

Irvin Sam Schonfeld

Assessing Stress in Teachers: Depressive Symptoms Scales and Neutral Self-Reports of the Work Environment

The focal interest of this chapter on teacher stress is methodologic. The purpose is fourfold. First, the chapter enumerates a number of defects in existing measures of job stress in teachers and, concomitantly, other helping professionals. Second, alternative ways of measuring stress in teachers are suggested and evaluated. In the section on these alternatives, the use of depressive symptom scales in concert with more “objective” measures of the work environment is discussed. Third, an application of the proposed alternative measurement strategy is described. Finally, the wider utility of the measurement strategy is briefly described.

Measuring Stress in Teachers

In research on the effects that adverse job conditions exert on the functioning of teachers and other helping professionals, investigators have long used measures of stress and burnout. Typically, research on stress in teachers is cross-sectional, although such designs have many documented weaknesses. However, even if research on teacher stress were longitudinal in design, the defects of stress and burnout measures would still seriously impede an investigator’s ability to draw valid conclusions. The problems of stress and burnout measures are manifold.

First, the creators of stress scales often view stress as an overinclusive construct embracing the working conditions that are suspected of provoking distress in teachers, as well as the distress that those conditions are thought to provoke (DeFrank & Stroup, 1989; Dunham, 1984; Dworkin, 1988; Dworkin, Hañey, & Telschow, 1988; Farber, 1984; Fimian, 1983; Galloway, Panckhurst, Boswell, Boswell, & Green, 1984; Kyriacou & Sutcliffe, 1978, 1979; Needle, Griffen, & Svendsen, 1981; Pette-

Preparation of this chapter was supported by NIOSH/CDC Grant R01 OH02571 and PSC-CUNY Award Program Grants 667401, 668419, 669416, and 661251. I gratefully acknowledge the cooperation of the individuals who participated in this study. I thank Danqing Ruan for her computer programming and my late friend Jim Johnson for his critical comments on an earlier draft of this manuscript. I extend a special note of appreciation to George Schonfeld.

All correspondence concerning this chapter should be sent to Irvin Schonfeld, EDFN, City College of New York, New York, NY 10031.

grew & Wolf, 1982; Seiler & Pearson, 1984). An example of a commonly used stress questionnaire having such an item structure comes from Kyriacou and Sutcliffe (1978, 1979): "As a teacher, how great a source of stress are these factors? Maintaining discipline; Shortage of equipment . . ."

Without independently measuring job conditions and the distress those conditions are hypothesized to provoke, the investigator forecloses the possibility of testing causal hypotheses that link job conditions to distress. A plausible alternative to the hypothesis that specifiable job conditions cause distress is that individuals who are distressed create (or overreport) the putative stressors, a variety of B. S. Dohrenwend and B. P. Dohrenwend's (1981) event-proneness model. One version of the event-proneness model holds that teachers with preexisting depressive symptoms, because of impaired interpersonal skills, may create classroom environments conducive to student rule breaking, a suspected stressor. Students might, as a consequence of being bored in a class headed by an impaired teacher, rebel or act out aggressive feelings. The impaired teacher would, in turn, be unable to enforce rules of civility. Case study evidence from Schonfeld and Ruan (1991) illustrates the plausibility of the event-proneness model of teacher stress.

A second problem with stress and burnout measures is that they often fail to provide information on the frequency with which teachers encounter given job conditions. Kyriacou and Sutcliffe's (1978, 1979) item concerning the extent to which maintaining discipline is a "source of stress" does not ask how often the teacher, say, broke up a fight. By the same token, Kyriacou and Sutcliffe's (1978, 1979) item asking the teacher to reveal the extent to which equipment shortages are a source of stress fails to ascertain how often the teacher is actually confronted with equipment shortages. These items are thus useless in ascertaining the frequency with which teachers encounter difficult work-related conditions, factors that may plausibly be viewed as the "independent variable."

A third problem with stress and burnout measures is that they lead the investigator to vacate the role of hypothesis tester. Instead, the subject is cast into that role: The subject is often asked to identify working conditions by the distress that those conditions may promote. Stress items of this type may be well suited for pilot research in which an investigator is attempting to generate hypotheses. They are not, however, well suited for hypothesis testing. Cohen, Karmarck, and Mermelstein (1983) wrote that "there is . . . evidence that people often misattribute their feelings of stress to a particular source when that stress is actually due to another source" (p. 387). For example, a teacher may easily misattribute his or her distress to acting out behavior in one or two children when overcrowding may be a more significant factor (see Worchel & Teddlie, 1976). By the same token, Kasl (1978) pointed out that police officers, in response to questions about job stresses, often identify administrative and court-related work. They rarely mention life-threatening aspects of their jobs.

In stress research on helping professionals such as teachers, *burnout* is a term frequently used to describe a tripartite syndrome consisting of emotional exhaustion, depersonalization, and a reduced sense of personal accomplishment resulting from the task of helping unwilling or ungrateful individuals (Cunningham, 1983; Farber, 1984; Gold, 1984, 1985; Iwanicki & Schwab, 1981; Johnson, Gold, & Knepper, 1984; Malanowski & Wood, 1984; Maslach & Jackson, 1981, 1984; McIntyre, 1984; Pier-

son-Hubeny & Archambault, 1985). The Maslach Burnout Inventory (MBI; Maslach & Jackson, 1981) is an instrument that has been commonly used in stress research. It has also been a source of items for teacher stress questionnaires (Farber, 1984; Fimian, 1983; Fimian & Santoro, 1983). Exemplary items include: "I feel frustrated by my work" and "Working with people directly puts too much stress on me."

Burnout items are vulnerable to attribution errors (Schonfeld, 1990b). It is possible for a teacher to agree with a Maslach burnout item, asserting that "working with people" is stressful when a teacher has been hounded by an authoritarian supervisor or when a child with conduct difficulties is frustrating well-planned lessons. The burnout literature tends not to identify work-related factors that increase the risk of psychological distress or ill health in teachers.

Schonfeld (in press) adduced evidence for the view that burnout scales largely reflect depressive symptoms. First, he demonstrated with correlational evidence that measures of psychophysiological symptoms, perceived health, job satisfaction, low self-esteem, and motivation to continue in the teaching profession—factors having known links to burnout scales (Kahill, 1988)—are similarly related to depressive symptoms. Second, he provided rational evidence that two of the three components of burnout, emotional exhaustion and a reduced sense of personal accomplishment, constitute symptoms of depression. He argued that the third component, depersonalization (cynical feelings that helping professionals may direct toward clients), is reflective of the hostility, friction, and aversive control that characterize the interpersonal relationships of depressed individuals (Coyne, Burchill, & Stiles, 1990; Coyne, Kahn, & Gotlib, 1987).

