CHAPTER 37

METHODOLOGICAL CHALLENGES AT THE SOCIAL/CLINICAL INTERFACE

Timothy W. Smith
Frederick T. Rhodewalt

The integration of social and clinical psychology has matured. Early landmarks in the development of this interface are best characterized as cogent discussions of the relevance of social psychological theory for the explication and even refinement of clinical enterprises such as psychotherapy (Frank, 1961/1973; Heller, 1963; Goldstein, Heller, & Sechrest, 1966; Carson, 1969; Brehm, 1976). These efforts clearly established the potential contribution of social psychological models to the study of human adjustment and therapeutic change. More recent reviews of the interface between these fields (e.g., Brehm & Smith, 1986; Weary & Mirels, 1982) reflect a different, later stage in the development of a discipline. The intervening years were associated with a wide array of empirical studies evaluating the previous theoretical contributions and several more recent ones. The promissory note of clinical relevance was paid, at least in part, by a large and growing body of research.

Despite this progress, considerable challenges remain. In the past 20 years, clinical psychology has become a more empirically oriented field. Forces from within and outside the discipline have encouraged and even required stronger evidence of the accuracy of theories of etiology, the validity of assessment procedures, and the effectiveness of interventions. To sustain and advance the process of integration, social-clinical research efforts must continue, and interface research must continue to grow more ambitious. Much of social-clinical research still consists of laboratory-based analog studies. Undergraduate samples are used, and the experimental manipulations are intentionally short-lived. Yet, the results of such studies are often interpreted as relevant to more severely disordered groups and the long processes of the development and remediation of dysfunction. In the years ahead, additional innovations in which social psychological concepts are applied to clinical topics can be tested appropriately in this manner, at least initially. However, a more substantial and lasting impact of the interface on mainstream clinical psychology will require more compelling evidence of the value of social psychological models. This more compelling evidence, in turn, requires more direct tests with clinical populations.
Research at the interface of social and clinical psychology must address the methodological problems common to the two component disciplines. But there are also methodological challenges that, if not unique, are particularly salient in integrative social-clinical research. If this enterprise is to continue and develop, research methods must evolve so as to yield additional evidence of the value of the integrative approach.

In this chapter, we attempt to articulate the methodological challenges faced by social-clinical researchers. The list is certainly not exhaustive, nor is the list of suggested solutions complete. However, our discussion should highlight the ways in which researchers must work to increase the scientific yield of their studies. Cook and Campbell (1979) provide a useful framework for organizing the methodological challenges at the interface. Statistical conclusion validity refers to inferences about the sensitivity and accuracy of tests of covariation among research operations, as well as presence and magnitude of covariation. Internal validity refers to inferences about whether or not apparent statistical covariation between research operations reflects spurious relations due to confounding factors. Construct validity refers to inferences about the extent to which research operations (and the covariation between them) actually reflect the construct(s) they are intended to reflect. Finally, external validity in the Cook and Campbell (1979) system refers to inferences about the extent to which the relationships observed generalize to other subjects, settings, and times. Although statistical conclusion validity is occasionally a problem in the social-clinical research literature (Sedlmeier & Gigerenzer, 1989), the major difficulties occur elsewhere. An overreliance on simple correlational methods weakens the internal validity of large portions of the social-clinical research literature, and many difficult challenges concern construct and external validity.

Our first major section describes methodological challenges in testing models of emotional and physical health. Difficulties in operationalizing predictors and outcomes, and associated threats to construct and external validity, are discussed. Common problems in testing the association between predictors and outcomes also are discussed. The second major section discusses problems in testing models of diagnosis and treatment. We conclude with a discussion of the more general issue of the relation between theory and research at the interface. Hopefully, this review of methodological challenges will serve as a guide for the maintenance and continued development of the empirical foundation of the social-clinical interface.

**TESTING MODELS OF EMOTIONAL AND PHYSICAL HEALTH**

Certainly one of the largest areas in the integration of social and clinical psychology is the study of psychopathology and adjustment. Social psychological models of a wide variety of adjustment problems have been developed and at least partially tested. In recent years these models have expanded beyond emotional adjustment and psychopathology to include physical health as well. As indicated in previous chapters in this volume, many traditional social psychological concepts and paradigms have been employed in the theoretical explanations of emotional and physical disorders.

In some cases, social psychological approaches have been accepted in the mainstream of psychopathology research. The learned helplessness model of depression (Seligman, 1975; Abramson, Seligman, & Teasdale, 1978) is an obvious example of this achievement. In other cases, social psychological models are gaining empirical support, but have not yet been accepted outside of the interface of social and clinical psychology. For the accepted models to retain their status and the developing models to achieve such acceptance, a variety of methodological challenges must be met. The causal constructs in the models—predictors of adjustment—must be operationalized in valid ways. The specific constructs explained—the adjustment outcomes—also must be assessed through valid operations. And finally, the association between the causal and outcome constructs must be tested in an appropriate and compelling manner. At first glance, these appear to be most elementary of methodological requirements. As we will argue, however, many of the current shortcomings in social-clinical research on adjustment involve these basic issues. As a result, the future of the interface will be determined, at least in part, by our ability to correct these limitations.

**Operationalizing Predictors**

Social psychological models of adjustment posit a variety of situational factors and person variables as causes or at least predictors of emotional and physical outcomes. The results of the studies testing these models are useful only to the
extent that the specific research operations reflect the hypothesized causal or predictive constructs. In several areas, the construct validity of operational predictor variables is open to serious question.

Studies of Life Stress

Many social psychological models of adjustment assign a central role to the concept of stress. Such models identify aspects of stressful stimuli, such as controllability, as important predictors of health outcomes. Other models identify aspects of individuals or their social environments that make them more or less vulnerable to the negative effects of stress. Obviously, empirical tests of these models require the manipulation or measurement of stressful life circumstances.

Laboratory studies of stress in humans have employed a wide variety of well controlled aversive or threatening stimuli over the years (Lazarus & Folkman, 1984). Threats of physical discomfort (e.g., electric shock) or self-esteem threats (e.g., intelligence tests, etc.) successfully arouse the physiological, affective, and even behavioral signs of stress responses. The precise experimental control available in the laboratory and the consequent gains in internal validity are presumed to permit unambiguous interpretation of the research findings obtained in such studies.

