Journal of Occupational Health Psychology
1996, Vol. 1, No. 2, 145–169

1076-8998/96/$3.00

Longitudinal Studies in Organizational Stress Research: A Review of the Literature With Reference to Methodological Issues

Dieter Zapf and Christian Dormann
University of Konstanz

Michael Frese
University of Amsterdam

Demonstrating causal relationships has been of particular importance in organizational stress research. Longitudinal studies are typically suggested to overcome problems of reversed causation and third variables (e.g., social desirability and negative affectivity). This article reviews the empirical longitudinal literature and discusses designs and statistical methods used in these studies. Forty-three longitudinal field reports on organizational stress were identified. Most of the investigations used a 2-wave panel design and a hierarchical multiple regression approach. Six studies with 3 and more waves were found. About 50% of the studies analyzed potential strain-stressor (reversed causation) relationships. In about 33% of the studies there was some evidence of reverse causation. The power of longitudinal studies to rule out third variable explanations was not realized in many studies. Procedures of how to analyze longitudinal data are suggested.

Occupational stress research has been quickly developing during the last two decades. This has led to many excellent studies—many of them longitudinal—that produced a wide range of knowledge in this field. However, even though this development is welcomed, we argue that the power of longitudinal designs has not yet been fully realized. This is because designs and analyses of the stress investigations are far from optimal. Thus, the review of the literature in this article on longitudinal stress research from a methodological angle is supposed to show the pitfalls and difficulties in design and analysis strategies but should also point out good examples of longitudinal research. Obviously, we are not the only ones who present a critique of stress research (e.g., Brief, Burke, George, Robinson, & Webster, 1988; M. J. Burke, Brief, & George, 1993; Contrada & Kranz, 1987; Frese & Zapf, 1988; Kasl, 1978, 1986, 1987; Kessler, 1987; Leventhal & Tomarken, 1987; Spector, 1992; Zapf, 1989). However, this critique differs from other articles because it is oriented toward that part of the literature that renders the best data and the most knowledge: the longitudinal studies.

Dieter Zapf and Christian Dormann, Faculty of Social Sciences, University of Konstanz, Konstanz, Germany; Michael Frese, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands.

We thank Harry Garst for a critical reading of a prior version of this article.

Correspondence concerning this article should be addressed to Dieter Zapf, Sozialwissenschaftliche Fakultaet, Universitaet Konstanz, Postfach 5560, D42, D-78434, Konstanz, Germany. Electronic mail may be sent via Internet to dieter.zapf@uni-konstanz.de.

In this article we first argue that because stress research is a multicausal field, causal effects of specific stressors on strain cannot be very high. Second, we summarize those methodological aspects that make it difficult to detect causal effects. Third, we summarize the problems of causal inference that are associated with typical research designs and applied statistical procedures. Fourth, we review existing longitudinal studies to assess how they addressed the question of causal inferences.

On the Size of Stressor-Strain Relationships

Some authors have argued (e.g., Cohen & Edwards, 1989; Kasl, 1978; Nelson & Sutton, 1990; Rabkin & Struening, 1976; Vossel, 1987) that investigators should be disappointed by the small amount of explained variance in stressor-strain correlations. In contrast to these authors, we argue that a small correlation should be expected both from a content and from a methodological point of view (cf. also Leventhal & Tomarken, 1987; Semmer, Zapf, & Greif, in press). Many factors influence physical and mental health. Among the factors discussed in the literature are physical constitution, past diseases or accidents (Kasl, 1983; Lipowski, 1975), personality traits (Depue & Monroe, 1986; Semmer, in press), health behaviors such as smoking or drinking (Steptoe & Wardle, 1995), leisure time stressors (Bamberg, 1992), family stressors (Gutek, Repetti, & Silver, 1988), environmental factors such as air pollution (Seeber & Iregren, 1992), age (Birdi, Warr, & Oswald, 1995; Warr, 1992), sex (Frankenhaeuser, 1991), social class related factors (such as financial or housing situation, Marmot & Madge, 1987; Syme & Berkman, 1976), social stressors (Zapf & Frese, 1991), critical life events (Dohrenwend & Dohrenwend, 1974), individual resources (e.g., coping behavior; Semmer, in press), institutional resources such as control (Frese, 1989) or social support (Cohen & Wills, 1985), and, finally, work-related stressors.