A fourth problem with stress and burnout measures is the absence of clear evidence demonstrating the validity of the constructs. Although sometimes not the purpose of the investigators who conducted the research, they have provided evidence to suggest that burnout and stress measures overlap considerably with depressive symptom scales.

Four studies conducted in different geographic areas illustrate the link between stress or burnout measures and depressive symptoms. First, in a sample of Los Angeles secondary school teachers, Hammen and deMayo (1982) found that a one-item teacher stress measure correlated .63 with the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977; Weissman, Sholomskas, Pottenger, Prusoff, & Locke, 1977), a validated measure of current depressive symptoms. Had a more reliable, multiitem measure of stress been used, the coefficient would likely have been higher. Second, Belcastro and Hays (1984) compared burned-out and normal Alabama schoolteachers on 12 self-reported illnesses. Although the rate of illness in the burned-out teachers exceeded that of the normal teachers in only 3 of the 12 disorders, depression was the illness in which the rates for the two groups of teachers differed most sharply. Third, Greenglass and Burke (1988) studied 780 Canadian teachers. Greenglass (personal communication, 1990) found that the MBI was significantly related to depressive ($r = .53$), anxiety ($r = .44$), and somatic ($r = .44$) symptoms. Anxiety symptoms and somatic complaints are frequent accompaniments of depressive symptoms (Schonfeld, in press). Fourth, Meier (1984), using a mid-western college faculty sample, found that the MBI, the Meier Burnout Assessment, and a self-rating of burnout correlated with measures of depression about as strongly as the instruments' reliabilities permitted.

The fifth problem refers to an artifact in many burnout and stress measures. Correlations between burnout and stress measures and working conditions are likely to be inflated because many burnout and stress measures refer to difficulties at work. For example, Kyriacou and Sutcliffe's (1978, 1979) measure of stress and Maslach and Jackson's (1981, 1984) measure of burnout refer to job conditions in evaluating distress in teachers and (in the case of the MBI) other helping professionals. Other teacher stress measures (e.g., Fimian & Santoro, 1983) ask the respondents to indicate how bothered or annoyed they are by various school or classroom conditions.

Alternative Avenues to Measuring Stress in Teachers

Paradoxically, one way to measure stress in teachers is to do away with the concept of stress altogether (see Kasl, 1987). Instead of measuring the global concept of teacher stress, I argue that it is preferable to measure hypothesized aversive environmental conditions (the stressors) and depressive symptoms (the distress) independently. The CES-D is a satisfactory depressive symptom scale; however, other depressive symptom scales may do equally well. One advantage of the CES-D is the availability of normative data from unselected general-population samples (Schonfeld, 1990b). Studies of Los Angeles (Hammen & deMayo, 1982) and New York (Schonfeld, 1990a, 1990b) teachers revealed CES-D scores that were elevated in comparison with the normative landmarks that characterize general-population samples. These findings suggest that teaching carries some psychological risk. An epidemiologic survey of Western Australian teachers (Finlay-Jones, 1986) also revealed elevated psychological morbidity using a different general-population measure of psychological distress.

Another advantage of the CES-D is that it makes no reference to working conditions. As described earlier, the wording of many stress and burnout measures increases the likelihood that correlations with working conditions are biased upward. The use of depressive symptom scales like the CES-D, instead of stress/burnout measures, would help to reduce the likelihood of artifactual correlations between workplace stressors and distress.

What Do Depressive Symptom Scales Measure?

Depressive symptoms scales are thought to be reflective of either of two constructs: nonspecific psychological distress or clinical depression. Dohrenwend and his colleagues (Dohrenwend, Levav, & Shrout, 1986; Dohrenwend, Shrout, Egri, & Mendelsohn, 1980) showed that depressive symptoms tend to correlate with a variety of symptom scales about as highly as the scales' reliability coefficients permit. These symptom scales include measures of guilt, anxiety, low self-esteem, poor perceived health, and psychophysiologic symptoms. Dohrenwend and his colleagues (Dohrenwend et al., 1986; Dohrenwend et al., 1980) advanced the view that in the absence of clinical disorder such symptom scales probably measure a construct they called *nonspecific psychological distress* or *demoralization*, after the work of Frank (1973). Nonspecific distress might be thought of as the type of psychological state that

motivates many otherwise “normal” people—individuals who do not meet diagnostic criteria for psychiatric illness—to seek help from psychotherapists. Such people suffer from problems with living, although they would not qualify for a *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.; American Psychiatric Association, 1987) diagnosis.

On the other hand, individuals with high scores on depressive symptom scales like the CES-D are at increased risk for clinical depression (Weissman, Sholomskas, Pottenger, Prusoff, & Locke, 1977). In the psychiatric epidemiology literature, the CES-D has been conceived as a preliminary screening device to improve the efficiency of case finding in follow-up clinical interviews (Boyd, Weissman, Thompson, & Myers, 1982; B. P. Dohrenwend & B. S. Dohrenwend, 1982; Radloff, 1977; Weissman et al., 1977). On the other hand, the extreme distress reflected in high scores on scales like the CES-D in the absence of clinical depression also constitutes a serious mental health problem (Link & Dohrenwend, 1980).

Measures of the Work Environment

Kasl (1978) cogently argued that stress researchers should move away from “the excessive operational circularity in stress and distress measures” (p. 36). He rejected studies that linked measures of perceived job demands to measures of worker unhappiness with those demands. Three alternative measurement strategies for obtaining reasonable information on the work environment are available.

Objective information. First, one can obtain “objective” information on the quality of the work environment. Such a strategy is particularly applicable to studies comparing the health of workers in different occupations. For example, in research linking working conditions to cardiovascular disease, Schwartz, Pieper, and Karasek (1988) obtained independently from one sample (the Quality of Employment Survey [QES; Quinn & Staines, 1979]) ratings of job dimensions like decision latitude and physical exertion that would be used to characterize the jobs of healthy and unhealthy workers in two different samples (the U.S. Health Examination Survey [U.S. Department of Health, Education and Welfare, 1965] and the U.S. Health and Nutrition Examination Survey [U.S. Department of Health, Education and Welfare, 1979]). In a study of work history and schizophrenia risk, Link, Dohrenwend, and Skodal (1986) obtained independent ratings of job dimensions from the *Dictionary of Occupational Titles* (DOT; U.S. Department of Labor, 1965), a document summarizing ratings made by occupational analysts on a great variety of jobs. By using ratings of work dimensions that are external to the study’s participants, relations between job dimensions and measures of mental health are uncontaminated by self-report bias (Kasl, 1981).

