.g., Link 1987).

Stigma Experienced by Family Caregivers

Greenberg et al. (1993) present new measures of subjective burden experienced by parents of people with mental illness, including stigma. The stigma measure was adapted from earlier work by Freeman and Simmons (1963) and consists of the mean of seven items asking the extent to which family members avoid having family and
friends over or avoided telling others about their child’s illness for fear of what others may think of them (alpha = 0.74). Evidence for validity comes from the fact that the scale is significantly correlated with co-residence with the ill child (r = 0.27), greater psychiatric symptomatology of the child (r = 0.26), and a 27-item measure of “objective burden” consisting of stressful life events for the family (r = 0.23).

Szmukler et al.’s (1996) Experience of Caregiving Inventory (ECI) includes a five-item stigma scale (alpha = 0.82) assessing the frequency of distress over covering up the relative’s illness, feeling unable to tell anyone about the illness or have visitors to one’s home, feeling stigma, and worrying about how to explain the illness to others. Martens and Addington (2001) found that the stigma scale of the ECI was strongly negatively correlated (r = -0.49; p < 0.01) with the psychological well-being of a sample of relatives of people with mental illness as assessed with the General Health Questionnaire (Bech 1993). Item wording can be found in Szmukler et al. (1996).

Most recently, Struening et al. (2001) developed a seven-item scale answered on a four-point scale (strongly agree to strongly disagree) to estimate the extent to which family caregivers believe that most people devalue families that include a person with a serious mental illness. The seven-item scale has an alpha of 0.71 in a sample of family members of people with schizophrenia and schizoaffective disorders (n = 180) and 0.77 in a sample (n = 281) of family members of people with bipolar I and II disorders. Item wording can be found in Struening et al. (2001).

Use of Vignettes in Research on Stigma

One of the most common methodological approaches employed in the study of the stigma of mental illness is the vignette. The approach was first used in this context by Shirley Star, in a nationwide study of public attitudes toward mental illnesses (Star 1955). Star constructed vignettes depicting paranoid schizophrenia, simple schizophrenia, alcoholism, anxiety neurosis, juvenile character disorder, and compulsive phobia and administered them to over 3,000 residents of the United States in 1950. After each vignette, Star inserted queries about the described person to ascertain the respondent’s judgment as to how serious the condition was and whether or not it was some kind of mental illness.

An important turning point in the use of vignettes came when Derek Phillips (1963) employed the Star vignettes in a survey experiment. Phillips varied five vignette disorders (four from Star and a “normal” man) and five help sources (no help source, clergy, physician, psychiatrist, and mental hospital) in a classic Greco-Latin Square experimental design. Phillips’ study showed that help source significantly influenced the desire for social distance, thereby indicating that rejection might be a consequence of seeking mental health treatment. From a measurement and methods point of view, Phillips’ innovation ensured the future use of vignettes by integrating their use with the experimental method. Subsequently, researchers have randomly varied symptoms, behaviors, labels, causal attributions, social statuses (gender, race/ethnicity, education), and other characteristics in vignette studies.

But the Star vignettes per se are rarely used any more. DSM-III (and subsequently DSM-III-R and DSM-IV) brought new and far more explicit criteria for diagnosing mental disorders that rendered the Star vignettes out-of-date as representations of mental illnesses as conceptualized by mental health professionals. Researchers at Indiana and Columbia Universities collaborated to produce the MacArthur Mental Health Module of the 1996 General Social Survey, which included four DSM-IV vignettes: alcoholism, major depression, schizophrenia, and cocaine abuse. An additional vignette describing a “troubled person” who did not meet criteria for mental disorder was constructed to serve as a point of comparison for the vignettes depicting disorders. The vignettes were then administered to a nationally representative sample of 1,444 respondents in an in-person interview. The exact wording of the vignettes employed in the survey is available in Link et al. (1999) and Pescosolido et al. (2000). Vignettes depicting disorders according to the International Classification of Diseases have also been developed by Angermeyer and Matschinger (1997) to study mental health literacy in Australia and by Jorm et al. (1997) to study public attitudes in Germany and by Jorm et al. (1997) to study mental health literacy in Australia. A vignette, then, is a form of stimulus that researchers can ask people to react to. Following a vignette description, many, many types of measures can be applied. Examples of questions investigators have asked following vignettes can be accessed by reviewing recent vignette-based studies (table 3). In addition, a report by Pescosolido et al. (2000) contains all of the vignette-based questions employed in the 1996 General Social Survey.