This list of 15 areas can certainly be extended. If we assume for the moment that we can measure all 15 causal areas error free and if we further assume that each of them has an equally strong effect, then any one area (e.g., work stressors) can explain only a maximum of less than 7% of the variance (corresponding to a correlation of .26). If attenuation because of unreliability is taken into consideration, the observed correlation would even be lower, decreasing to .21, if stressors and strains are measured with reliabilities of .80. Moreover, it should be noted that there are multiple work stressors. Thus, this correlation refers to the multiple correlation coefficient, and the impact of any single work stressor should be much lower (cf. Semmer et al., in press) and only the measurement of the full set of stressors leads to the multiple correlation mentioned earlier. It follows from the multifactorial determination of health that any one factor cannot explain the "lion's share" of variance in health variables. Thus one suspects that high stressor-strain correlations, for example .30 and even .50 and higher (e.g., R. J. Burke & Greenglass, 1995; Caplan, Cobb, French, Harrison, & Pinneau, 1975; Poulin & Walter, 1993; Tetrick & LaRocco, 1987), are due to common method variance leading to an inflation of the reported correlations (M. J. Burke et al., 1993; Spector, 1992; Spector & Brannick, 1995; Zapf, 1989). Such high correlations occurred only with self-report data of independent and dependent variables. It is interesting to note that the causal effects reported in longitudinal studies and analyzed with reasonable methods never reach such magnitudes.

Moreover, there are other factors that lead to a reduction of stressor-strain relationships. One issue is related to moderator variables. Because of moderating effects, causal stressor-strain relationships may be valid only for subgroups of individuals, for example, for individuals with insufficient coping skills (Parkes, 1990), commitment (Begley & Czajka, 1993), Type A behavior (Caplan & Jones, 1975; Orpen, 1982), or subgroups of jobs with low control (Frese & Semmer, 1991; Wall, Jackson, Mullarkey, & Parker, in press) or with low social support (Cohen & Wills, 1985; Frese, 1995; Fusilier, Ganster, & Mayes, 1987). The moderating effects may lead to a substantial amount of explained variance in certain subsamples but attenuate the effects in the total sample (cf. Wall et al., in press).

Another reason for attenuated stressor-strain relationships is the healthy worker effect (Waldron, Herold, Dunn, & Staum, 1982). Because it is likely that workers who are seriously ill will stop working, there is a restriction of range in the dependent variable (health). A similar effect occurs when a worker quits the job because he or she feels that the job is too stressful (cf. Kessler, 1987).

Another group of reasons that leads to low empirical causal stressor-strain relationships is the specific time relationships of stressors and strains (Frese & Zapf, 1988). In that article the authors presented different models of how a stressor may affect ill health in the course of time. The authors discussed five types of the so-called exposure time model, which assume that a stressor has some impact on psychological and psychosomatic dysfunctioning, however, in various ways:

1. According to the stress reaction model the impact of a stressor increases psychological dysfunctioning with exposure time. Once the stressor is removed, there is an improvement in psychological functioning. An assumption of this model is that the increase of a stressor would have the same impact on psychological dysfunctioning as its decrease, simply the sign of the impact should be reversed.

2. According to the accumulation model, strain comes about as a result of the accumulation and it does not go away even after the stressors have been reduced. Such effects might appear in shift workers after a certain "breaking point" has been achieved (Frese & Okonek, 1984). The mechanism of the accumulation model must be considered because longitudinal designs rely on changes of the respective variables. According to the accumulation model, however, increases of stressors should have to be treated differently from decreases (see below).

3. The dynamic accumulation model, in contrast to the accumulation model, assumes that there is an inner dynamic that leads to a further increase even after the stressor has been removed, although this increase is probably decelerated. This may be so because the original stressor had a general weakening effect on the psychophysical system so that new stressors have a higher impact than normal. In such a case an individual may also be more vulnerable to stressors not related to the work situation at all (e.g., in his or her leisure time; Bamberg, 1991).