It is more difficult to study within-occupation variation because objective external measures of job characteristics such as the DOT and the QES provide average values on work dimensions that characterize specific occupations. They do not provide data on within-occupation variability. In my own research on teachers, I attempted to obtain within-occupation, “objective” measures on the quality of the New York City public schools in which many of the study participants have worked. One confidential document that I secured with the help of the local teachers union

provided an independent, within-teachers measure of the stressfulness of their work environments. The document provided yearly information on the school-by-school rates at which teachers suffered assaults, robberies, and sexual offenses. One might argue that these statistics index the quality of the teachers' work environments by reflecting the average level of violence in the schools. The data were external to the study participants, who were newly appointed teachers, and were largely compiled before the participants obtained jobs.

This research strategy, however, is not without shortcomings. M. Gillespie, the director of City College's Principals Center and a former New York City public school principal, reported (personal communication, 1990) that principals vary in their willingness to permit such crimes to be aggregated into official records. I have known teachers to be assault victims and under pressure from their principals not report the assaults.

Paired study participants. An alternative within-teachers strategy would capitalize on the circumstance in which two study participants might obtain jobs in the same school. One teacher's report on his or her working conditions could be treated (in, for example, a regression equation) as the working conditions encountered by the other teacher, and vice versa. If sufficient numbers of pairs of teachers at the same level of experience obtain jobs in the same schools, this strategy may prove workable. A strength of such a strategy is that it provides ratings on the school environment that are independent of the teacher reporting distress. A weakness of the strategy is the loss of classroom-level data on the teacher's work environment. Furthermore, approximate matching by grade taught is required: A first-grade teacher's experience does not adequately reflect a fifth-grade teacher's experience even if they work in the same school.

Neutral self-reports. A third type of strategy involves measuring the teacher's immediate work environment with the help of self-report items that minimize the extent to which the teacher projects his or her feelings when responding. Kasl (1987) argued that neutral self-report items that capture the frequency with which workers encounter specifiable conditions are superior to items that ask workers to indicate how bothered or annoyed they are by those conditions. For example, in the longitudinal study of newly appointed teachers, participants were asked to estimate the frequency, within a specified period of time, with which they encountered students engaged in a fight. By contrast, the teachers were *not* asked to rate how annoyed or angered they were by student fighting. The latter type of item is probably confounded with symptoms. A weakness of neutral self-report items is that they probably cannot fully eliminate confounding with symptoms; evidence provided by Schonfeld and Ruan (1991), however, suggests that in scales developed from such items, confounding with symptoms can be kept to a reasonable minimum.

An Application

An ongoing longitudinal investigation of newly appointed teachers began before the teachers entered the work force (Schonfeld, 1991, 1992). The teachers were

recruited while they were in their final semesters of college. Preemployment (summer) and postemployment (fall and spring) depressive symptoms were measured with the CES-D. Neutral self-report items were used to obtain information on the teachers' work environments. The data were collected by questionnaires that were pilot tested in two veteran-teacher samples (e.g., Schonfeld, 1990b). The study of newly appointed teachers, which is described in detail in Schonfeld (1992), involved four graduating cohorts, the classes of 1987, 1988, 1989, and 1990, although the findings summarized here are limited to 255 women from the first three cohorts. Men and individuals who did not become teachers were excluded. About 90% of those eligible signed letters of informed consent, and 86% of the women recruited participated in the summer data collection.

Two reliable environmental stressor scales ($\alpha \geq .82$, fall and spring), the Episodic Stressor Scale and the Strain Scale, were constructed by aggregating neutral self-report items. The Episodic Stressor Scale was a measure of the rate at which the teacher encountered episodically occurring job-related events. An Episodic Stressor Scale score was calculated by computing the teacher's mean on the items assessing the frequency with which she encountered such events (e.g., threat of personal injury, encounter with an insolent student). Each scale item was scored as follows: (0) not at all, (1) once per month, (2) once per week, (3) 2–4 times per week, or (4) daily.

The Strain Scale, which was named for the pioneering work of Pearlin and Schooler (1978), pertained to ongoing types of stressors. A Strain Scale score was calculated by computing the teacher's mean score on items assessing ongoing conditions (e.g., an overcrowded classroom, unmotivated students attending class). Each scale item was scored as follows: (0) not at all, (1) to a minimal extent, (2) to a small extent, (3) to a moderate extent, or (4) to a great extent.

Both the Episodic Stressor Scale and the Strain Scale included (reverse scored) items reflecting positive aspects of the work environment (e.g., "a parent praised you") to counter any tendencies toward response set.

Three sets of analyses were conducted to examine the link between the school environment and depressive symptoms. In the first set of analyses, the sample was arbitrarily divided into three, about equal-sized groups consisting of the lowest, middle, and highest scorers on the fall job environment scales. The purpose of this set of analyses was to evaluate mean group differences on the summer ($n = 247$), fall ($n = 244$), and spring ($n = 210$) CES-D. The findings for the Episodic Stressor Scale, depicted in Figure 1, revealed no summer differences among the sample, but significant differences were found for the fall and spring. Similar findings were obtained with the Strain Scale. It is of interest that the preemployment symptom picture in women who obtained jobs in the best schools improved.

In the second set of analyses, the fall CES-D was regressed on each of the two job environment scales ($n = 238$ with listwise deletion for missing data). In both regression equations, preemployment CES-D, social class of origin, life events, marital status, race, preemployment social support, and age were controlled. The correlation increase associated with the Episodic Stressor Scale was .20. In the full regression equation, only preemployment CES-D and the Episodic Stressor Scale were significantly related to the outcome. The unstandardized regression weight ($B = 8.33$) for the Episodic Stressor Scale indicated that, on average, a unit increase