It is also true, however, that laboratory studies typically do not entail severe or prolonged stressful stimulation. Thus, laboratory studies typically do not include manipulations of stress equivalent to those discussed in many models of emotional and physical disorder. Experimental operations may not adequately capture the theoretical constructs, and generalization to more severe stressors is unknown. This is one important incentive to study naturally occurring stressful life circumstances. Most of the work of this type has involved the measurement of stressful life events (e.g., death of a loved one, divorce, loss of a job, etc.) and the examination of the association between levels of life stress and either emotional or physical disorder. Several self-report instruments are now in wide use for measuring the number and even perceived stressfulness of recent life events (Monroe, 1982; Perkins, 1982).

As straightforward as such procedures might seem, they are the source of considerable controversy. Several authors have suggested that many of the individual stressful events might be the result of emotional or physical disorder (e.g., breakup of a relationship, loss of a job) rather than an independent cause (Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1982; Schroeder & Costa, 1984). Thus, concurrent associations between self-reports of stressful life events and adjustment may reflect precisely the opposite causal association as is central to most of these models.

A very similar criticism suggests that reports of stressful life events may be symptoms of psychological disorder. In an interesting test of this notion, Dohrenwend, Dohrenwend, Dodson, and Shout (1984) asked a large sample of clinical psychologists to rate the items in several measures of stressful life events as to the extent to which they reflected symptoms of psychological disorder. Many of the items were rated as reflecting emotional disorder, and this confounding of predictor and outcome was particularly true of measures of minor stressful events (i.e., daily hassles; Delongis, Coyne, Dakoff, Folkman, & Lazarus, 1982). Thus, the correlations between measures of stressful events and emotional outcomes may be due, at least in part, to the fact that the two sets of research operations assess a single construct.

In a very important discussion of this issue, Depue and Monroe (1985) have underscored the possible role of chronic disorders and enduring personality characteristics in the life events–symptoms association. Individuals with chronic psychological disorders or high levels of certain personality characteristics such as neuroticism or negative affectivity are likely to report high levels of life stress and current emotional distress. Neuroticism or negative affectivity refers to a stable individual difference in the tendency to experience negative moods. As a result, the association between stressful event scores and current distress may reflect the operation of a third variable—the presence versus absence of chronic disorder or individual differences in neuroticism.

The problems of reverse causal direction, confounding of measures of events and symptoms, and chronic disorder or personality are all obviously serious threats to the validity of concurrent correlational studies of stressful life events. Despite these interpretive ambiguities, such studies are common at the interface of social and clinical psychology. Given the limitations of such designs, the results do not provide compelling evidence about the accuracy of the models being tested. The assessment and statistical control of chronic conditions and relevant personality dispositions seems essential in concurrent studies. Prospec-
tive studies, also including measurement and statistical control of possible confounding factors, would produce a substantially less ambiguous and therefore more compelling form of evidence.

Current models of stress and adaptation often emphasize the importance of the individual's appraisal of the stimulus (Lazarus & Folkman, 1984). Similar events may not be equally stressful for different people because of differences in the level of subjectively perceived harm or threat. A large body of research is consistent with this view. As a result, many stressful life event surveys (e.g., Sarason, Johnson, & Siegel, 1978) ask respondents not only to indicate whether or not they have experienced the listed events but also to rate the valence or impact of the event. While this approach is consistent with cognitive mediation models of stress and adaptation, it also has been criticized as confounding the event measure with the outcome under study (Dohrenwend et al., 1982). Subjective impact ratings may reflect emotional distress, thereby artificially inflating the association between predictor and outcome.

The potential gains from life stress research are substantial. As a result, this general approach certainly should remain as a methodological tool in testing social-clinical models. However, its uncritical use will add little truly useful information. Refined methodologies that attempt to rule out the alternative interpretations we have discussed will hopefully become more commonplace. Studies of predictable, naturally occurring stressful events such as childbirth or elective surgery provide an alternative way to operationalize the construct of stress in prospective designs with less interpretive ambiguity.

Studies of the Social Environment

One of the most widely studied moderators of the effects of stress is social support. As in the case of life stress research, the most common approach to operationalizing this variable has been self-report inventories (Heitzman & Kaplan, 1988). Respondents rate the quantity and/or quality of supportive social relationships available to them. Similar interpretive difficulties as those found in studies of life stress are associated with most studies involving self-reported social support. For example, in the Dohrenwend et al. (1984) study, clinical psychologists rated many social support scale items as likely to be symptoms of emotional disorder. Further, Depue and Monroe (1986) argued that subjects who score low on reported social support and high on measures of stressful life events are likely to be suffering from chronic disorder or to display high levels of neuroticism or negative affectivity. Because both life stress and social support are likely to be correlated with these stable subject characteristics, people scoring high on both types of inventories as opposed to only one are more likely to display some chronically dysphoric condition. Thus, the cell of the high versus low life stress by high versus low social support factorial design that is predicted to display the highest level of distress is also the most likely to include chronically disordered or high negative affectivity subjects.

The self-report methodology and the measured as opposed to manipulated nature of the social support variable are the source of these interpretive ambiguities. Descriptions of the social environment as unsupportive may reflect personality or chronic disorder rather than the actual situation. Even if veridical, however, such reports may still reflect enduring characteristics of the person rather than the independent influence of the social environment. As a result, studies of social support or related environmental characteristics must articulate and rule out these alternative explanations as best as possible.

There is an obvious intuitive appeal and a growing body of empirical support for the notion that a supportive social network has salubrious effects on emotional and physical health (Cohen & Wills, 1985; House, Landis, & Umberson, 1988). Researchers testing social-clinical models involving this construct must be aware of the interpretive difficulties associated with the most commonly used approaches. Although cumbersome, random assignment to supportive relations is possible. Two recent studies of expectant mothers, for example, found that random assignment to a condition in which a previously unacquainted, supportive companion was available throughout the labor and delivery process produced a 50% reduction in delivery complications and a similarly large reduction in the duration of the process, even among uncomplicated births (Sosa, Kennell, Klaus, Robertson, & Urrutia, 1980; Klaus, Kennell, Robertson, & Sosa, 1986). Thus, experimental approaches to the study of social support during significant stress are available.