4. The adjustment model is related to the stress reaction model because there is at first a linear

increase of dysfunctioning with the duration of the stressor. However, after a certain point, an adjustment process sets in and the dysfunctioning decreases although the stressor is still present. The adjustment model can be described quite well within Lazarus' theory (Lazarus & Folkman, 1984): One develops coping strategies toward the stressors (e.g., denial or help seeking), which reduce ill health. If one does not control for coping strategies (or uses job tenure as an indirect control) one may find stressor-strain relationships that underestimate the true causal association. For example, Schonfeld (1992) controlled for job tenure in a sample of newly employed teachers who did not yet accommodate to the stressors.

5. The sleeper effect model implies that dysfunctioning appears a long time after exposure to the stressor, as for example, in post-traumatic stress disorders (Leymann & Gustafsson, in press; Theorell, Leymann, Jodko, Konarski, & Norbeck, 1994). It is possible that the stressor is not present any more when dysfunctioning appears. It also assumes that the effect is higher with longer exposure to and stronger intensity of the stressor.

Most of the longitudinal studies reviewed for this article typically used some sort of linear models (regression or analysis of variance). These methods imply a parallelism between changes in stressors and strains. That is, if a stressor goes up for x units, strain goes up for y units. If a stressor goes down for x units, strain goes similarly down for y units. As some of the above-mentioned time course models suggest, this is not necessarily so. Some models do not assume a decline of strain if the stressors are reduced. Other models suggest to treat the time frame of increasing stressors differently from decreasing stressors. One would then expect causal relationships for increasing stressors, and, for example, zero relationships in the case of decreasing stressors or different time frames between developing and recovering from ill health. Thus, linear data analysis methods would usually underestimate the true strength of the stressor-strain association. We are not aware of articles taking such considerations into account.

In all, there are many reasons why causal effects in organizational stress research cannot be very high. If effects are not high to begin with, then problems of the empirical design such as a wrong time lag, third variables, measurement issues, or insensitive statistical procedures can easily contribute to the failure to detect causal relationships. This has to be kept in mind when we discuss strengths and weaknesses of designs and statistical methods below.

Problems of Testing Causal Effects in Longitudinal Studies

The bulk of empirical research (clearly more than 90%) on stressors and health is cross-sectional. The weaknesses of such a design are widely acknowledged and researchers are well aware that it is usually impossible to demonstrate causal relationships in such designs. It has been suggested that longitudinal studies can reduce the problems associated with cross-sectional studies, in particular, the problem of reverse causation and the treatment of third variables.

In line with Cook and Campbell (1979), we speak of a causal effect as existing (a) if there is covariation of the stressor with ill health, (b) if the stressor appeared before ill health developed, and (c) if other plausible explanations can be ruled out (e.g., a weakening of the body's constitution that led first to more stress and later to an increase in the manifest ill health). Thus, causal inferences cannot be proven (H. M. Blalock, 1961; Dwyer, 1983; Holland, 1986) but can be made plausible by ruling out alternative explanations.

Longitudinal studies have one important restriction, namely, that it makes no sense to assume that a variable at Time 2 has an impact on a variable at Time 1. This restriction can be used to argue for a causal inference under certain circumstances. If there is a relationship between two variables, there are basically only two alternative explanations to a causal inference of x on y: First, there is a reverse causation of y on x, and second, a third variable z influences both x and y and thus produces the relationship between x and y. Typically, longitudinal research is seen to be able to exclude reversed causal relationships hypotheses. However, as argued below, interpretation of longitudinal data are also not immune against third variable problems.