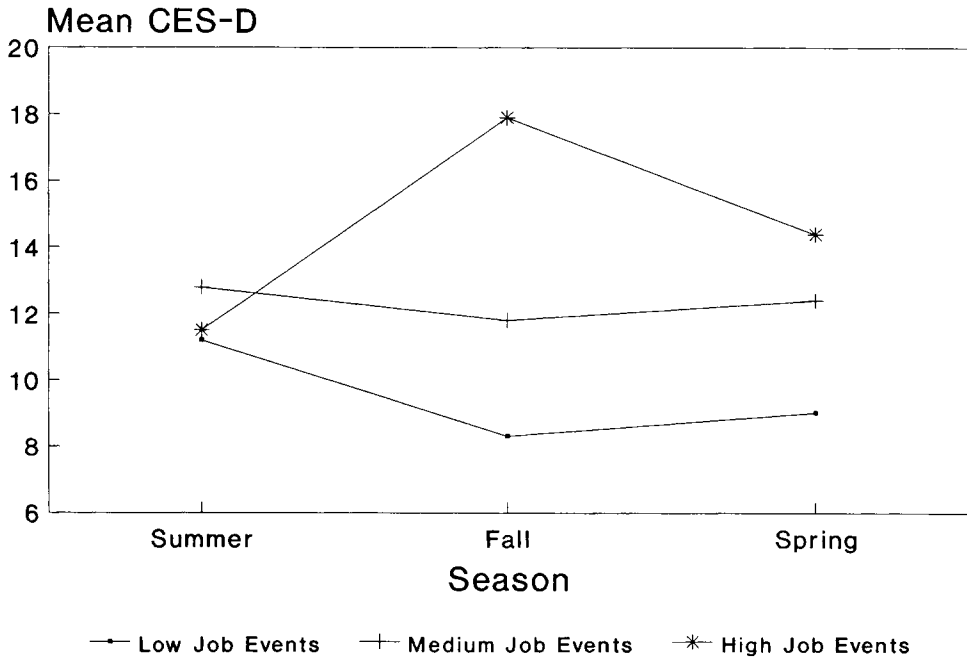


Figure 1. Depressive symptoms in female teachers in the fall low-, medium-, and high-event groups. (Depressive symptoms were measured by the CES-D. The high-event group represents the teachers in the most adverse school environments, and the low-event group represents the teachers in the “best” school environments. Significant group effects were found in the fall [$p < .001$] and spring [$p < .01$] but not the summer. From Schonfeld, 1992. Copyright 1992 by Pergamon Press. Reprinted by permission.)

on the Episodic Stressor Scale was associated with an 8-point increase on the CES-D. A unit increase may be thought of as the difference between a scale score of 1, representing a once-a-month occurrence of a variety of adverse school-related events, and a scale score of 2, representing a once-a-week occurrence of such events. To contextualize this finding, note that the median mean score on the CES-D in general-population samples is about 8 (Schonfeld, 1990b), and 16 is considered to be a marker of clinical significance (Boyd et al., 1982). The results were similar when the Strain Scale replaced the Episodic Stressor Scale in the regression equation.

Despite the inclusion of control variables measured at an earlier point in time (e.g., summer CES-D, social support), the regression procedures constituted concurrent analyses: Fall symptoms were measured at the same time as school conditions. The relatively large effect sizes for the school environment variables did not rule out the hypothesis that preexisting depressive symptoms somehow “caused” the occurrence of the putative environmental stressors (the event-proneness model described earlier). An event-proneness explanation of the findings, however, did not hold because the zero-order correlations between the summer CES-D and the fall and spring school environment measures did not differ significantly from zero.

The third and final set of analyses ($n = 196$) used LISREL software (Jöreskog & Sorbom, 1989) to develop structural equation models (Hayduk, 1987) linking stressors and symptoms during the course of the teachers’ first year on the job.

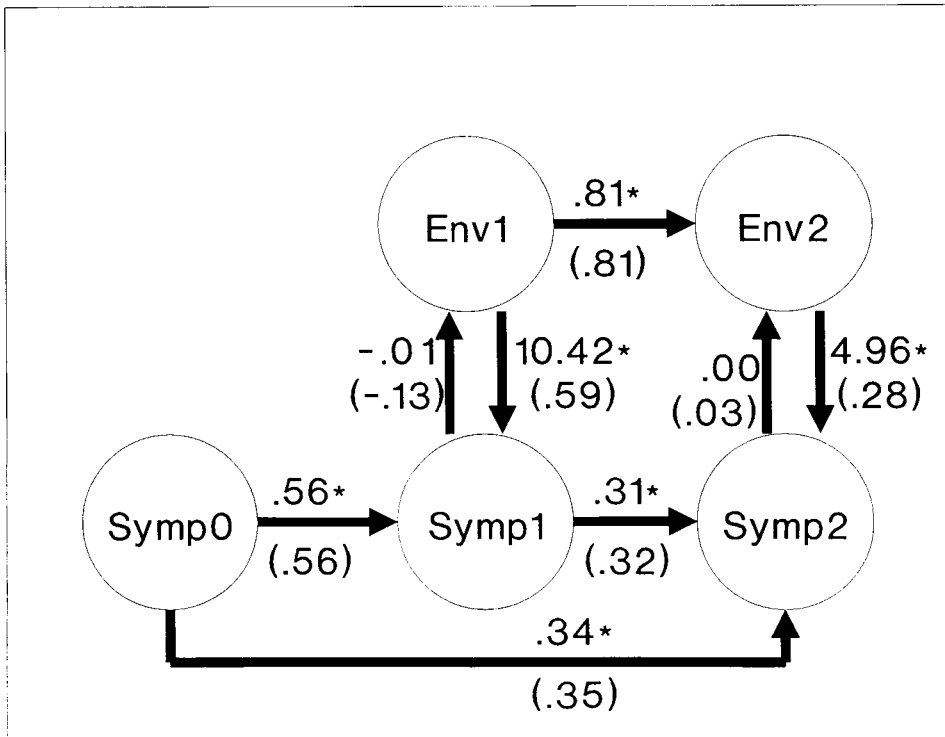


Figure 2. LISREL model of reciprocal effects in newly appointed female teachers. (Env1 and Env2 represent adversity in the school environment in the fall and spring, respectively. Symp1, Symp2, and Symp3 represent depressive symptoms in the summer, fall, and spring, respectively. Unstandardized coefficients are presented above each path, and standardized coefficients are presented below each path in parentheses. Asterisks indicate a significant causal path [$p < .001$]. From Schonfeld, 1992. Copyright 1992 by Pergamon Press. Reprinted by permission.)

Model building was simplified by excluding all control variables except the summer CES-D. This decision was justified in view of the regression analyses, which failed to demonstrate significant effects for control variables other than the summer CES-D. A latent fall (Time 1) school environment variable (Env1 in Figures 2 and 3) was constructed using both the fall Episodic Stressor and Strain Scales as indicators. Similarly, a latent spring (Time 2) school environment variable (Env2 in Figures 2 and 3) was constructed using both the spring Episodic Stressor and Strain Scales as indicators. Env1 and Env2 were forced to assume the same units as the Episodic Stressor Scale. Because both the Time 1 and Time 2 environment factors each had two indicators, LISREL estimated the error terms for each school-stressor indicator.