Studies of Personality and Individual Differences

A common theme running through much of the social-clinical literature is that certain aspects of personality increase vulnerability to the develop-
ment of emotional or physical dysfunction. Type A behavior, attributional style, and pessimism are only a few of the relevant examples of this general paradigm. Empirical tests of these models are valid only to the extent that the measures of the personality characteristic under study are valid assessment techniques. If scores on the individual difference assessment instrument do not reflect the construct under study, then the model has not been tested and results are not relevant to the theory. If the construct validity of the relevant assessment device is unknown or incompletely established, then obviously the relationship of research findings using the device to the theory is unknown. Evaluation of the construct validity of personality and individual difference assessment procedures has long been an important aspect of clinical psychology. These fundamental tenets of personality research methodology are, surprisingly, a source of difficulty in current social-clinical research.

As we discussed in our chapter on Type A behavior in this volume, the lack of traditional construct validity studies posed a problem in the early years of Type A research. The major measures of this construct are quite modestly related and show quite distinct patterns of association to other personality dimensions. Yet, early reviews of this literature pooled findings using this array of assessment procedures to evaluate the overall relationship between the construct of Type A behavior and coronary heart disease. Subsequent research has explored and even benefited from the differences between purported measures of this single construct. It is also true, however, that some of the waxing and waning of acceptance of the association between Type A and heart disease and the resulting confusion and controversy in both scientific and popular literature can be directly attributed to a lack of thorough, a priori construct validation research.

More recent entries into the list of individual difference dimensions in social-clinical research also suffer from this problem of incomplete construct validation. One example is the construct of optimism and the device intended to assess this dimension, the Life Orientation Test (LOT; Scheier & Carver, 1985). Consistent with self-regulation models of the process of coping and adjustment (Carver & Scheier, 1982), individual differences in optimism are important when confronting taxing situations. Optimists, given their positive expectations, are likely to cope actively with the difficulty and as a result are less likely to suffer the emotional and physical consequences of stress. In contrast, pessimists, given their negative expectations, are likely to cope regressively when confronted with difficulty. Given this maladaptive response, they are more likely to suffer the emotional and physical effects of stress.

Scheier and Carver (1985) developed the eight-item LOT to assess individual differences in optimism-pessimism. This self-report scale correlates as expected with self-reports of coping behavior and physical symptoms (Scheier & Carver, 1985; Scheier, Weintraub, & Carver, 1986). As for evaluations of the construct validity of the scale, Scheier and Carver (1985, 1987) reported that moderate correlations with measures of depression, hopelessness, and similar dimensions provide evidence of convergent and discriminant validity. The authors argued that the correlations are large enough to be meaningful, (i.e., convergent validity), but not so large as to indicate that the LOT is assessing the same dimensions as assessed by the other scales (i.e., discriminant validity).

Traditional procedures for evaluating the convergent and discriminant validity of assessment devices are more specific than those employed by Scheier and Carver. At least two measures of at least two distinct constructs are employed (cf. Campbell & Fiske, 1959). Convergent validity is demonstrated by a large correlation between two measures of a given trait, while discriminant validity is demonstrated by correlations between two measures of separate traits that are smaller than the correlations between the measures of a single trait.

In a recent study, we evaluated the construct validity of the LOT using this more traditional approach (Smith, Pope, Rhodewalt, & Poulton, 1989). We selected the Generalized Expectancy of Success Scale (GES; Fidel & Hale, 1978) as a second measure of optimistic expectations about the future. The construct described by these authors is very similar to optimism as articulated by Scheier and Carver (1985). The second trait selected for the validation procedure was neuroticism (Eysenck & Eysenck, 1964) or negative affectivity (Watson & Clark, 1984). We selected this dimension because it has similar self-report coping and physical symptom correlates as those reported for the LOT (Costa & McCrae, 1987; McCrae & Costa, 1986; Vitalliano et al., 1987; Watson & Pennebaker, 1989). Thus, if the LOT displayed poor discriminant validity in relation to measures of neuroticism, then it is possible that the previously reported findings regarding coping
and health reports reflect this older, more established personality trait as opposed to the construct of optimism.

In three separate samples, the LOT and GESS were significantly correlated (average $r = .56$). However, the LOT was at least as closely correlated with measures of neuroticism (average $r = -.62$). Thus, given this lack of discriminant validity, it could be argued that the LOT is simply an inversely scored measure of neuroticism. Although we were able to replicate the previously described correlations of the LOT with coping and symptom reports, these associations were largely eliminated when neuroticism was controlled through partial correlation. Control of optimism scores, however, did not eliminate the significant correlations of neuroticism with coping and symptom reports. Therefore, it is quite plausible that the earlier findings reflect the trait of neuroticism rather than optimism.

Obviously, our results do not challenge the conceptual model that generated this line of research on optimism. Rather, our findings challenge the interpretation of the previous studies as supporting that model. Because of the questionable construct validity of the measure of optimism, the previous results are not clearly relevant to the model.

It would be misleading to imply that the LOT is a unique example of this problem. For example, very similar critiques have been presented concerning the construct of hardiness (Kobasa, 1979). Scales purported to assess this personality scale have been criticized for related measurement problems, with consequent ambiguities regarding the interpretation of previous tests of hardiness theory (Allred & Smith, 1989; Funk & Houston, 1987; Rhodewalt & Zone, 1989). Even though they have not been criticized, measures of other individual difference dimensions have not been subjected to thorough construct validation research. For example, Peterson, Seligman, and their colleagues have reported several very impressive prospective studies of the emotional and physical health consequences of explanatory style (Chapter 14). Even though multiple methods are available for assessing this construct, the obvious convergent-discriminant validation studies have not been reported. Without such independent evidence of construct validity, interpretation of this impressive set of findings as reflecting the effects of explanatory style is tentative at best.

