A LONGITUDINAL STUDY OF OCCUPATIONAL STRESSORS AND DEPRESSIVE SYMPTOMS IN FIRST-YEAR FEMALE TEACHERS

IRVIN SAM SCHONFELD

City College of New York and Columbia University, U.S.A.

Abstract—Three cross-sectional studies (Finlay-Jones, 1986; Hammond & deMayo, 1982; Schonfeld, 1990) at best suggest that veteran teachers are at some risk for experiencing above-average levels of psychological distress. The paper advances three reasons for the view that longitudinal studies with newly appointed teachers provide a better means with which to examine the link between working conditions and psychological distress: (1) the comparative weakness of cross-sectional studies in distinguishing among causal hypotheses; (2) the relative absence in veteran teacher samples of individuals who quit in response to adverse school conditions; (3) the need to control preemployment symptoms. A longitudinal study involving 255 newly appointed female teachers showed that job conditions were related to postemployment depressive symptoms independently of preemployment symptoms and other risk factors. Regression and LISREL analyses revealed that the effects of working conditions on symptoms are relatively immediate. Moreover, other analyses suggested that there may be positive mental health effects, in relation to preemployment levels, associated with teaching in “benign” school environments. Suggestions for future progress in teacher-stress research include the use of neutral self-reports (Kasl, 1987; Schonfeld, 1990) to assess school conditions and a greater reliance on standardized instruments to measure independent and dependent variables.

Inquiry into the impact of work environments on psychological functioning can benefit from a look at epidemiologic methods. Epidemiology is the study of the health of defined populations. The key term is “defined.” By carefully defining the population of interest, an investigator can obtain insights into job-related risk factors for ill health. The population of interest for the study reported in this paper is newly appointed schoolteachers. The advantages of studying newly appointed, in comparison to veteran, teachers will be described later. Studies of psychological distress in veteran teachers, however, served as a prelude to research on distress in new teachers.

One cross-sectional study (Schonfeld, 1990) of New York City teachers with an average of 12 years experience revealed relatively high levels of depressive symptoms as measured by the Center for Epidemiologic Studies — Depression Scale (CES-D; Radloff, 1977; Weissman, Sholomskas, Pottinger, Prusoff, & Locke, 1977). The CES-D was developed at the National Institute of Mental Health to be used in unselected general-population samples. Elevated scores on the CES-D reflect increased risk for affective illness and high scores without affective illness generally reflect high levels of nonspecific psychological distress (Dohrenwend, Shrout, Egri, & Mendelsohn, 1980; Schonfeld, in press).

The average CES-D of the veteran teachers in
the cross-sectional sample was about 13. To put that mean score in perspective consider that the median mean score in 12 different community surveys is eight and that a score of 16 is considered to be a clinical cutoff, a value at or above which there is increased risk of major depressive disorder (Schonfeld, 1990). Thirty-two percent of the sample obtained scores above 16, about twice the proportion found in epidemiologic surveys. Eleven percent of the sample obtained scores at or above 24, the mean found in a sample of psychiatric patients (Radloff, 1977).

The findings from the New York veteran teacher sample were in keeping with other cross-sectional studies. A study (Hammen & deMayo, 1982) of veteran Los Angeles secondary school teachers revealed CES-D scores that were elevated in comparison to the normative landmarks that characterize general population samples. An epidemiologic survey of Western Australian teachers (Finlay-Jones, 1986) also revealed elevated levels of psychological morbidity using a different, general-population measure of psychological distress.

Because he used a parallel instrument, Schonfeld (1990) found that the level of job satisfaction that characterized the New York teachers compared unfavorably with levels found in the Quality of Employment Surveys (QES; Quinn & Staines, 1979). The QES consisted of three randomly selected national samples of American workers.

Evidence adduced by Finlay-Jones (1986), Hammen and deMayo (1982), and Schonfeld (1990) has been, at best, suggestive of high levels of distress in veteran teachers although cross-sectional findings have not unequivocally supported that view (Eaton, Anthony, Mandel, & Garrison, 1990). In response to the limitations inherent in cross-sectional research with veteran teachers, a longitudinal study of new teachers was put into the field. Briefly, there are at least three reasons for mounting such a longitudinal study. First, cross-sectional studies constitute the weakest type of research aimed at testing hypotheses concerning cause–effect relations. With longitudinal data there is a greater opportunity to test causal models.