Ever since Star and Phillips, vignettes have enjoyed a very prominent position in research on the stigma of mental illness. There are two major reasons for this popularity. First, vignettes allow the researcher to present a more elaborate stimulus to respondents than is afforded in measurement approaches that simply ask people about “mental illness,” a “psychiatric hospital
patient,” or a “mental health consumer.” Second, vignettes allow the use of random assignment and bring the power of the experimental method to hypothesis testing. Furthermore, because vignettes can be used in survey research, vignette experiments can be administered to randomly selected general population samples and therein achieve somewhat better external validity than is typical of many laboratory experiments that employ college students as subjects.

These advantages have ensured and will continue to ensure the use of vignettes in future research. At the same time, it is important to recognize that vignettes are hypothetical and abstracted from “real life” experience. For example, a member of the public rarely encounters a vignette-type situation in which a person systematically displays all the requisite symptoms for a diagnosis of schizophrenia. Moreover, the respondent is not in the presence of a real person, is not gleaning information from appearance and other nonverbal cues, and cannot assess the described person’s responses to initial gestures that might affect reactions in “real” situations. In the absence of information of this sort, people rely on cognitive schemas or scripts to answer questions about vignettes that may not map on to their actual behavior in real situations. Although vignettes have drawbacks, we believe the best approach to developing future knowledge in this area will be built on information derived from multiple methodological approaches with different strengths and weaknesses. Vignette-based research can be one important component of such a multimethod approach.

Measurement of Stigma-Relevant Behavior

The final component of the Link and Phelan (2001) conceptualization of stigma is status loss and discrimination. Clearly, assessing either status loss or discrimination involves the measurement of behavior. Although most of the measures described above imply that discrimination is a likely consequence of the attitude, belief, or behavioral intention expressed, very few aim to assess discriminatory behavior directly. Why? The reason is that when we seek to measure discrimination, our intent is to determine whether and to what extent people with mental illnesses are denied access to the good things our society affords and are differentially exposed to the bad things it confers. The measures we use to assess discrimination can include employment status; social network ties; access to medical treatment; hiring decisions; influence in group situations; or being shunned, put down, or ignored. None of these are in and of themselves “stigma” measures, and they are not, as a result, reflected in a listing of such measures. They are brought into the domain of stigma research when labeling, stereotyping, and/or being set apart are shown to produce the good or the bad outcome assessed. This can occur either through direct discrimination, through structural discrimination, or via processes that operate through the stigmatized person him- or herself (Link and Phelan 2001). The task for stigma researchers, then, is to select outcome measures from the vast array of possible good or bad outcomes extant in our society. A good selection requires a creative match of an outcome measure to the theory being tested. In the remainder of this section, we review behavioral measures that have been incorporated into experimental studies and then turn to non-experimental studies that have measured discrimination.

Laboratory Experiments and Behavioral Measures Indicating Discrimination. As shown in table 1, most of the empirical studies we located employed nonexperimental survey methods. However, the laboratory experiment, in which the behavior of research participants is observed and measured, has an important tradition in stigma research, especially in the field of psychology. This research has examined phenomena such as the disruptive effect of labeling on social interactions and the impact of labels on the tendency to discriminate, punish, or help. Researchers interested in the stigma attached to mental illness may also draw on methods and measures employed in experimental studies of other stigmatized characteristics such as minority racial status and low socioeconomic status.

Experiments have demonstrated the pernicious effects that the introduction of a mental illness label can have on social interaction. In one study, Sibicky and Dovidio (1986) randomly assigned college students to be unwittingly labeled as a psychotherapy client or as an introductory psychology student. These “target” students then interacted with a “perceiver” student, to whom the label had been disclosed. When the target was labeled as a psychotherapy client, blind observers rated the perceiver as behaving more negatively (i.e., unenthusiastically, cruelly, artificially) toward the target and rated the target as behaving in a less socially desirable manner (see also Farina et al. 1971). Other experiments have measured the effect of labeling on discrimination in housing (Page 1977) and punishment via “shocks” in learning situations (the shocks were not actually administered).