Reverse Causal Hypothesis

For the reversed causation hypothesis, there are at least two plausible groups of explanations possible: the so-called drift-hypothesis and the true strain-stressor hypotheses. First, according to the drift-hypothesis (Frese, 1982; Kohn & Schooler, 1983) individuals with bad health drift to worse jobs, for example, first by becoming unemployed and then by getting a worse job because of their personal record of frequent absenteeism. Or, people with high absenteeism are transferred to positions with less responsibilities, which go along with higher work stressors. Moreover, in selection, preferred employees are those with higher levels of social competence, self-esteem, and stress tolerance for skilled jobs. Thus, healthier people get the better jobs. In a longitudinal study this would lead to a causal impact of strain on stressors. Second, stressors may sometimes be affected by strain, as exemplified in the relationship between social stressors and depression. One could argue that the level of depression is related to the quality of important personal relationships. Depressed people tend to assess their environment more negatively, thus contributing to a more negative group climate (Beck, 1972). This may cause an increase in conflicts between coworkers in the group, leading to higher social stressors at work.

Thus, reversed causal influences of strains on stressors are plausible (Leventhal & Tomarken, 1987). There are also cases where reciprocal causal relationships seem plausible (positive feedback loop). For example, an increase of social stressors caused by the worker's depression can contribute in turn to an increase of depression.

Third Variables

The other problem for the interpretation of causal relationships is third variables. Third variables may affect stressors and strains by using the same methods, thus producing common method variance, for example, through social desirability, acquiescence, or negative affectivity. Other third variables affect stressors and strains independently from the method used, for example social status (bad housing conditions, higher environmental pollution, financial problems), education, sex, and age. For longitudinal research the stability of third variables over time is crucial. Occasion factors, background variables, and nonconstant variables can be differentiated.

Occasion factors. Occasion factors are hypothetical and usually unmeasured variables that have an impact on stressor (independent) and strain (dependent) variables. As Dwyer (1983) emphasized, such "occasion factors are the Achilles' heel of cross-sectional designs, where they must be measured and controlled explicitly to avoid bias in the estimation of structural coefficients" (p. 360). Examples of occasion factors are weather, time of the day, or mood variables. If participants of a study are in a good mood, they might, for example, see themselves as less depressed or report less psychosomatic complaints and simultaneously assess the work stressors as less pronounced in contrast to participants in a bad mood who may exaggerate strain and stressors. This creates an artificial correlation between stressors and strain. Because it is unlikely that participants are in a similar mood again when they answer the same questionnaire several months later, there are no true occasion factor correlations over time.

While influences of mood can exaggerate cross-sectional stressor-strain relationships, there is a twofold effect for longitudinal studies: On the one hand, mood works like error variance, thus, attenuating the observed effects. On the other hand, a part of the effect is carried by the stability of the dependent variable. Assume for example a true zero correlation between stressor x_1 and strain y_1 and no causal effect of x_1 on y_2, but an effect of mood at Time 1, which leads to an observed correlation of .30 between x_1 and y_1. If the stability of the dependent variable (y_1, y_2) is .50, then one would still find an observed correlation of .15 (.30 times .50) between x_1 and y_2, although the true effect is zero. This effect, however, disappears when y_1 is partialled out from the correlation between x_1 and y_2.

Background variables. Another type of third variables is assumed to be completely stable over time. They are called *background variables* (Dwyer, 1983), particularly sociodemographic variables such as age, sex, education, social status (Frese, 1985), or personality traits such as negative affectivity (Brief et al., 1988). The effects of background variables are carried over time; that is, the correlation between stressors Time 1 and strain Time 2 is as exaggerated as the correlation of stressors Time 1 and strain Time 1. However, by partialling out strain Time 1, as is done in hierarchical multiple regressions, such third variable effects are controlled for.

Nonconstant variables. The most problematic type of variables are those that have some stability over time and that influence both the independent and dependent variable (Dwyer, 1983). One can argue that social desirability does not have to be constant over time (although it will always have a certain amount of stability). We assume that social desirability is influenced by a person's sense of insecurity. Sense of insecurity may vary over time because of reasons in and outside the person, and, therefore, a similar variation in social desirability will occur. This means that varying social desirability may affect stressors and strains differentially over time.

We refer to background variables and to nonconstant variables that have not been explicitly measured as *common factors* (Dwyer, 1983).