Second, studies involving veteran teacher samples are likely to include individuals who have made relatively successful adaptations to their jobs because such samples do not include individuals who quit. Compared to their better functioning colleagues, the casualties of occupational stresses are more likely to have quit their jobs before the investigator arrives on the scene to begin a longitudinal study of a veteran worker sample (Kasl, 1983). The individuals not found in veteran teacher samples because of attrition (Esrig, 1987; Harris, Kagay, & Leichenko, 1986; Korshavn, 1991) would likely include a disproportionate number of the casualties of job stressors. A shortcoming associated with research on newly appointed teachers is limited generalizability. Most teachers are veteran teachers. By studying longitudinally newly appointed teachers, however, an investigator can identify the casualties of adverse job conditions as those casualties develop.

Third, with veteran teacher samples it is difficult to control for key preemployment factors. The bane of epidemiologic research is selection. More or less distressed individuals may select themselves, or be selected by administrative gatekeepers, into different occupational roles. Occasionally in research on physical health there is some record of preemployment morbidity. In research on mental health outcomes there is usually no information on preemployment psychological symptoms (Schonfeld & Ruan, 1991). Newly appointed teachers were studied in order to measure and control factors like preemployment depressive symptoms. The preemployment symptom measure would prove helpful in creating “instrumental” variables (Kenny, 1979) required for developing causal models of effects.

Parenthetically, the published literature on teacher stress almost exclusively involves samples of veteran teachers. Korshavn’s (1991) work on occupational longevity is an important exception because it is one of the only studies to examine large numbers of newly appointed teachers.

In the present study, the investigator measures episodically occurring school-related events and ongoing conditions. Teacher-stress researchers have largely neglected to differentiate between evenful experience and life
strains (Pearlin, Lieberman, Menaghan, & Mullan, 1981). Stressful life events are episodic, undesirable, and unscheduled (Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1982; Pearlin et al., 1981). Strains refer to enduring, threat arousing problems (Pearlin & Schooler, 1978). For example, the threat of involuntary transfer would constitute a strain and an episode of vandalism, an event. Both undesirable events and strains are related to distress in veteran teachers (Schonfeld, 1990).

The focal interest of the longitudinal study presented here is to assess the effects of occupational conditions exert on depressive symptoms in newly appointed female teachers. The study design capitalizes on variability among the schools in which the teachers worked. Within-occupations studies play an important role in stress research. For example, a study involving air traffic controllers at airports differing in traffic density has been helpful in linking job stressors to increased risk of hypertension (Cobb & Rose, 1973).

Method

Sample. The sample frame and the sampling procedures are described in detail elsewhere (Schonfeld, 1991; Schonfeld & Ruan, 1991). Briefly, as part of a larger study, subjects were recruited during spring semesters in 1987, 1988, and 1989 while they were upper seniors attending their last education seminars at leading teacher-training institutions in New York City. The aim of the recruitment procedures was to obtain a highly representative sample of newly appointed teachers. For this paper, the sample consisted of 255 women who were teachers in the fall following recruitment. Their average age was 28 and 25% were non-white. Males who became teachers were too few in number (n = 26) to be included in the analyses. More than 90% of the individuals eligible to be selected signed letters of informed consent and 86% of the women who signed such letters participated in the summer round of data collection. The women were contacted in the summer prior to their entering the work force and twice more, once in the fall and once in the following spring. Women who did not enter the teaching profession are not described.

Measures. The summer, fall and spring questionnaires supplied information on depressive symptoms and nonoccupational stressors (fateful, undesirable life events such as the death of a loved one). Depressive symptoms were measured with the Center for Epidemiologic Studies — Depression Scale (CES-D; Radloff, 1977; Weissman et al., 1977). In the present sample, the α coefficients for the summer, fall, and spring CES-Ds exceeded .89. Social support was measured in the summer with eight positively and negatively worded items from Cohen’s (Cohen, Karmack, & Mermels Siein, 1983; Cohen & Wills, 1985) Likert-type revision of the Interpersonal Support Evaluation List (α = .79). The demographic section of the questionnaire provided information on age, marital status, social class of origin (Hollingshead, 1974), and race.

Two measures of the adversity of the school environment were developed: (1) the Episodic Stressor Scale; and (2) the Strain Scale. The Episodic Stressor and Strain Scales reflect the distinction made in the stress literature between eventful experience and chronically occurring conditions (Pearlin & Schooler, 1978). Both scales employed neutrally worded self-report items assessing the frequency with which the teachers encountered different types of stressors. In contrast to traditional stress and burnout items that assess the extent to which the teachers are annoyed, bothered, or otherwise disturbed by stressors, neutrally worded items are less likely to be confounded with symptoms (Kasl, 1987; Schonfeld, 1990).