Experimental studies addressing other forms of stigma and discrimination can also be profitably examined by mental illness—stigma researchers. For example, according to Steele and Aronson’s (1995) concept of
"stereotype threat," people in stigmatized groups are familiar with the stereotypes that might be applied to them, and in situations in which these stereotypes are salient, performance on tasks relevant to the stereotype can be impaired. For example, focusing on the stereotype that African-Americans are less intelligent than members of other groups, Steele and Aronson found that, controlling for initial differences on SAT scores, African-American college students performed worse than white students on a test when participants were led to believe that the test measured intellectual ability. In contrast, when the same test was not labeled as being diagnostic of ability, African-Americans scored as well as whites.

Another body of work—that pertaining to "expectations states" and "status construction"—is also highly relevant to the study of stigma, particularly the component of "status loss" (Link and Phelan 2001). This research has shown how status hierarchies based on characteristics such as gender or race are constructed and maintained—and can be altered—through social interactions. Studies have shown that status characteristics affect estimation of a person’s overall social worth as well as specific performance capacities, even when the status characteristic is irrelevant to the task at hand. These expectations in turn lead to behaviors such as who listens and who speaks more, and whose ideas are accepted by others (Berger et al. 1980, 1998). Moreover, an arbitrarily selected characteristic (e.g., whether one prefers paintings by Klee or Kandinsky) can be made a status characteristic by treating a person with that characteristic with more or less deference, and this status can be "taught" to third-party observers (Ridgeway and Erikson 2000).

Thus, there exist many strategies for assessing behavior within the experimental context. One issue regarding these measures is that they come in response to temporary stigma conditions or in designs that make one person think the other is stigmatized. Responses might be different if the person in question was someone who had been in the stigmatized circumstance for some time. In addition, most of the behaviors assessed are short-lived discriminatory consequences that do not have far-reaching effects outside the experimental context. Inducing truly serious consequences for study participants would not be ethical. Nevertheless, several of the studies involve dimensions that are directly connected to real-world outcomes such as assessments of achievement, power, and influence. Stigma researchers can use the examples provided above both as direct models for the measurement of outcomes and as more abstract schemas that can assist in creating new behavioral measures for research on the stigma of mental illness.

Observational Studies and Behavioral Measures Indicating Discrimination. Observational studies do not enjoy the power of the experimental method but are better able to directly measure real-world outcomes. One nonexperimental approach involves asking consumers of mental health services about their experiences of discrimination. In Wahl’s measure (described above) consumers are asked how often discrimination experiences (being denied a job, an educational opportunity, etc.) occurred when the person’s consumer status was known. From such a measure, one can learn about self-reported discrimination in a variety of important life domains. Note that this strategy asks the consumer to make the connection between the label (their psychiatric treatment) and the discriminatory behavior. In a different approach, study participants are asked about employment status, earned income, or social network ties without any reference to the involvement of labels or stereotypes (Link 1982, 1987; Link et al. 1989; Perlick et al. 2001). Because people with mental illnesses tend to be disadvantaged in areas like these, the question becomes whether stigma processes play a role in creating this disadvantage. To gauge the potential importance of stigma processes, independent measures of labeling and stereotyping are related to variation in outcome variables while controlling for other possible influences on those outcomes. An excellent example of this approach is a study by Druss et al. (2000). Their question was whether persons with a label of schizophrenia would receive equal access to optimum treatment approaches for heart attacks. They not only found that people with schizophrenia were less likely to experience optimum treatment but that this difference could not be explained by a comprehensive set of measures assessing the availability of the procedures in question or the physical status of the persons studied.

Still another approach to assessing differential treatment that may indicate discrimination is to include behavioral measures in surveys of knowledge, attitudes, beliefs, and behavioral intentions. For example, in a study by Corrigan et al. (1999), after completing the Psychiatric Disability Attribution Questionnaire (PDAQ), research participants were given the opportunity to sign a petition objecting to stereotypic portrayals of people with mental illnesses in the media. Participants were provided with copies of the petition and instructed to sign it if they wished and deposit it in a box at the front of the room. In another study, Penn and Nowlin-Drummond (2001) used a behavioral measure (willingness to participate in a seminar with mental health consumers and providing a contact number) to help assess the impact of psychiatric labels.

In sum, as with experimental studies, there exist several examples of approaches to assessing behaviors
indicative of differential treatment of people with mental disorders that can be used to study the discrimination component of the stigma process.