Strategies to Test Causal Effects

In the following discussion we summarize the reasons why longitudinal designs do not automatically prove causality and do not automatically reject third variable explanations (see, e.g., Cook & Campbell, 1979; Dwyer, 1983; Kenny, 1979; Kessler, 1987; Kessler & Greenberg, 1981). We discuss typical strategies to analyze causal effects, and to concretize what we mean we apply the various analysis strategies to the data reported by Frese (1985). This study looked at the causal relationship of psychological stressors at work with psychosomatic complaints based on a two-wave design with a time lag of 16 months and a sample of 79 blue-collar workers. Both psychological stressors (a composite measure of time pressure, uncertainty, danger of accidents, organizational problems and environmental stressors; only the results of the questionnaire form, which is one of the three measures by Frese, 1985, is analyzed here) and psychosomatic complaints were measured with a questionnaire (for details, see Frese, 1985).

Stressor Time 1 and Strain Time 2 Designs

Some authors (e.g., Richter, Schirmer, & Dettmar, 1989) have analyzed causal effects with designs that measured the stressors at Time 1 (referred to as x_1) and strain at Time 2 (referred to as y_2). Correlational or multiple regression analyses were used to analyze such data (cf. Figure 1). Such an analysis strategy does not have advantages compared with a cross-sectional design. The fact that a variable was measured later than another variable is little proof of a causal relationship. Consider, for example, a father's social status and his son's income. Even if the father's social status was measured years after the son's income, nobody would assume that the son's income had a causal impact on the father's social status. If one would like, for example, to find support for the drift hypothesis in Frese's (1985) data, this analysis strategy would find a significant ($p < .01$) correlation of .30 between complaints Time 1 and stressors Time 2. This would be taken as evidence that the drift hypothesis was supported. However, as we shall see,

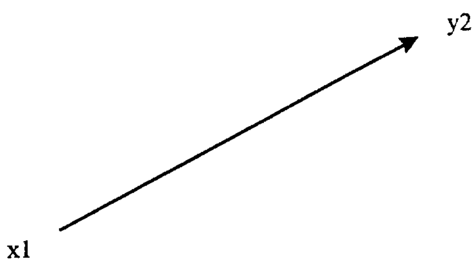

Figure 1. Effects of x_1 on y_2.

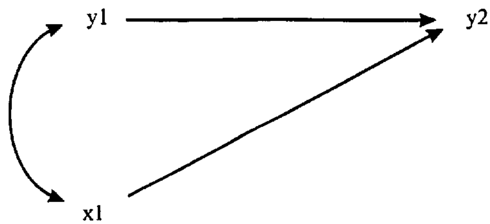

Figure 2. The hierarchical multiple regression approach: Regression of y_2 on y_1 and x_1.

there is no evidence for the drift hypothesis in this study, but there is a strong effect in the opposite direction. In summary, with such a design it is impossible to rule out alternative causal hypotheses.

Incomplete Two-Wave Panel Design

A design measuring the dependent variable only at Time 2 was not used frequently in the literature. More often, the independent variable x_2 was measured as well. However, the analyses often did not take x_2 into consideration. Usually, hierarchical multiple regressions of the following structure were performed: First, third variables such as age or sex and, in addition, strain Time 1 (y_1) were introduced into the equation to partial out their effects. Second, stressor Time 1 (x_1) was introduced into the equation. The dependent variable was strain Time 2 (y_2). The causal effect was seen to be supported if the R^2 change of the last step was significant (cf. Figure 2).

With regard to reverse causation and third variables problems the following can be said:

1. Occasion factors and background variables are ruled out as a source of spurious dependency between x_1 and y_2 through partialling y_1. Unmeasured nonconstant third variable (common factor) effects, however, are not controlled with such a design.

2. The reverse causal effect is not tested. This analysis strategy cannot exclude that the regression of stressors Time 2 on strain Time 1 with stressors Time 1 controlled will also lead to significant effects and will be even higher than the effect of stressors on strain.