Summer (preemployment), fall, and spring CES-Ds were used to construct the Time 0, 1, and 2 symptoms measures, respectively. Each Time 0, 1, and 2 symptoms factor was forced to have the same scale as its CES-D indicator. An error term derived from the reliability coefficient ($\alpha \geq .89$) for each CES-D indicator was included in the model. Preemployment (Time 0) symptoms constituted an "instrumental" variable required for estimating reciprocal effects (Kenny, 1979). The re-

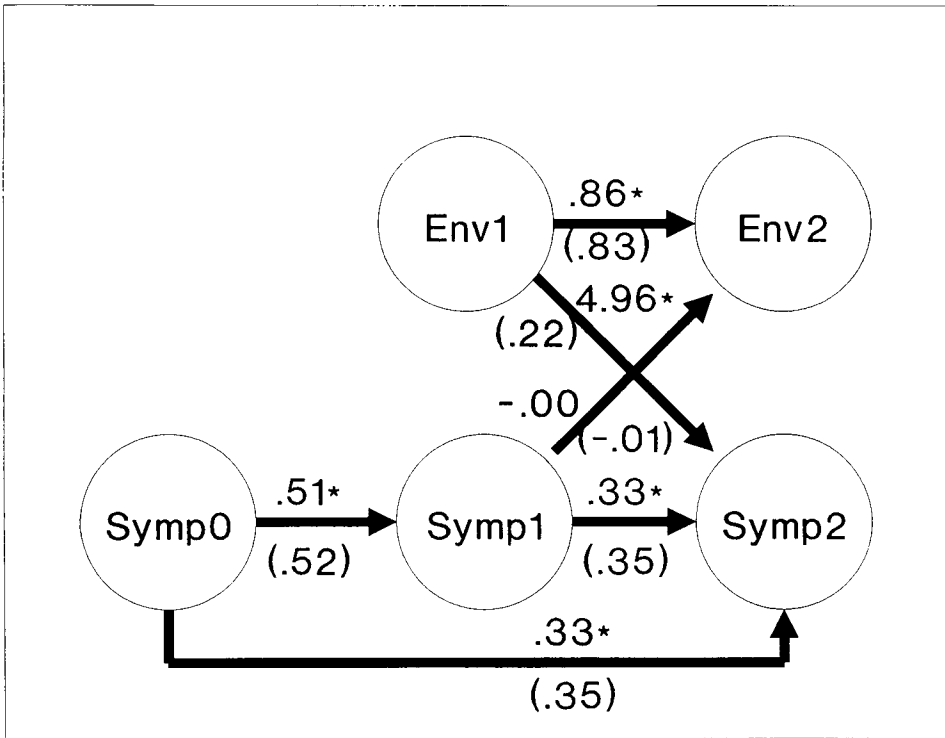


Figure 3. LISREL model of lagged effects in newly appointed female teachers. Env1 and Env2 represent adversity in the school environment in the fall and spring, respectively. (Symp1, Symp2, and Symp3 represent depressive symptoms in the summer, fall, and spring, respectively. Unstandardized coefficients are presented above each path and standardized coefficients are presented below each path in parentheses. Asterisks indicate a significant causal path [$p < .001$]. From Schonfeld, 1992. Copyright 1992 by Pergamon Press. Reprinted by permission.)

ciprocal effects model shown in Figure 2 was consistent with the data. At Time 1 (fall) and again at Time 2 (spring), the model shows reciprocal effects between depressive symptoms and the work environment. The effect from environment to symptoms was at each time considerably greater than the small (“halo”) effects from symptoms to environment (*ns*). A rival lagged-effects model was also tested but rejected (see Figure 3).

The three sets of analyses jointly suggest that the effects that working conditions exert on depressive symptoms in newly appointed female teachers is relatively immediate. Adverse effects (and perhaps beneficial effects associated with obtaining a job in low-events schools) begin to appear during the teachers’ first year on the job. It would have been too late to have begun a study of the effects of working conditions later in the teachers’ careers. A strategic feature of the study, the measurement of depressive symptoms prior to the participants’ entry into the work force, afforded an opportunity to evaluate and reject an event-proneness explanation of the findings. An advantage of research designs that follow newly employed, in comparison with veteran, workers is the opportunity that such designs furnish for

obtaining preemployment measures of health (see Kasl, 1983; Schonfeld & Ruan, 1991). Preemployment health or symptom measures may constitute instrumental variables that could help to disentangle the effects that develop after individuals enter the work force.

Self-Report Items

An important feature of the study was the creation of job environment variables based on neutral self-report items. This chapter advances the view that carefully worded self-report items can sometimes be as helpful as many "objective" indicators in constructing measures of the work environment. This is not to say there will be no risk of confounding with prior psychological symptoms. The extent to which there will be confounding is an empirical question that should be studied. The issue of which items are likely to be confounded with prior symptoms ought to be resolved in pilot research prior to the substantive study.

I am not arguing for doing away with so-called objective measures. On the contrary, objective indicators of airport noise near a school may be superior to subjective measures (Cohen, Evans, Stokol, & Krantz, 1986). There are, however, occasions when objective measures of the work environment are not objective enough, as in the case of official school-by-school records of assaults against teachers. Alternatively, neutral self-report items may inquire into teachers' exposures to putative stressors (e.g., pupil fighting, vandalism). Coupled with satisfactory longitudinal data, an investigator may examine the independence of postemployment stressors from preemployment health and the effects of exposures to such stressors on future health.

The Role of the CES-D in Identifying Psychiatric Cases

Depressive symptom scales are one of a number of possible outcome variables. Earlier mentioned advantages of well-standardized measures like the CES-D include normative landmarks from general-population samples and the absence of reference to the work environment. Scales like the CES-D have another advantage absent in traditional stress and burnout measures. Because high scores on scales like the CES-D suggest elevated risk for affective illness and low scores decreased risk, these measures make it possible to add, economically, clinical case-finding procedures to research on occupational stress.

An example of the utility of such a procedure comes from a current pilot study of the 1990 cohort. The purpose of the study is to examine the relation between incident psychiatric disorder and the working conditions of newly appointed teachers. The psychiatric study was put in the field as a "graft" onto the longitudinal study described in this chapter. Because of budget constraints, experienced female diagnosticians conducted telephone interviews with the subjects selected ($n = 69$ female teachers and nonteachers, 61 interviews completed as of the writing of this chapter). Ideally, face-to-face interviews are to be preferred. The use of the telephone interviews, however, was not considered a drawback because participants were

accustomed to relatively frequent mail and telephone contact in connection to their participation in the main study.