Many researchers have called for greater empirical scrutiny of measures of personality in studies of emotional and physical health, and most of these authors have explicitly suggested that new measures be validated in reference to measures of established dimensions of personality such as neuroticism (Costa & McCrae, 1987; Depue & Monroe, 1986; Holroyd & Coyne, 1987; Watson & Pennebaker, 1989). This type of research would not only facilitate the interpretation of many existing and future studies, it would also “reduce the risk that we would reinvent constructs under new labels” (Holroyd & Coyne, 1987, p. 367). In any case, researchers who develop or employ individual difference measures should thoroughly evaluate and establish the construct validity of their trait measures. Incomplete efforts clearly undermine the value of the accumulated findings.

**Studies of Cognitive Processes**

As is the case in much of psychology, much of the activity at the social-clinical interface has a strong cognitive orientation (Brehm & Smith, 1986). Emotional and physical health are seen as influenced by a variety of cognitive processes. The research and theory regarding optimism or explanatory style, for example, could easily be viewed as representing cognitive models. Attributing a causal role to cognitive constructs creates a special challenge for the investigator. Valid methods of manipulating or measuring cognitive processes must be developed.

Many tests of cognitive models of adjustment are by necessity correlational. Predicted differences on a relevant cognitive dimension are tested in comparisons of functional and dysfunctional groups, such as nondespressed and depressed college students. These groups are often initially distinguished by selection of extreme scores on self-report measures of the dysfunctional condition. Unfortunately, the cognitive process under study also is often operationalized as scores on a self-report measure. For example, depressed and non-depressed groups might be compared as to their scores on a measure of dysfunctional attitudes, or socially anxious and nonanxious groups might be compared as to their scores on a paper-and-pencil measure of anxious self-statements. Coyne and Gotlib (1983) have pointed out that such studies often contain a “thinly veiled tautology” (p. 46). Groups that are distinguished on the basis of their relative endorsement of questionnaire items purported to measure the dysfunctional condition (e.g., depression, social anxiety, etc.) are com-
pared on their relative endorsement of very similarly worded items intended to assess a cognitive process. Rather than a correlation between a dysfunctional emotional condition or disorder and a cognitive process or structure, such correlations may simply reflect the convergence of scores on similarly worded questionnaires (cf. Nicholls, Licht, & Pearl, 1982).

Interpretive ambiguities associated with studies of this type are not unique to the social-clinical interface. They are common in the literature testing many cognitive models of emotional disorder (Coyne & Gotlib, 1983; Smith, 1989b). At the very least, the producers and consumers of this type of cognitive research should explicitly acknowledge the interpretive problems inherent in such studies, and when possible the alternative explanations should be tested directly (e.g., Zurawski & Smith, 1987).

Researchers also should consider alternatives to self-report assessments of cognitive constructs. A variety of information-processing paradigms have been employed to test cognitive models of dysfunctional conditions (Ingram, 1986). Such studies have often produced results consistent with cognitive models of depression (Ingram, Smith, & Brehm, 1983), social anxiety (Smith, Ingram, & Brehm, 1983), and Type A behavior (Stern, Harris, & Elvenum, 1981) without the alternative interpretations associated with self-reports of cognition. Such procedures also circumvent difficulties arising from the influence of cognitive processes that are not easily tapped by verbal reports. The rapidly growing literature on social cognition (Taylor & Fiske, 1984) contains many experimental paradigms with demonstrated or at least potential utility for testing cognitive models of adjustment.

**Operationalizing Outcomes**

We stated earlier that tests of social psychological models of adjustment are useful only to the extent that the specific operational definitions of the predictors or causes of dysfunction measure or manipulate the construct these operations are intended to reflect. The same is true of the other end of the equation. These empirical tests are useful only to the extent that the operations used in measuring or classifying dysfunctional states or conditions have sufficient construct validity. The social-clinical literature includes several threats to our ability to interpret research operations as reflecting the adaptational constructs under study.

**Emotional Functioning**

Perhaps the largest part of the social-clinical literature consists of studies of dysfunctional emotional conditions such as depression and anxiety. Studies testing social psychological models of emotional adjustment often contain measurement problems that limit their contribution.

A common problem concerns the issue of the continuity of emotional disorders along the dimension of severity. Many studies test models of maladjustment by comparing groups of college students or groups from other essentially normal populations. The groups differ on the basis of their scores on a measure of a dysfunctional state, such as the Beck (1967) Depression Inventory (BDI). One possible problem with this procedure is that the difference in depression levels between groups of college students with high versus low BDI scores may be much smaller than the difference between clinically depressed and normal groups. Insufficient variability on the independent variable would pose a threat to statistical conclusion validity in the Cook and Campbell (1979) scheme. However, the continuity problem is somewhat different. Simply put, the question of continuity concerns whether or not the difference between groups selected in this manner resembles the difference between a group of clinically dysfunctional individuals and a group of demographically similar normal controls not simply in magnitude but in form. In the continuous model, the assumption is that the causes of mild dysfunctional conditions are similar to the causes of severe, clinically significant conditions. The discontinuous model suggests that the outward similarity between the two conditions is misleading, and in fact the subclinical and clinical dysfunctions represent different phenomena with distinct etiologies.

It is now generally agreed among psychopathology researchers that the continuity of seemingly similar dysfunctional states varying in outward severity might be fruitfully tested but should rarely if ever be assumed (Depue & Monroe, 1983, 1986). Analog studies have an obvious heuristic purpose, but are insufficient if a model of a clinical disorder is to be tested. Unfortunately, many social psychological models of emotional adjustment are tested only with college student populations. Although studies of college students selected on the basis of high scores on measures of social anxiety or depression might be relevant to our understanding of clinically significant social
phobia or affective disorder, we cannot assume this to be true.

One response to this situation is to note that variations in emotional distress within the normal range are a worthy topic of study in their own right. The emotional problems of essentially healthy, adjusted individuals are certainly not trivial. Therefore, one could solve the problem posed by the threat of discontinuity by avoiding even the temptation to speculate about, let alone generalize to, clinical populations.