3. A relationship between stressors and strain based on "synchronous effects" will not be detected if the independent variable is not stable. It should be noted that synchronous does not necessarily mean that the independent variable immediately affects the dependent variable. Rather, synchronous can embrace a considerable period of time depending on the design. For example, Kohn and Schooler (1982) used a time lag of 10 years. Were there a true effect of time

pressure on anxiety with a time lag of 1 year, a model comprising a synchronous path of time pressure on anxiety at Time 2 would approximate the true effect better than a model with a 10-year lagged effect of time pressure Time 1 on anxiety Time 2. In studies that use a time lag of 1 year, a true 3-months time lag should be better represented by synchronous effects than by lagged effects. Therefore, not being able to test these types of synchronous stressors-strain effects is a serious weakness of this approach.

4. In addition to the objections just mentioned, there are further problems such as assumptions of uncorrelated measurement errors (cf. Dwyer, 1983).

Applying hierarchical multiple regression to Frese's (1985) data leads to the results in Table 1: There is a significant effect of stressors on complaints but no effect of complaints on stressors. The regression analyses show that the correlation between complaints Time 1 and stressors Time 2 can be explained by the cross-sectional correlations and the stabilities of the variables.

Full Two-Wave Panel Design

Several strategies are used to analyze data of a two-wave panel design: (a) cross-lagged panel correlation analysis; (b) hierarchical regression with lagged effects, hierarchical regression with synchronous effects, and comparing the regressions of y_2 and x_2; (c) simultaneous estimation procedures such as structural equation approaches.

Cross-lagged panel correlation (CLPC) technique. The CLPC technique (R. Blalock, 1962; Campbell, 1963; Campbell & Stanley, 1963) comprises six correlations: the cross-sectional correlations at Time 1 and Time 2, the stabilities or autocorrelations $r(x_1, x_2)$ and $r(y_1, y_2)$, and the cross-lagged correlations $r(x_1, y_2)$ and $r(y_1, x_2)$. The core element of the CLPC technique is the statistical comparison of the two cross-lagged correlations of $r(x_1, y_2)$ and $r(x_2, y_1)$ in Figure 3, for example, tested with Steiger's (1980)

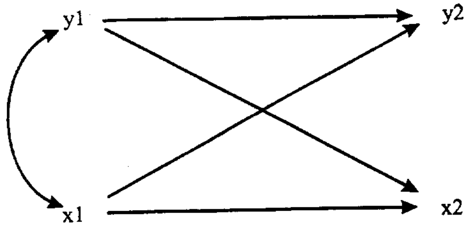

Figure 3. Two-wave panel design with lagged effects.

Formula 15. Researchers who use this technique are searching for the causal predominance of either of the variables.

Although the logic of CLPC is intuitively appealing, several articles have argued against using the CLPC technique (e.g., Dwyer, 1983; Feldman, 1975; Link & Shrout, 1992; Locascio, 1982; Rogosa, 1980; Williams & Podsakoff, 1989).

It can be algebraically shown that the difference of the cross-lagged correlations is directly dependent on the stabilities of x and y. Thus, a difference in cross-lagged correlation opposite to the true direction can appear if the difference in stability path coefficients more than offsets the difference in cross-lagged causal parameters (Locascio, 1982, p. 1031). Other problems are differences in variances and differences in cross-sectional correlations indicating that assumptions of the CLPC approach are not met (see Rogosa, 1980, or Williams & Podsakoff, 1989, for algebraic details).

It has also been shown that CLPC is only able to reject certain types of third variable models. A significant difference of the cross-lagged correlations leads to the conclusion that the correlation between x and y is not only due to a synchronous common factor. Critical assumptions that have to be made for this conclusion is that the causal influence does not change over time and the common factor exerts its effects on both variables with identical time lags. The

Table 1
Application of the Hierarchical Multiple Regression to the Frese (1985) Data (Standardized Regression Coefficients)

Independent variable	Dependent variable: Complaints at Time 2 β	Independent variable	Dependent variable: Stressors at Time 2 β
Complaints at Time 1	.66**	Stressors at Time 1	.64**
Stressors at Time 1	.19*[a]	Complaints at Time 1	.08[b]

[a] Adjusted $R^2 = .56$.
[b] Adjusted $R^2 = .41$.
* $p < .05$. ** $p < .01$.