To qualify for the psychiatric study, a female teacher must have obtained a score on the preemployment CES-D that is below the clinical cutoff (<16). Scores greater than or equal to the cutoff are associated with increased risk of clinical depression, and scores below the cutoff are associated with decreased risk (Boyd et al., 1982; Radloff, 1977; Weissman et al., 1977). Any teacher whose preemployment CES-D was greater than or equal to the cutoff did not qualify for the clinical study because the study's aim was to assess the incidence of disorder after the women obtain jobs of varying quality. This preemployment "gate" reduces the likelihood of identifying individuals with preexisting disorders. Every subject whose (a) preemployment (summer) CES-D was below the cutoff and (b) fall CES-D was greater than or equal to the clinical cutoff was scheduled for a clinical interview after the project's receipt of the completed fall CES-D. These subjects might be aptly called the *low-high group* based on their summer-fall CES-D profile.

Members of a second group, a low-low group consisting of women whose scores on the CES-D were below the cutoff in both the summer and fall, were also scheduled for interviews. Because there were many more low-lows than low-highs, the low-lows were randomly selected and matched to the low-highs for date of the project's receipt of the completed fall CES-D. There are approximately equal numbers of

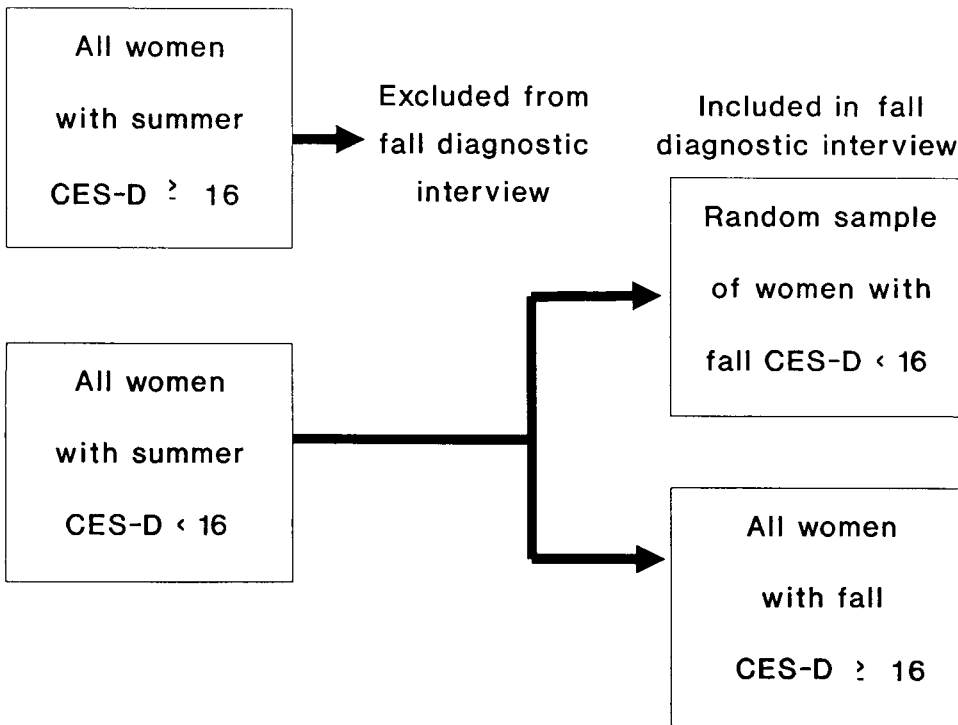


Figure 4. Description of the two-stage research design to ascertain the incidence of affective illness in teachers: fall procedures. (CES-D = Center for Epidemiologic Studies Depression Scale.)

low-highs and low-lows. The clinical interviewers were blind to all information available in the completed questionnaires, including the subjects' scores on both CES-Ds. The study design is depicted in Figure 4.

It was more preferable to sample from the low-low group than to interview every member of that group because depressive illness, the disorder of primary interest, is rare among individuals with low scores on the CES-D. By contrast, it was desirable to interview all low-high subjects because that group was at considerably higher risk than was the low-low group. These twin procedures made for an economy of scale while maximizing the chances of identifying true cases. An extended set of procedures, described in Figure 5, was used for data collected in the spring.

It is important to bear in mind that the CES-D reflects a continuous dimension of psychopathology. Although the CES-D provides good data on this dimension, the scale does not provide sufficient information to diagnose psychiatric disorder (B. P. Dohrenwend et al., 1986). To diagnose psychiatric illness, the second clinical interview stage of data collection is required.

According to B. P. Dohrenwend and B. S. Dohrenwend (1982), in their important article on psychiatric epidemiology, two-stage procedures that combine psychometrically valid screening instruments in Stage 1 with standardized psychiatric interviews in Stage 2 constitute an optimum method for ascertaining "true" incidence and prevalence rates. An advantage that two-stage procedures have over methods