Applying Qualitative Methods To the Study of Stigma

Quantitative assessments using measures such as those described above have contributed significantly to our understanding of attitudes, behaviors, emotions, and beliefs that form the expression of and response to stigma. Measuring constructs quantitatively may not fully reflect the intricacies of the lived experience of stigma (Schneider 1988; Kleinman et al. 1995), however. Studies using qualitative methods sometimes directly reflect such experience and thereby deepen our understanding of stigma processes. Thus, qualitative inquiry provides another set of methodologies that can offer rich insights into the subjective experience of stigma and the complexity of social systems that produce stigma.

Seventeen of the identified papers in this review report the results of qualitative studies. In this section we discuss, first, the uses of qualitative research and its relevance for the study of stigma. Second, we examine lessons learned from classic studies of stigma that used qualitative methods. Third, we describe selected qualitative studies, focusing on the concepts of stigma addressed and highlighting specific studies that have deepened our understanding of stigma.

Uses of Qualitative Research and Relevance for the Study of Stigma. Bryman and Burgess (1999) define qualitative research broadly as "[a] strategy of social research that deploys several methods and displays a preference for the interpretation of social phenomena from the point of view of the meanings employed by the people being studied" (Introduction, p. x). Qualitative studies enable researchers to obtain an insider's view of the situation under investigation (Maxwell 1996). Data collection that may occur through in-depth interviews (semi-structured or unstructured), chart review, participant observation, life history, oral history, documentary research, diaries, film, video, or photography captures the complexities of life circumstances. Data analysis prioritizes depth of understanding (Patton 1990; Bryman and Burgess 1999).

The study of stigma lends itself to qualitative methods of investigation. These research methods permit the investigator to understand how stigma is constructed in social interaction and how people interpret their experiences and their behavior (Bryman 1999). These methods are essential for appreciating the subtle, damaging effects of stigma. Structural discrimination, for example, can only be understood when the institutions through which it operates are well characterized, the history of the relationship with the stigmatized group understood, the policies of the institution examined, and the attitudes of its leaders explored. Out of this in-depth exploration may come confirmation of previous theories as well as new findings that contribute to new theory development (Maxwell 1996; Strauss and Corbin 1998).

Qualitative research may also be exploratory or may form the basis of further study (Yang and Fox 1999; Bryman and Burgess 1999). Some studies in the current selection illustrate the use of qualitative methods in conjunction with quantitative methods for three purposes identified by Hammarsley (1996): (1) methodological triangulation (each method serves to validate the findings of the other), (2) facilitation (one method serves as research groundwork for the other), or (3) complementarity (qualitative and quantitative methods are used to explore different aspects of one question).

Classic Qualitative Studies of Stigma. Qualitative research studies dramatically influenced the early social science literature on mental illnesses. Major studies focused on the social structure and social climate of psychiatric hospitals (Stanton and Schwartz 1954; Caudhill 1958; Goffman 1961), family identification of psychosis (Yarrow et al. 1955), social reactions to symptoms (Lemert 1962), and involuntary commitment proceedings (Scheff 1964). While most of these studies were not specifically about stigma, the processes identified by our current conceptualization (see above) were very prominent in them. They all addressed the social arrangements used to manage and control people with mental illnesses. The endeavor invariably involved issues of stigma as people with mental illnesses experienced labeling, faced powerful stereotypes, were set distinctly apart, experienced status loss and discrimination, and were assigned little personal power in many of these processes. One finds in these studies many of the elements we think about today when we discuss stigma and its consequences.

Review of Current Selected Qualitative Studies. As in the classic studies described above, several of the qualitative studies in the current review do not focus primarily on stigma related to mental illness. As researchers apply qualitative methods to study the experiences and attitudes of people with mental illness, families, or health care providers, references to stigma or shame frequently
emerge. We encounter descriptions of stigma in studies that investigate community views toward seeking mental health care (Fuller et al. 2000), attitudes toward the treatment of depression (Priest et al. 1996), the experience of help seeking (Kai and Crosland 2001), determinants of disclosure of mental health problems (Williams and Healy 2001), and experiences working with people with psychiatric diagnoses (Lyons and Ziviani 1995), as well as in the experiences of hospitalization (Sayre 2000) and the community life of persons diagnosed with mental illness (Bedini 2000).