CLPC technique is, however, not able to reject occasion-factor models that lead to an artifactual relationship of the variables at Time 2. Moreover, although differences in cross-lagged correlation can sometimes indicate a causal interdependence, the assessment of the direction of the causal influence is exclusively based on theoretical assumptions. This is so because the logic behind this technique is much less intuitive as it seems to be from the simple comparison procedure (see Kenny, 1975). Also it is difficult to determine an effect size from the results of the cross-lagged panel technique. Finally, the CLPC-technique does not provide information on the time lag, although in practice, significant effects are mostly interpreted as lagged effects by most of the researchers. Rather, they can also stem from synchronous effects.

Suggestions have been made to partial out the stabilities (e.g., Pelz & Andrews, 1964). This, in our opinion, implies a shift from the CLPC logic toward the logic that is inherent in hierarchical multiple regression, which also tests the reversed hypothesis.

Applying CLPC to the Frese (1985) data has already been done in the original paper and needs not be repeated here. Because *stationarity* (i.e., a stable causal structure over time) and equality of the variances over time existed (in addition, stabilities were partialled out), it was legitimate to apply a CLPC procedure. There were significant effects of stressors on complaints, shown by comparing the lagged correlations of $r = .47$ (stressor Time 1 − complaints Time 2) with $r = .30$ (complaints Time 1 − stressors Time 2). Thus, there is a causal effect of psychological stressors on psychosomatic complaints.

Hierarchical regression analyses. The general structure of multiple regression equations for the analysis of longitudinal data has already been briefly described. The effects of third variables are controlled by including them into the regression in the first step. The more potential third variables are included, the more explanations of spuriousness can be ruled out. However, there is still the problem that synchronous causal effects cannot be identified by this approach. Comparisons between reversed hierarchical regression with synchronous effects have also been suggested to identify synchronous causal relationships. In such regressions, the independent variables are put into the hierarchical regression in the following order: (a) third variables, (b) strain Time 1, (c) stressor Time 1, (d) stressor Time 2. This approach is prone to occasion factors because their influence is not neutralized as in the case of lagged effects. However, it is advantageous that changes in the independent variables are related to changes in the dependent variable. Although such a design allows for comparisons of the stressor-strain hypothesis with the reversed hypothesis, this is often not done (e.g., Billings & Moos, 1982) because the opposite hypothesis is a priori rejected by the authors.

Structural equation approaches. Structural equations approaches (most often referred to as LISREL analyses; Jöreskog & Sörbom, 1989, 1993) have several advantages with respect to the other analytical approaches discussed so far, although they are sometimes criticized (e.g., Brannick, 1995, for application problems and wrong interpretations).

Four advantages of LISREL models should be mentioned:

1. Measurement errors can be accounted for by the introduction of measurement models. Causal relationships between variables are modeled on the basis of latent constructs that are considered to be error free.
2. Structural equation models allow simultaneous estimates of causal relationships for all latent variables. Thus, multivariable–multiwave models can be analyzed.
3. Reciprocal relationships can be introduced into the models.
4. Various method and third variable problems can be modeled such as occasion factors and common factor models that account for effects of unmeasured third variables.

In short, everything that can be done with cross-lagged panel correlations and regression analyses can also be done with structural equation models (see introductions by Dwyer, 1983, Kessler & Greenberg, 1981, or Williams & Podsakoff, 1989). A reanalysis of Frese's (1985) data is presented below because it also looks at the issue of spuriousness.

The Treatment of Third Variables in Longitudinal Studies

The treatment of third variables in longitudinal designs is not easy. Whereas in cross-sectional studies researchers are well aware of this problem, little attention is paid to ruling out third variable explanations in longitudinal investigations (Link & Shrout, 1992). Effects in two-variable panel designs can always be explained by models that assume no causal relationship between these variables. That is, without additional assumptions, it is difficult to demonstrate unambiguously that causal relationships exist in two-wave panel studies (Dwyer, 1983; Link & Shrout, 1992).

A typical path model representing causal effects in a two-wave design was shown in Figure 3. An alternative model representing spuriousness in such a design is presented in Figure 4.