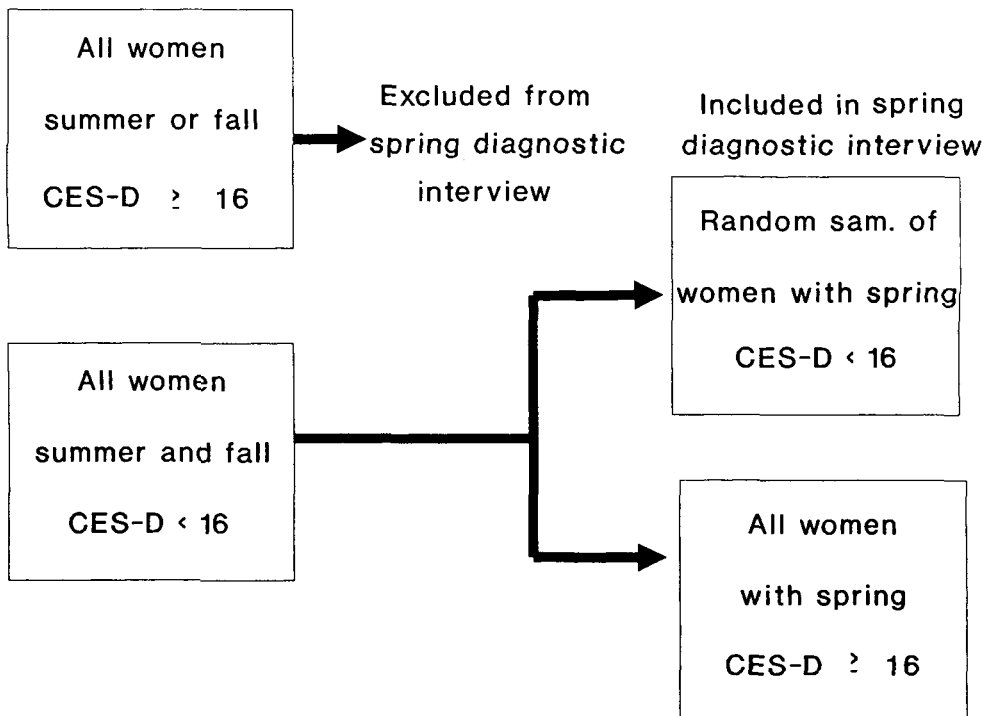


Figure 5. Description of the two-stage research design to ascertain the incidence of affective illness in teachers: spring procedures. (CES-D = Center for Epidemiologic Studies Depression Scale.)

in which entire samples are administered psychiatric interviews in one step without prior screening is that the Stage 1 screening instrument, here the CES-D, serves as a check on the diagnostic information ascertained in Stage 2. With the important control feature of making the Stage 2 diagnosticians blind to the Stage 1 CES-D scores, individuals identified as having a disorder in Stage 2, compared with individuals without a Stage 2 disorder, should have significantly higher scores on the Stage 1 screening instrument.

The diagnostic interviewers used the *Structured Clinical Interview for DSM-III-R (SCID)*; Spitzer, Williams, Gibbon, & First, 1989). The *SCID* was reorganized slightly for the purposes of this study. To put the interviewee at ease, a number of questions about fairly common "medical" conditions (taken from the section on somatization disorder) were asked before the interviewer turned to the sections on major depression, dysthymia, and a number of anxiety diagnoses. Sections pertaining to psychiatric disorders not relevant to this study (e.g., bipolar illness, schizophrenia) were not included.

Another check on the quality of the data is built into the *SCID*. The *SCID* ascertains the time of onset of severe symptoms. This dating procedure served as a further check on the hypothesis that the diagnoses ascertained reflected incident cases as would be expected when participants with high preemployment CES-D scores were screened out.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.
- Belcastro, P. A., & Hays, L. C. (1984). Ergo . . . ergophobia . . . ergo . . . burnout? *Professional Psychology: Research and Practice*, *15*, 260–270.
- Boyd, J. H., Weissman, M. M., Thompson, W. D., & Myers, J. K. (1982). Screening for depression in a community sample: Understanding the discrepancies between depression syndrome and diagnostic scales. *Archives of General Psychiatry*, *39*, 1195–1200.
- Cohen, S., Evans, G. W., Stokol, D., & Krantz, D. S. (1986). *Behavior, health, and environmental stress*. New York: Plenum Press.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*, 385–396.
- Coyne, J. C., Burchill, S. A. L., & Stiles, W. B. (1990). An interactional perspective on depression. In C. R. Snyder & D. O. Forsyth (Eds.), *Handbook of social and clinical psychology: The health perspective* (pp. 327–349). New York: Pergamon Press.
- Coyne, J. C., Kahn, J., & Gotlib, I. H. (1987). Depression. In T. Jacob (Ed.), *Family interaction and psychopathology* (pp. 509–533). New York: Plenum Press.
- Cunningham, W. G. (1983). Teacher burnout—Solutions for the 1980s: A review of the literature. *Urban Review*, *15*, 37–51.
- DeFrank, R. S., & Stroup, C. A. (1989). Teacher stress and health: Examination of a model. *Journal of Psychosomatic Research*, *33*, 99–109.
- Dohrenwend, B. P., & Dohrenwend, B. S. (1982). Perspectives on the past and future of psychiatric epidemiology. *American Journal of Public Health*, *72*, 1271–1279.
- Dohrenwend, B. P., Levav, I., & Shrout, P. E. (1986). Screening scales from the Psychiatric Epidemiology Research Interview (PERI). In M. M. Weissman, J. K. Myers, & C. E. Ross (Eds.), *Community surveys of psychiatric disorders* (pp. 349–375). New Brunswick, NJ: Rutgers University Press.
- Dohrenwend, B. P., Shrout, P. E., Egri, G., & Mendelsohn, F. S. (1980). Nonspecific psychological distress and other dimensions of psychopathology: Measures for use in the general population. *Archives of General Psychiatry*, *37*, 1229–1236.