However, this solution certainly undermines the potential utility of many social psychological models of emotional functioning, and probably does not honestly reflect the ambitions of most social-clinical researchers. If the social-clinical integration is to move from an innovation to more lasting, mainstream contribution, promising analog research should be followed by more difficult studies of clinical populations. Although both types of research are useful, a lasting impact for the interface will require a solid foundation of clinical research.

A second major challenge to the validity of outcome measurements concerns the problem of specificity. If a model of depression is being tested, depressed and nondepressed groups are typically compared. If a model of social anxiety is under scrutiny, anxious and nonanxious groups are compared. Predicted differences are interpreted as supporting the model of a particular emotional problem under consideration.

Potential problems arise due to the fact that individuals scoring high on measures of one dysfunctional condition (e.g., depression) often have similarly extreme scores on measures of other emotional disorders (e.g., anxiety). Several studies have demonstrated that measures of different maladaptive emotions such as anxiety, depression, and others are so closely correlated in normal samples as to be virtually indistinguishable (Gotlib, 1984; Tanaka-Matsumi & Kameoka, 1986; Watson & Clark, 1984). As a result, groups selected on the basis of depression scores are likely to differ by almost as much, if not just as much, on anxiety and other characteristics. The obtained correlates of depression are quite likely to be just as closely related to anxiety.

The fact that the correlates of a particular emotional condition are not specific to that type of dysfunction may or may not be a problem depending on the theory being evaluated. If the theory is nonspecific (i.e., it concerns emotional ad-justment broadly defined), then no difficulty is present. If, however, the theory purports to explain a specific form of emotional disorder, then the findings must reflect that level of specificity if they are to be considered supportive. For example, the reformulated learned helplessness model (Abramson et al., 1978) concerns depression as opposed to more general forms of distress. Some studies have found the predicted attributional style to be present in clinically depressed but not other clinical groups (Heimberg, Vermilyea, Dodge, Becker, & Barlow, 1987; Johnson, Petzel, & Munic, 1986), suggesting that depressive attributional style is specific to depression. Other research, however, has suggested that this attributional style is a correlate of global psychopathology in clinical populations rather than being unique to depression (Hamilton & Abramson, 1983; Miller, Klee, & Norman, 1982; Heimberg, Klosko, Dodge, Shadick, Becker, & Barlow, 1989).

Thus, if the theory to be tested makes predictions about a specific disorder, then additional controls are needed. Two-group comparisons are inadequate if the results are to be interpreted as relevant to a single type of pathology. If correlational designs are used literally, correlations with other dysfunctional dimensions must be tested simultaneously for specific interpretations to be justified. We should note that the failure to demonstrate predicted specificity may reflect poor discriminant validity of the measures of emotional functioning rather than the absence of unique effects. Thus, researchers must be cautious to avoid falsely rejecting specificity hypotheses.

One final difficulty in measuring emotional adjustment concerns interface research in medical populations. Several social psychological models have been developed to explain the development of depression or other emotional adjustment problems in the medically ill. This common and important clinical problem is certainly worthy of study, and the social psychological models offered to date are quite promising.

The primary threat to the validity of studies addressing this problem concerns the use of depression measures developed in psychiatric samples. In psychiatric populations, somatic symptoms such as lethargy, difficulty concentrating, concerns over physical appearance, reduced libido, and changes in appetite typically reflect depression. Endorsement of related items on self-report inventories validly reflect depression in this instance. In the medically ill, however, these
symptoms may reflect the illness itself or the side effects of medication. As a result, scores on inventories including such items are likely to overestimate the severity of depression and create ambiguity in the interpretation of effects involving depression scores (e.g., Pincus, Callahan, Bradley, Vaughn, & Wolfe, 1986; Peck, Smith, Ward, & Milano, 1989).

**Physical Functioning**

Increasing amounts of the conceptual and empirical activity at the social-clinical interface concern physical rather than emotional health. As a result, the measurement of physical health outcomes becomes an essential methodological challenge. A key, often neglected distinction in operationalizing physical health is the difference between illness and illness behavior (Cohen, 1979; Mechanic, 1972). Illness refers to the tissue or physiological changes indicative of disease. Fever, infection, and elevated blood pressure are all examples of possible measures of illness. Illness behavior, in contrast, refers to all of the things people might do when they are ill, such as visit a physician, report symptoms or stay in bed.

Illness and illness behavior are obviously correlated. This is true logically, but there is also empirical evidence to indicate that subjective health reports are concurrently and prospectively related to objectively measured health status (Kaplan & Comacho, 1983). However, illness and illness behavior are far from perfectly correlated. Stoic individuals deny symptoms and refuse to see a physician even when clearly ill. Hypochondriacal persons, in contrast, report many symptoms and frequently seek medical attention when they are well. Thus, while measures of illness and illness behavior share some variance, they have a considerable amount of unique variance as well.

The problem arises when we use measures of illness behavior to operationalize physical health. It is easy to see why researchers might substitute measures of illness behavior for actual measures of illness. The two types of measures are related, and illness behavior is likely to be easier and less expensive to assess. However, there is no guarantee that the portion of illness behavior variance accounted for by predictor variables in a study is the same portion of variance in illness behavior scores that overlaps with actual illness. The findings observed might reflect the variance in illness behavior attributable to actual illness, but might instead reflect the variance in illness behavior that is independent of actual illness. If the latter is true, it would obviously be inappropriate to interpret the operational outcome as reflecting the outcome construct of physical health. Interpretive ambiguities arise from the fact that if only illness behavior is assessed, then it is virtually impossible to determine whether or not the covariation between predictors and outcome involves actual illness.