The Episodic Stressor Scale was created by computing the teacher’s mean score on items assessing the frequency with which he or she encountered episodically occurring stressors (e.g., threat of personal injury, confrontation initiated by an insolent student, episode of vandalism). Each item was scored: 0: not at all; 1: once per month; 2: one per week; 3: 2–4 times per week; and 4: daily. The Strain Scale was created by computing the teacher’s mean score on items assessing ongoing stressors (e.g., overcrowded classroom, unmotivated students attending class, tendency of admini-
strators not to enforce rules against disruptive pupils). Each item was scored: 0: not at all; 1: to a minimal extent; 2: to a small extent; 3: to a moderate extent; and 4: to a great extent. Both the Episodic Stressor and Strain Scales included positively worded items (e.g., “a parent praised you”) in order to break any tendencies toward response set. These items as well as positively worded items found in the CES-D were reverse scored. α Coefficients for the fall and spring school environment scales were .83 or greater.

Results

Mean scores on the CES-D. The teachers were divided into low-, medium-, and high-events groups using the 33rd and 67th percentile ranks on the fall Episodic Stressor Scale. One-way analyses of variance described in Table 1 indicate that the groups did not reliably differ in their scores on the summer, preemployment CES-D but did differ significantly on the fall and spring CES-Ds, with the high-event group manifesting the highest levels of depressive symptoms. Parallel findings were obtained when the teachers were similarly divided into low-, medium-, and high-strain groups on the basis of their scores on the fall Strain Scale.

Table 1

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean CES-D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Summer</td>
</tr>
<tr>
<td>Low events</td>
<td>11.22</td>
</tr>
<tr>
<td>Medium events</td>
<td>12.85</td>
</tr>
<tr>
<td>High events</td>
<td>11.50</td>
</tr>
</tbody>
</table>

Test statistics

<table>
<thead>
<tr>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>2, 244</td>
<td>22.04*</td>
<td>.001</td>
</tr>
<tr>
<td>2, 207</td>
<td>5.68#</td>
<td>.01</td>
</tr>
</tbody>
</table>

*aTukey tests indicated that each group was significantly different from every other group, p < .05.

Figure 1. Depressive symptoms in female teachers in the low-, medium-, and high-event groups. Significant group effects were found in the fall (p < .001) and spring (p < .01) but not the summer. Note. 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.

It is notable that the trajectory of symptom scores beginning with the summer and progressing into the fall and spring took on a fan-like shape (see Figure 1). A priori tests involving correlated t-tests were consistent with expectations: Within the high-event group the CES-D increased significantly from summer to fall (p < .001) and from summer to spring (p < .001); within the low-event group the CES-D declined significantly from summer to fall (p < .004) and from summer to spring (p < .05); within the medium-event group there were no significant changes. All tests were one-tail. In summary, women who obtained jobs in the most difficult schools, the schools with the highest levels of adverse events (and the lowest levels of the reverse-scored positive events such as praise from a parent or administrator), showed an increase in depressive symptoms from the summer. Women who obtained jobs in schools with the fewest reported adverse events (and the most positive events) showed decreases in symptom scores. Finally, the symptom scores of women who obtained jobs in schools that were intermediate in adversity remained about the same.

Regression analyses. In order to obtain more precise estimates of the effects school conditions exerted on depressive symptoms, ordinary least squares (OLS) regression analyses were conducted. The fall CES-D was regressed on the Episodic Stressor Scale as well as a number
of control variables including preemployment CES-D. The other control variables included social class of origin, life events, marital status, race, social support, and age, all factors having known links to depressive symptoms (Schonfeld & Ruan, 1991). Only the preemployment CES-D \( (B = .49; \beta < .48, p < .0001) \) and the Episodic Stressor Scale were significantly related to outcome \( (B = 8.33; \beta = .45, p < .0001) \). With the Strain Scale replacing the Episodic Stressor scale in the OLS analyses, the Strain Scale was significantly \( (B = 7.34; \beta = .41, p < .0001) \) related to the CES-D controlling for all other factors.

The unstandardized regression weight for the Episodic Stressor Scale reveals that a unit increase, as in the difference between classrooms in which different types of episodic events occur at a rate of about once per month (a scale score of 1) and classrooms in which such stressors occur at a rate of about once per week (scale score of 2), was, on average, associated with an 8.3-point adjusted (for preemployment symptoms, etc.) increase in the CES-D. Considering that a score of 8 is the median score in epidemiologic surveys of community residents (Schonfeld, 1990) and that a score of 16 is considered a marker of clinical significance (Weissman et al., 1977), the adjusted unit increase represented by the unstandardized regression weight is sizable.