Themes of status denial and discrimination, avoidance of labeling, and stereotyping predominate. Of the 17 studies, 9 focus primarily on the stigmatized. They describe the perspectives of persons with a psychiatric diagnosis or mild mental health problem or the perspectives of family members of people with mental illness (Mohr, 1998; Scheff 1998; Bedini 2000; Sayre 2000; Williams and Healy 2001; Kai and Crosland 2001; Treasure et al. 2001; Muhlbauser 2002; Dalgin and Gilbride 2003). Seven studies convey the perceptions and experiences of the potential stigmatizer (Lyons and Ziviani 1995; Priest et al. 1996; Bailey 1998; Secker et al. 1999; Fuller et al. 2000; Shor and Sykes 2002; Pinfold et al. 2003). Williams and Taylor (1995) conducted a content analysis of newspaper articles, examining the portrayal of mental health issues in the media. Semistructured and unstructured interviews were the most frequently used methods of qualitative data collection. Of these 17 studies, we highlight 2 that illustrate how stigma functions in a particular context of mental health service provision.

Sayre (2000) studied the mechanisms that patients use to protect against status loss and stigma when faced with psychiatric hospitalization. The analysis of interviews, field notes, and informal interviews reveals that most patients opt to reject the psychiatric explanation of their problems and the negative social implications of psychiatric hospitalization. Themes that emphasize the social process of managing self-worth emerge in explanations of the reasons for hospitalization. Sayre describes six attribution accounts (problem, disease, crisis, punishment, ordination, violation) that enable patients to manage their self-worth and protect against status loss. By attributing hospitalization to external stressful events (i.e., crisis), patients avoided acknowledging that they had a mental illness and were thus somewhat protected from the negative stereotypes and stigma associated with mental illness.

The second study (Kai and Crosland 2001) investigated the experiences and perceptions of health care for 34 respondents with chronic mental illness. Stigma issues that emerged focused on victimization secondary to being identified as a psychiatric patient, social isolation, and fear. Significantly, this study suggests that the hazards of living in low-income, high-crime environments, and having few economic resources compounded the stress related to stigma. These individuals experienced social exclusion because of stigma related to their mental illness and because of poverty secondary to their disability. Thus, stigma related to mental illness does not operate in isolation, but in concert with other challenges in the lives of psychiatric patients.

Qualitative studies in the literature on stigma of mental illness vary widely in their treatment of stigma—from relatively superficial reports of the experience of "stigma" or the fear of being stigmatized, to deeper exploration of certain components (usually status loss, discrimination, stereotyping, and labeling). Of note, these studies are few in number in the current review (table 1). The underutilization of qualitative methodologies is significant given that certain aspects of stigma can best be explored through the use of qualitative research.

Gaps in Measurement

Our review is intended not only to help researchers locate suitable measures but also to identify gaps in the use or availability of measures. Our review revealed several prominent areas of inquiry that appear to be understudied.

Structural Discrimination. Earlier in this article, we referred to the concept of structural discrimination, which we conceptualize as institutional practices that work to the disadvantage of stigmatized groups and that allow extensive disparities in outcomes even when direct person-to-person enactment of discrimination is absent. We found this form of discrimination almost entirely unaddressed in the log stigma papers we reviewed. While issues of structural discrimination (e.g., insurance parity for people with severe mental illnesses) are discussed in the literature, they have not been integrated into the literature on stigma. Any complete accounting of the processes that disadvantage people with mental illnesses must incorporate such phenomena. One way to do this is through ethnography. Classic studies such as Goffman's (1961) Asylums, Caudhill's (1958) The Psychiatric Hospital as a Small Society, and Estroff's (1981) Making It Crazy are excellent examples of how this can be achieved. At the same time, quantitative researchers interested in stigma should attend to this by investigating, for example, the
revealed a shortage of nonvignette experimental studies for two reasons. First, the field was powerfully addressing the stigma of mental illness. This is surprising given the centrality of feelings of shame, humiliation, and embarrassment in the area of stigma. We are aware of two projects seeking to expand this domain of measurement, but at the time of this writing neither haved been published.

Assessment of the Emotional Responses of Patients/Consumers. Our review found only four studies assessing the emotional responses of patients/consumers and few systematic measures for this domain of inquiry (see Link et al. 2002 for existing examples). This is odd and unfortunate given the centrality of feelings of shame, humiliation, and embarrassment in the area of stigma. We are aware of two projects seeking to expand this domain of measurement, but at the time of this writing neither haved been published.