This problem becomes a larger threat when one considers the content of many social-clinical models of physical health. These models often include the constructs of life stress, social support, and a variety of personality dimensions. As mentioned in our discussion of operationalizing each of these classes of predictor variables, measures of these constructs are often substantially correlated with the personality trait of neuroticism or negative affectivity. Recent research suggests that this trait, in turn, is consistently related to somatic complaints or physical symptom reports but not actual illness (Costa & McCrae, 1987; Watson & Pennebaker, 1989). Therefore, associations between many predictor measures and health measures may reflect a correlation between neuroticism and somatic complaints in the absence of actual illness. Obviously, this possibility requires reinterpretation or at least reconsideration of associations of life stress, social support, and many personality traits with symptom reporting. Further, if social psychological models of influences on physical health are to be tested in a more compelling fashion, the possible confounding effects of neuroticism must be considered.

The cautionary note should not be taken to indicate that illness behavior in the absence of actual illness is simply a nuisance variable. Such behavior is extremely interesting, important, and even costly. As a result, it is certainly worth researching. Rather, we simply wish to point out the problems inherent in using measures of illness behavior to operationalize physical illness.

**Testing Associations Between Predictors and Outcomes**

From our discussion so far, it is clear that the value of social-clinical research depends on how predictors and outcomes are operationalized. The value is also influenced by the nature of the tests of their association. Here again, this statement is perhaps both obvious and far from unique to the social-clinical interface. Yet, given many of the research strategies employed in this area, this issue
and its incumbent challenges are worth reviewing. We are not referring to basic threats to statistical conclusion validity (Cook & Campbell, 1979), such as low power, experimentwise error rates, and the reliability of outcome measures and treatment implementation. Rather, we are referring to the general approach to hypothesis testing in many social-clinical studies.

Theoretical Risks

It is perhaps unnecessary to point out that theories—or the less assuming structures, models—are never proven correct. Rather, the strongest statement one can make about the extent to which a theory has been corroborated is that it has been subjected to serious risk of disconfirmation and has survived. Popper (1959) argued that the degree of corroboration depends on “the severity of the various tests to which the hypothesis in question can be, and has been, subjected” (p. 267). The severity of the test depends, in part, on how likely it is that a prediction derived from the theory will be disconfirmed. Severe tests subject the prediction to clear risk of disconfirmation. Simple directional predictions would seem too sufficiently clear as to obviously risk refutation.

However, many such simple, directional tests provide a rather illusory form of risk of refutation. In his critique of “soft psychology” (the social-clinical interface is an example of this form of psychology), Meehl (1978) argued that one factor inhibiting the systematic accumulation of compelling theory-based scientific knowledge is the fact that hypotheses are rarely confronted with a “grave risk of refutation” (p. 821). For example, in testing the attributional reformulation of the learned helplessness model, a researcher might simply predict that higher scores on the measure of dysfunctional attributional style will be associated with higher scores on a depression scale. When one considers the host of methodological artifacts, potential confounds, as well as alternative processes that would possibly produce the hypothesized effect, it is clear that the effect is quite likely to occur, even if the theory is false. A supportive result is a very weak corroboration of the theory, because of its high prior probability given its many possible explanations. Thus, support for simple predictions often does not provide much in the way of corroboration because of their high prior probability of occurrence from many factors other than those specified by the theory.

Meehl (1978) argued that more complex or precise predictions entail a greater risk of refutation and provide more potential corroboration as a result. If the theory generates the prediction that depressed subjects should differ in their explanations of negative events from anxious subjects who will, in turn, be similar to normal controls, a somewhat larger degree of corroboration results from a predicted finding. A somewhat smaller pool of possible explanations produces a smaller prior probability of the pattern of differences. An even greater degree of corroboration results from a predicted finding indicating that these differences in attributions across groups are precisely opposite from a different class of events. A more complex prediction, with a resulting lower prior probability, is more useful.

Meehl’s (1978) suggestion that psychology is often satisfied with increasing numbers of “tabular asterisks” (i.e., significant statistical tests of simple directional predictions) is no less true of research at the social-clinical interface. His recommendation that the accumulated body of research will be more valuable if more complex predictions are tested and theories, as a result, placed at greater risk is also potentially quite helpful for interface research.

On a practical level, this means the articulation of more complex pattern predictions rather than simple directional tests. Interactional or nonlinear predictions are preferred over linear main effects. Although the level of complexity Meehl suggests may well be unattainable, the criticism clearly describes possible improvements. Social psychological theories can generate complex predictions, and the empirical tests should reflect this richness. An equally valuable aspect of Meehl’s advice is the acceptance of falsification when predicted effects fail to materialize. Of course, this requires the researcher to accept the validity of the operationalizations of the constructs under consideration. Such acceptance is prudent only after thorough empirical evaluations of construct validity.

Potentially Misleading Patterns

Although more complex, pattern predictions may ultimately produce a richer harvest, such patterns must not be accepted uncritically. Even a more complex effect can reflect something other than the constructs and processes under consideration. We have already referred to one example. Depue and Monroe (1986) described how the apparent interaction between life stress and social support might reflect something other than the buffering effect on emotional health. This pattern could reflect the fact that more chronically
METHODOLOGICAL CHALLENGES

disturbed people—identified through the convergence of measures of problems associated with chronic disturbance—are more likely to experience current distress than are groups without a chronic disorder.

In a study of this type, Blumenthal, Burg, Barefoot, Williams, Haney, and Zimet (1987) reported that social support moderated the concurrent association between Type A behavior and the severity of coronary artery disease. This pattern could reflect, as the authors suggest, that a supportive social environment reduces the health risk associated with Type A behavior. It is also possible, however, that the combination of low levels of social support and the presence of Type A behavior effectively identified hostile individuals. Hostile people, in turn, may have had more severe coronary disease (cf. Cohen & Matthews, 1987). Thus, more complex predictions certainly do not remove potential interpretive ambiguities.

Recognizing Tested and Untested Patterns

In many cases, the association between predictors and outcomes that is actually tested is far less complex than the model that spawned the study. The results may be interpreted, however, as supporting the more complex conceptual structure. Krantz and Hedges (1987) have noted this potential problem in discussing research on personality and health. The theoretical models in this area typically assert that a given personality characteristic is associated with a health outcome by means of an intervening psychophysiological mechanism. For example, Type A behavior may lead to coronary disease through repeated episodes of exaggerated cardiovascular reactivity. Subsequent research may establish that (a) Type A behavior predicts the development of coronary disease, (b) Type A behavior is associated with reactivity, and (c) that reactivity predicts the development of coronary disease. Even if these three features of the nomological net are clearly established, the basic theory has not been tested directly.