- Dohrenwend, B. S., & Dohrenwend, B. P. (1981). Life stress and illness: Formulation of the issues. In B. S. Dohrenwend & B. P. Dohrenwend (Eds.), *Stressful life events and their contexts* (pp. 1–27). New Brunswick, NJ: Rutgers University Press.
- Dunham, J. (1984). *Stress in teaching*. Beckenham, Kent, England: Croom Helm.
- Dworkin, A. G. (1988). *Teacher burnout in the public schools*. Albany: State University of New York Press.
- Dworkin, A. G., Haney, C. A., & Telschow, R. L. (1988). Fear, victimization, and stress among urban public school teachers. *Journal of Organizational Behavior*, 9, 159–171.
- Farber, B. (1984). Stress and burnout in suburban teachers. *Journal of Educational Research*, 77, 325–331.
- Fimian, M. J. (1983). A comparison of occupational stress correlates as reported by teachers of mentally retarded and non-mentally retarded handicapped children. *Education and Training of the Mentally Retarded*, 18, 62–68.
- Fimian, M. J., & Santoro, T. M. (1983). Sources and manifestations of occupational stress as reported by full-time special education teachers. *Exceptional Children*, 49, 540–543.
- Finlay-Jones, R. (1986). Factors in the teaching environment associated with severe psychological distress among school teachers. *Australian and New Zealand Journal of Psychiatry*, 20, 304–313.
- Frank, J. D. (1973). *Persuasion and healing*. Baltimore: Johns Hopkins Press.
- Galloway, D., Panckhurst, F., Boswell, K., Boswell, C., & Green, K. (1984). Mental health, absences from work, stress, and satisfaction in a sample of New Zealand primary school teachers. *Australian and New Zealand Journal of Psychiatry*, 18, 359–363.
- Gold, Y. (1984). The factorial validity of the Maslach Burnout Inventory in a sample of California elementary and junior high school classroom teachers. *Educational and Psychological Measurement*, 44, 1009–1016.
- Gold, Y. (1985). The relationship of six personal and life history variables to standing on three dimensions of the Maslach Burnout Inventory in a sample of elementary and junior high school teachers. *Educational and Psychological Measurement*, 45, 377–387.
- Greenglass, E. R., & Burke, R. J. (1988). Work and family precursors of burnout in teachers: Sex differences. *Sex Roles*, 18, 215–229.
- Hammen, C., & deMayo, R. (1982). Cognitive correlates of teacher stress and depressive symptoms: Implications for attributional models of depression. *Journal of Abnormal Psychology*, 91, 96–101.
- Hayduk, L. A. (1987). *Structural equation modeling with LISREL*. Baltimore: Johns Hopkins Press.
- Iwanicki, E. F., & Schwab, R. L. (1981). A cross validation study of the Maslach Burnout Inventory. *Educational and Psychological Measurement*, 41, 1167–1174.
- Johnson, A. B., Gold, V., & Knepper, D. (1984). Frequency and intensity of professional burnout among teachers of the mildly handicapped. *College Student Journal*, 18, 261–266.
- Jöreskog, K. G., & Sorbom, D. (1989). *LISREL 7: User's reference guide*. Mooresville, IN: Scientific Software.
- Kahill, S. (1988). Symptoms of professional burnout: A review of the empirical evidence. *Canadian Psychology*, 29, 284–297.
- Kasl, S. V. (1978). Epidemiological contributions to the study of work stress. In C. L. Cooper & R. Payne (Eds.), *Stress at work* (pp. 3–48). New York: Wiley.
- Kasl, S. V. (1981). The challenge of studying the disease effects of stressful work conditions. *American Journal of Public Health*, 71, 682–684.
- Kasl, S. V. (1983). Pursuing the link between stress life experiences and disease: A time for reappraisal. In C. L. Cooper (Ed.), *Stress research* (pp. 79–102). New York: Wiley.
- Kasl, S. V. (1987). Methodologies in stress and health: Past difficulties, present dilemmas, future directions. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 307–318). New York: Wiley.
- Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.
- Kyriacou, C., & Sutcliffe, J. (1978). Teacher stress: Prevalence, sources, and symptoms. *British Journal of Educational Psychology*, 48, 159–167.
- Kyriacou, C., & Sutcliffe, J. (1979). Teacher stress and satisfaction. *Educational Research*, 21, 89–96.
- Link, B., & Dohrenwend, B. P. (1980). Formulation of hypotheses about the true prevalence of demoralization in the United States. In B. P. Dohrenwend, M. S. Gould, B. Link, R. Neugebauer, & R. Wunsch-Hitzig (Eds.), *Mental illness in the United States* (pp. 114–132). New York: Praeger.

- Link, B., Dohrenwend, B. P., & Skodal, A. E. (1986). Socio-economic status and schizophrenia: Noisome occupational characteristics as a risk factor. *American Sociological Review*, *51*, 242–258.
- Malanowski, J. R., & Wood, P. H. (1984). Burnout and self-actualization in public school teachers. *Journal of Psychology*, *117*, 23–26.
- Maslach, C., & Jackson, S. (1981). *Maslach Burnout Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., & Jackson, S. (1984). Burnout in organizational settings. *Applied Social Psychology Annual*, *5*, 133–153.
- McIntyre, T. C. (1984). The relationship between locus of control and teacher burnout. *British Journal of Educational Psychology*, *54*, 235–238.
- Meier, S. T. (1984). The construct validity of burnout. *Journal of Occupational Psychology*, *53*, 211–219.
- Needle, R. H., Griffen, T., & Svendsen, R. (1981). Occupational stress: Coping and health problems of teachers. *Journal of School Health*, *51*, 175–181.
- Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, *19*, 2–21.
- Pettegrew, L. S., & Wolf, G. E. (1982). Validating measures of teacher stress. *American Educational Research Journal*, *19*, 373–396.
- Pierson-Hubeny, D., & Archambault, F. X. (1985). Role stress and perceived intensity of burnout among reading specialists. *Reading World*, *24*, 41–52.
- Quinn, R. P., & Staines, G. L. (1979). *The 1977 Quality of Employment Survey: Descriptive statistics with comparison data from the 1960–70 and the 1972–73 surveys*. Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.
- Schonfeld, I. S. (1990a). Coping with job-related stress: The case of teachers. *Journal of Occupational Psychology*, *63*, 141–149.
- Schonfeld, I. S. (1990b). Psychological distress in a sample of teachers. *Journal of Psychology*, *123*, 321–338.
- Schonfeld, I. S. (1991). Dimensions of functional social support and psychological symptoms. *Psychological Medicine*, *21*, 1051–1060.
- Schonfeld, I. S. (1992). A longitudinal study of occupational stressors and depressive symptoms in first-year teachers. *Teaching and Teacher Education*, *8*, 151–158.
- Schonfeld, I. S. (in press). Burnout in teachers: Is it burnout or is it depression? *Human Stress: Current and Selected Research*.
- Schonfeld, I. S., & Ruan, D. (1991). Occupational stress and preemployment measures: The case of teachers. *Journal of Social Behavior and Personality*, *6*, 95–114.
- Schwartz, J. E., Pieper, C. F., & Karasek, R. A., Jr. (1988). A procedure for linking job characteristics to health surveys. *American Journal of Public Health*, *78*, 904–909.
- Seiler, R. E., & Pearson, D. A. (1984). Stress among accounting educators in the United States. *Research in Higher Education*, *21*, 301–316.
- Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1989). *Structured clinical interview for DSM-III-R: Non-patient edition*. New York: New York State Psychiatric Institute.
- U.S. Department of Health, Education and Welfare. (1965). *Coronary heart disease in adults: U.S., 1960–62* (DHEW Public Health Series No. 1000, Series 11). Washington, DC: U.S. Government Printing Office.
- U.S. Department of Health, Education and Welfare. (1979). *Plan of operation of the Health and Nutritional Examination Survey, U.S., 1971–73* (DHEW Publication No. 79-1310, Series 1, No. 10b). Washington, DC: U.S. Government Printing Office.
- U.S. Department of Labor. (1965). *Dictionary of occupational titles* (3rd ed.). Washington, DC: U.S. Government Printing Office.
- Weissman, M., Sholomskas, M., Pottenger, M., Prusoff, B., & Locke, B. (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology*, *106*, 203–214.
- Worchel, S., & Teddlie, C. (1976). The experience of crowding: A two-factor theory. *Journal of Personality and Social Psychology*, *34*, 30–40.