As expected, with the preemployment CES-D (depressive symptoms) in the regression equation, the control variables were not significantly related to the postemployment symptoms. The other control variables' relation to the fall CES-D was absorbed by the strong relation between preemployment and postemployment depressive symptoms (see Cohen & Cohen, 1983). The \( R^2 \) for the equation containing the control variables but not the Episodic Stressor Scale was .25; the \( R^2 \) increase associated with entering the Episodic Stressor Scale (Strain Scale) into the regression equation last was .20 (.16), further supporting the view that the relation between working conditions and depressive symptoms is strong.

Although the OLS analyses included control variables (e.g., summer CES-D, social support) that were measured earlier than the fall CES-D, the analyses as they pertain to the relation of working conditions to symptoms were concurrent: Fall symptoms and school conditions were measured contemporaneously. The OLS analyses could not, by themselves, be used to rule out the hypothesis that preexisting depressive symptoms somehow "cause" or pave the way for the occurrence of hypothesized school-related stressors, a version of the event proneness model described by Dohrenwend and Dohrenwend (1981). Examples of event proneness explanations are provided in the Discussion section below.

An event proneness model would predict a significant positive correlation between preexisting symptoms and later adverse school conditions. The zero-order correlations between the summer CES-D and the two fall and two spring measures of the school environment (Episodic Stressor and Strain Scales) were, however, nonsignificant \( (r \leq 1.081) \). These correlations and the earlier described analyses of variance of the summer CES-D suggest that adversity in the school environment was more or less independent of preexisting symptoms.

**Structural equation models.** Structural equation models of possible causal links between school conditions and depressive symptoms were developed with the help of LISREL (Hayduk, 1987; Joreskog & Sorbom, 1989) software. To simplify model building, all control variables except the summer CES-D were excluded in view of OLS analyses that failed to demonstrate significant effects for the other variables. In order to examine school-environment effects on symptoms throughout the teachers' first academic year, the fall and spring Episodic Stressor and Strain Scales were included in the model building.

Both the fall Episodic Stressor and Strain Scales served as indicators of a latent fall (Time 1) school-environment variable (Env1 in Figures 2 and 3). In parallel fashion, both spring school environment scales served as indicators for the latent spring (Time 2) school-environment variable (Env2 in the Figures). LISREL allowed the investigator to force Env1 and Env2 to assume the same units as the Episodic Stressor Scale. Since each (Time 1 and 2) Env factor had two indicators, LISREL was allowed to estimate the error term for every one of the four indicators. On the theoretical grounds that the Strain Scales measure ongoing
stressors, the epsilon terms for the two Strain indicators were the only two error terms that were allowed to be correlated.

The summer (preemployment), fall, and spring CES-Ds served as indicators for the Time 0, 1, and 2 Symptoms factors, respectively (Figures 2 and 3). Each Symptoms factor was forced to assume the same scale as its CES-D indicator. An error term, derived from the reliability coefficient, for each CES-D indicator was entered into the model.

Time 0 (preemployment) Symptoms, in this analysis an "instrumental" variable, made possible the estimation of reciprocal effects (Kenny, 1979). The simultaneous effects model depicted in Figure 2 fitted the data satisfactorily \( \chi^2(8) = 11.52, \ p < .17 \). The model shows reciprocal effects between symptoms and the school environment in the fall (Time 1) and again in the spring (Time 2). The fall environment-to-symptoms coefficients in the LISREL analysis were larger than the corresponding (unstandardized and standardized) regression coefficients in the OLS analyses. Two conditions of the LISREL analysis suggest that the effect sizes associated with the Episodic Stressor and Strain Scales in the OLS analyses were underestimates. First, LISREL, unlike regression analyses, takes into account measurement error. Second, the scale units for the Symptoms and Env variables reflect the units of the CES-D and Episodic Stressor Scale, respectively.

The environment-to-symptoms effect (\( p < .001 \)) was at each time substantially greater than the small symptoms-to-environment ("halo") effects (ns). When the paths representing the halo were dropped, the model's fit was improved slightly; the paths representing the halo were retained, however, in support of the view that the effect of the school environment on teachers' depressive symptoms was substantial, controlling for individual differences in the teachers' tendencies to report on their work environments.