It is also true that some researchers claim falsification when a pattern prediction from the theory has not actually been tested. For example, many conceptual models concerning personality factors hypothesized to increase vulnerability to illness are interactional models. In the presence of certain environmental or situational factors (e.g., high levels of stress, etc.), the personality dimension is related to disease risk. The operational tests of these models often fail to assess the relevant situational factor (Matthews, 1983; Smith, 1989a). As a result, the theoretically more accurate and possibly statistically more powerful interaction effect goes untested. Thus, researchers must take care not to claim support or disconfirmation for incompletely tested models.

Mediational Analyses and Theoretical Risk

Interactional predictions in which one independent variable moderates the effects of a second are one way to increase theoretical risk and consequent scientific yield. Moderator effects are, by definition, statistical interactions. As a result, the more complex prediction produces greater risk and potential corroboration. A second way involves tests of models where one variable mediates the effects of a second (Baron & Kenny, 1986). Mediational statistical techniques can test predictions regarding how predictors influence outcomes. The predicted set of associations is inherently more complex, and as a result a greater degree of corroboration is possible. In many instances mediational variables are specified in social psychological theories of adjustment, and recent developments in data analytic techniques make such hypotheses easily testable.

TESTING MODELS OF DIAGNOSIS AND TREATMENT

Social psychological models are not only useful in suggesting possible etiological and maintaining factors in adjustment difficulties. Many conceptual approaches to the clinical processes of assessment and intervention have been offered in recent years. Research in these areas is, unfortunately, less common than are studies of emotional and physical dysfunction. This is unfortunate because the potential practical benefits of social psychological models in this area are substantial. As noted above, however, the current zeitgeist in clinical psychology emphasizes empirical tests of such models in clinical populations. For the potential benefits of social psychological approaches to assessment and treatment to be realized, additional research is necessary. Such research, in turn, must address several methodological challenges.

Studies of Clinical Assessment

From a historical perspective, assessment has been perhaps the most central activity of clinical psychologists. As described by many authors, clinical assessment is essentially a special case of social judgment. Evidence is assembled and re-
viewed, dispositional characteristics are attributed to the person assessed, and predictions about future behavior are offered. Meehl (1960) was one of the earliest proponents of subjecting the cognitive activity of the clinician to empirical study. The research that has accumulated to date documents the fact that despite their professional training, clinicians often make similar mistakes in social judgment as those observed in nonclinical social knowing (see Chapter 21). Their active collection and recall of clinical evidence is likely to be influenced by their hypotheses to produce a confirmatory bias. Their judgments are susceptible to anchoring effects and illusory correlations, and they are prone to attribute client characteristics to dispositional rather than situational factors. Additional research would certainly be useful in readdressing some of these drawbacks in the clinical assessment process.

Many studies of clinicians' cognitive processes in assessment contexts have employed analogue samples and role-play procedures. Obviously, studies of actual clinicians provide more compelling and relevant information. Interestingly, some studies have found that actual clinicians are more prone to cognitive errors in the assessment process than are untrained, naive subjects (e.g., Friedlander & Phillips, 1984). Other studies indicate that some biases, such as confirmatory hypothesis-testing strategies, actually increase with additional clinical experience (Hirsch & Stone, 1983). Thus, analog studies may underestimate the problematic features of social cognition in clinical assessment. As a result, studies of actual clinicians are important.

It may be useful to duplicate clinicians' typical working environments in studies of this type. Many clinicians, for example, must cope with large caseloads. Increasing the information-processing demands in clinical assessment tasks has been found to produce greater evidence of cognitive errors, such as illusory correlations (Leugner & Petzel, 1979). Increased ecological validity of studies of clinicians' judgments, therefore, may underscore the importance of the problem and its remediation.

One final methodological challenge to social psychological studies of the assessment process concerns the interactive nature of actual clinical assessment. Initial impressions or even the wording of referral questions may lead to the selection of some assessment procedures over others or some interview questions over other lines of inquiry. The resulting additional information elicited from the client shapes additional assessment selections in an at least temporarily ongoing interactive cycle. In contrast, many studies of the social psychology of clinical judgment are more static. A set of materials is presented, with some feature varied across experimental conditions. Although such studies are certainly valuable, this approach also may fail to capture the extent of the problem. Confirmatory hypothesis testing in an ongoing interaction as opposed to a more static review of test protocols is likely to produce compounding errors with stronger ultimate effects. Also, as the amount of available clinical material increases, clinicians' confidence in the accuracy of their judgments outstrips their actual accuracy (Oskamp, 1965). Thus, open-ended, clinician-guided assessment procedures seem to present the greatest opportunity for problems. Once again, the study of actual clinicians in as ecologically valid a setting and process as possible would be very useful.

Studies of Clinical Intervention

Many of the conceptual contributions of social psychology to clinical practice have been directed toward a better understanding and increased effectiveness of psychotherapy and behavior-change procedures. The empirical evaluation of the effectiveness of therapeutic interventions is an increasingly important aspect of clinical psychology, as is the study of therapy process. The result of these trends is that the potential contribution of social psychological models is great, but the methodological price of admission is steep.

Analog Studies of Psychotherapy

Large controlled trials of psychotherapy and related behavior-change techniques using clinical samples seeking treatment and experienced therapists are time consuming, expensive, and cumbersome. Given these disincentives, it is not surprising that many therapy studies are conducted as analogs of actual intervention procedures. Kazdin's (1978) discussion of the potential differences between analog therapy studies and more realistic procedures is essential reading for all researchers considering such an approach for testing a social psychological model of behavior change. Many times less severe target problems are addressed, college student populations are studied, subjects are recruited through incentives other than a desire to reduce their distress, and less experienced therapists are typically used. Whether or not these
differences from actual clinical settings and procedures preclude generalization from analog research to clinical practice is an empirical question. Unfortunately, there are many more analog studies than there are answers to the empirical question about external validity.