Assessment of Children’s Knowledge, Attitudes, Beliefs, and Behaviors and Children’s Experience of Stigma. We found only 4 studies out of the 109 we reviewed that focused on children or adolescents. This is remarkable given the salience of the issue for understanding the mechanisms through which our society perpetuates stereotypes. One of the studies is both clear and disturbing in indicating that children are exposed to stereotypes through the cartoons they watch (Wilson et al. 2000), and another suggests that developmental issues are likely central by indicating that third graders have learned that mental illnesses are “bad” but have yet to develop specific stereotypes (Adler and Wahl 1998). These important studies break new ground, but we need to understand much more about when children develop conceptions of mental illnesses, what forms their perceptions, and what kinds of educational experiences might move conceptions in a positive direction.

Equally as striking, we found no studies assessing children’s experience of stigma; all the studies of patients/consumers were studies of adults. Moreover, none of the measures discussed were aimed at assessing children’s experience of stigma.

Use of Experimental Approaches. Our review also revealed a shortage of nonvignette experimental studies addressing the stigma of mental illness. This is surprising for two reasons. First, the field was powerfully influenced by the work of Amerigo Farina and his experimental paradigm for the study of mental illness stigma (e.g., Farina et al. 1968, 1971). Second, experimental studies concerning stigmatizing conditions other than mental illnesses would seem to be on the rise as reflected in the excellent work of social psychologists (see Crocker et al. 1998 for a review). As we hope our section on experimental studies makes clear, there is a great opportunity to adapt strategies and apply theories from these approaches.

Cross-Cultural Approaches. Although we identified a number of studies conducted in non-Western societies, many of these applied concepts and instruments that were developed in the West. Studies that identify how stigma is constructed and acted upon in non-Western cultures would contribute greatly to our understanding of these processes in the United States. Such studies would help us understand how stigma components of labeling, stereotyping, separating, emotional reactions, status loss, and discrimination are similar and different in other cultures. This kind of information could help illuminate how we might change stigma processes in our own culture.

We began our review by referring to the Surgeon General’s Report on Mental Health and to the prominent position he assigned to stigma in that report. Proceeding from his statement is the idea that improvement in the nation’s mental health and mental health care will depend in some large part on whether stigma processes are identified and effectively addressed. One way our review contributes to this goal is by providing information about the measurement of stigma that can guide efforts to identify stigma processes and evaluate efforts to address those processes. We hope that improved measurement of stigma will play a direct role in shaping the kinds of broad policies that are developed to improve the lives of people with mental illnesses and their families.

References (Data Base)

3. Angermeyer, M.C., and Matschinger, H. The effect of personal experience with mental illness on the attitude


53. Kirmayer, L.J.; Fletcher, C.M.; and Boothroyd, L.J. Inuit attitudes toward deviant behavior: A vignette study.
78. Perlick, D.A.; Rosenheck, R.A.; Clarkin, J.F.; Sirey, J.A.; Salahi, J.; Struening, E.L.; and Link, B.G. Stigma as a barrier to recovery: Adverse effects of perceived stigma


References (Text)


**Acknowledgments**

Support for this article was provided through Purchase Order No. 263–MD–20281 from the National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services; as well as grants 5T32 MH 13043 to Dr. Link, NIMH K02 MH 65330 to Dr. Phelan, and NIMH K01 MH01691 to Dr. Collins. We thank Matthias Angermeyer, Patrick Corrigan, David Penn, Elmer Struening, and Otto Wahl for comments on previous drafts. We thank Richard Carpiano for his careful review of the manuscript and for assistance in constructing the data base of stigma articles. Special thanks to Emeline Otey for comments on previous versions and her extensive support for this undertaking.

**The Authors**

Bruce G. Link, Ph.D., is Professor of Epidemiology and Sociomedical Sciences, Columbia University, New York, NY; and Research Scientist, New York State Psychiatric Institute, New York, NY. Lawrence H. Yang, Ph.D., is Post-Doctoral Fellow, Columbia University. Jo C. Phelan, Ph.D., is Associate Professor of Sociomedical Sciences, Columbia University. Pamela Y. Collins, M.D., M.P.H., is Assistant Professor of Epidemiology, Columbia University; and Research Scientist, New York State Psychiatric Institute.