A rival lagged-effects model was also tested, but rejected. The chi-square statistic for the rival model indicated that this lagged model did not fit the data (see Figure 3; \( \chi^2(10) = 85.57, \ p < .0001 \)). The effect of the fall school environment on spring symptoms, however, was stronger than the effect fall symptoms exerted on the spring school environment.

The OLS and LISREL analyses together suggest that the school environment exerts effects not long after the commencement of the academic year. The LISREL analyses also reveal significant effects from Time 0 to Time 1 and Time 2 symptoms and from Time 1 to Time 2 symptoms, indicating that depressive symptoms, whether present before the teachers went to work or provoked by adversity in the school environment, tend to endure.

**Discussion**

Three sets of analyses were conducted. First, one-way analyses of variance and correlated t-tests linked depressive symptoms to school con-
ditions. The analyses revealed that women who worked in the most adverse school environments showed the most depressive symptoms although there were no preemployment differences on the CES-D. By contrast, women who obtained jobs in the “best” schools tended to show the fewest symptoms. Correlated t-tests suggested that in relation to preemployment symptoms adverse school conditions may have detrimental effects on mental health and more benign work environments may be related to better mental health.

Second, the OLS analyses revealed that the effect of school conditions on symptoms was (conservatively!) quite sizable when other risk factors were controlled. The OLS analyses as they pertain to the relation between school conditions and depressive symptoms could not, however, rule out an event-proneness explanation of the findings. One plausible event-proneness explanation encompasses the view that teachers with preexisting depressive symptoms exercise poor classroom management skills, thus creating the conditions in which stressors like pupil fighting thrive. Another version of the event-proneness hypothesis holds that, given their pessimistic outlook, teachers with high prior levels of depressive symptoms tend to overreport the occurrence of school-related stressors. Zero-order correlations, however, indicated that preemployment depressive symptoms were not related to any of the four (fall and spring) school environment variables, findings that are incompatible with event-proneness explanations of the link between teachers’ symptoms and adverse school conditions.

Finally, the LISREL analysis indicated that the causal paths from the school environment to symptoms dominated the paths from symptoms to the environment. The reciprocal path model fitted the data markedly better than the model that included lagged paths. Both the OLS and the LISREL analyses suggest that the effects of school conditions on symptoms are relatively immediate.

The link between adverse school conditions and depressive symptoms is not surprising. Both qualitative (Blase, 1986) and quantitative research (Schonfeld, 1990) indicate that many teachers’ work environments may be characterized by danger, disappointment, and lack of control, risk factors for depressive illness (Seligman, 1975). These findings are consistent with life stress research indicating that events that demonstrate to the individual a strong sense of personal disappointment and thwarted goals are related to elevated risk for clinical depression (e.g., Brown & Harris, 1989). By the same token, school conditions in which teachers are free of danger and allowed to exercise control over curricular and other matters may foster high morale.

In a final note on methodology, three features of the study design improved the quality of the research. First, by creating job environment variables that were based on neutral self-reports (Kasl, 1987; Schonfeld, 1990), school conditions could be assessed relatively independently of prior symptoms. Second, the measurement of depressive symptoms prior to the women’s entry into the work force helped in evaluating an event-proneness explanation of the findings. The advantage of research designs that follow newly employed, in comparison to veteran, teachers is the opportunity such designs afford in obtaining preemployment measures of key factors that are related to outcomes (see Schonfeld & Ruan, 1991, for a discussion of the importance of preemployment measures of health). Such preemployment measures constitute potential instrumental variables that could help to disentangle the network of effects that develop after individuals enter the work force.

Third, this study and the author’s earlier study of veteran teachers (Schonfeld, 1990) suggest that there are advantages for teacher stress researchers when they use instruments standardized in other samples. This observation does not constitute a plea for instrument conformity throughout the teacher-stress research community. Investigators should always remain free to develop and try out new instruments, especially when the new instruments measure highly specific local conditions. It should, however, be recognized that the use of QES job-satisfaction items or the CES-D would allow for extensive comparisons with samples of adults with a wide variety of occupational and life histories. Moreover, the use of standardized instruments to measure both independent (e.g., the school environment) and dependent variables (e.g., psychological distress, job
satisfaction) across different samples of teachers would help to make cross-sample comparisons of effect sizes more meaningful. Such a situation will contribute more to progress in teacher-stress research than would a situation in which many different investigators measure outcomes with a variety of stress and burnout instruments.

References


Received 11 November 1991