Certainly analog therapy research has many uses, such as initial hypothesis testing. However, the potential impact of social psychological models will be reduced if interface researchers do not address the generalizability question directly. Psychotherapy researchers and practicing clinicians are much more likely to be persuaded by more realistic therapy studies than by analog research. Certainly research along the entire continuum from basic social psychological laboratory research to comparative psychotherapy trials involving social psychological models is useful (Brehm & Smith, 1982). Nevertheless, the future and potential contribution of the social-clinical interface would be well served by pushing the activity farther along the continuum more frequently.

Testing Social Psychological Therapies

As mentioned above, psychotherapy research has become increasingly more sophisticated in recent years. The central issues in this area are relevant to any researcher wishing to test a psychotherapy or related interventions derived from social psychological principles.

One basic issue concerns the selection of appropriate comparison or control conditions. At a simple level, much psychotherapy research asks the question, "Is treatment X effective?" An important follow-up question is, "Compared to what?" Historically, psychotherapy research design has emphasized controls for the natural history of the disorder under treatment (i.e., spontaneous remission) and for the nonspecific factors common to virtually all therapies (e.g., attention, expectancy for improvement, etc.). As a result, two logical choices for control groups were "no treatment" or "waiting list" controls and "nonspecific factors" or "placebo" controls (Hersen, Michelson, & Bellack, 1984).

Current discussions of appropriate comparison groups have questioned this long-standing practice and the associated conceptual models (Horvath, 1988; Parloff, 1986; Basham, 1986). Withholding treatment and administration of inert therapies presented as effective raise unpleasant ethical questions. It is also relatively rare that the question, "Is this therapy better than nothing?" is important given the large treatment literatures amassed for most disorders. Further, one researcher's "nonspecific factor" is another's innovative social psychological active ingredient. As a result, needed tests of social psychologically oriented therapies must carefully articulate the question to be asked, ensure that it is useful given the existing treatment literature, and select the necessary comparison condition so that groups differ only in the key construct under consideration.

Another recent development in psychotherapy research is the focus on the clinical significance of change (Jacobsen & Reuensté, 1988). Statistical differences may answer important conceptual questions, but in this arena the practical importance of observed changes is equally important. Evaluating the magnitude and meaning of observed changes is an evolving topic of clinical research, and social-clinical studies of psychotherapy should at least be informed by this work.

Social Psychological Studies of Therapy Process

Unique therapies may be derived from social psychological models, and they can be tested meaningfully by the methods mentioned above. A different role for interface models, however, involves explanation of the mechanisms that underlie the effectiveness of existing interventions. Mediational models and associated analytic techniques are an established aspect of current social psychology (Baron & Kenny, 1986). Similar conceptual and statistical approaches are extremely relevant to the explication of psychotherapy process (e.g., Hollon, DeRubeis, & Evans, 1987).

The controlled evaluation of therapies derived from social psychological concepts is unnecessary in this pursuit. Rather, social psychological constructs hypothesized to mediate therapeutic effectiveness are assessed in the context of traditional therapy research designs, and included in appropriate mediational statistical tests.

CONCLUSIONS AND CURRENT ISSUES

Sound and increasingly ambitious research is necessary to sustain the impressive integration of social and clinical psychology observed in recent decades. Theories must be pressed to yield more precise and complex, falsifiable predictions. The operational definitions must be increasingly compelling in order that both positive and negative results have greater utility. Promising and intrigu-
ing results from analog research should be followed by conceptual replications of more direct clinical relevance, and the cumbersome complexity of research problems in psychopathology, psychotherapy, and clinical assessment must be confronted and managed rather than avoided. This tall order must be met if the integration of social and clinical psychology is to continue its movement from a primarily academic activity to an influence on mainstream clinical psychology and actual clinical practical.

Of course, there are already many examples of the kinds of social-clinical studies we are recommending. Some are decades old and others are quite recent. Thus, rather than suggesting a new type of research activity at the interface, we are simply recommending an increase in the proportion of research of this more involved and informative type.

Although they are not strictly methodological, there are several related issues that we have not addressed. One issue concerns the direction of the currents flowing at the interface. Our discussion has largely focused on generating more compelling tests of social psychological models of clinical problems and procedures. This type of interdisciplinary activity has had an important impact and will continue to do so. It is also true, however, that social psychological paradigms can offer useful tools for testing rich conceptual traditions in clinical psychology. For example, the concepts and methods of attribution research and specifically self-handicapping (Jones & Berglas, 1978) provided an opportunity to test Adler’s decades-old notions about the strategic use of symptoms (Snyder & Smith, 1982). Recent advances in cognitive psychology provide an opportunity to operationalize many psychodynamic constructs pertaining to unconscious processes (Kihlstrom, 1987). Current cognitive-social approaches to personality (e.g., Cantor & Kihlstrom, 1981) similarly provide methods for testing a long theoretical tradition of interactional approaches to dysfunctional behavior (e.g., Sullivan, 1953; Carson, 1969; Anchin & Kiesler, 1982). Thus, social and clinical psychology both bring rich reserves of theory that when explored with the other discipline’s research methods can yield a wealth of new information.

Even if the enormous amount of empirical activity we have outlined actually occurs, the explanation and remediation of dysfunctional conditions will not become a precise science. As a result, we must acknowledge that the more compelling and difficult research we have discussed must be combined with an additional type of interface activity. In discussing the relationship between the science of psychology and the explanation and alteration of individual behavior, Manicas and Secord (1983) pointed out that empirically substantiated theories must be translated to have relevance for specific situations. This translation, of course, is based on the empirical science, but requires a great deal of additional historical, social, biographical and even biological information.

The methodological challenges we have discussed might help to refine and extend the science, but they do little to produce sophistication in the technical process of application. A very different, and perhaps less easily defined, set of methodological challenges must be met to produce compelling applications. Of course, high-quality application is precisely the kind of activity displayed by many of the pioneers at the social-clinical interface (Brehm, 1976; Heller, 1963). Coming full circle then, hopefully a more compelling body of empirical research will stimulate additional, artful applications to clinical activities.

REFERENCES