On the basis of statistics alone, one cannot reject the spuriousness model of Figure 4 (Dwyer, 1983; Kenny, 1975, 1979; Link & Shrout, 1992) because seven parameters (a–g in Figure 4) are involved in the associations of the observed variables x_1, x_2, y_1, and y_2. However, a cross-lagged panel design provides only six observed correlations. Consequently, the model in Figure 4 cannot be estimated unambiguously from the empirical data. Such a model is unidentified. As a result, more than one causal model is consistent with a set of observed relationships. Therefore, any assessment of causality of a panel design rests on the logic of the inquiry and the persuasiveness of tests proposed to rule out alternative hypotheses (Link & Shrout, 1992).[1]

In LISREL analyses several theoretical models can be built and tested against each other (Bollen, 1989; Jöreskog & Sörbom, 1989). We have reanalyzed Frese's (1985) data using the LISREL approach and tested a series of models. The results are presented in Table 2. Model A, the null model, encompasses stabilities of stressors and complaints and a correlation at Time 1, but no lagged or synchronous causal paths between stressors and complaints. In Models B through E, there is one additional causal path in each case: a lagged causal effect of stressors on complaints in Model B, the reverse lagged causal effect of complaints on stressors in Model C, a synchronous causal effect of stressors on complaints in Model D, and a synchronous reverse effect in Model E. The comparison of the overall fit indexes shows that Model D with a synchronous causal effect of psychological stressors on psychosomatic complaints has an excellent model fit with the goodness of fit and adjusted goodness of fit equal to 1. A comparison of the four causal Models B through E shows that the models with an effect of stressors on complaints generally fit better than the models with an effect of complaints on stressors. It should be noted that the result of the CLPC analysis in the original article led Frese to conclude that there is a lagged effect of stressors on complaints. The LISREL analysis clearly demonstrates that there is a synchronous effect. We believe that this is a typical finding: CLPC results are usually interpreted to indicate lagged effects if the lagged correlations are significantly different. LISREL analysis, however, is able to reveal synchronous effects.

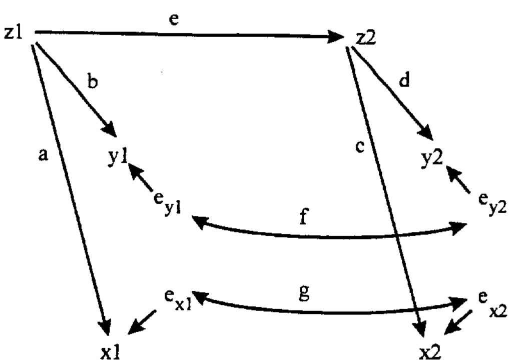

Figure 4. Spuriousness model.

In addition, Table 2 comprises five third variable models that is only a selection of many more possible models. The first four models assume that stressors and complaints are indicators of one latent construct (cf. z in Figure 4). In addition, there are the following restrictions: The variances of the latent variables are constrained to 1 (i.e., the variance does not change over time); in addition, there are equal autocorrelations of stressors and complaints ($f = g$ in Figure 4). The four models differ with regard to stationarity and the stability of the common factor (i.e., the latent third variable). Model 1 comprises proportional stationarity and free stability (i.e., the ratio of the factor loadings of stressors and complaints are equal; $a/b = c/d$ in Figure 4), and the stability of the common factor is estimated (e in Figure 4). Model 2 differs from model 1 in that there is perfect stationarity (i.e. the factor loadings of stressors and complaints on the common factor are equal at Time 1 and Time 2; $a = c$ and $b = d$ in Figure 4).[2] Models 3 and 4 differ from Model 1 in that there is one common factor that does not change over time ($e = 1$ in Figure 4).

In terms of content, the models described correspond, for example, to the critique of observed item overlap by Kasl (1978): Stressors and complaints are

[1] The problem of identification is obvious if one analyzes two-wave panel data with structural equations techniques. Testing theoretical models against the empirical data is possible only if some restrictions can be accepted. For example, a very liberal model can assume reciprocal causal effects between stressors and strains. Such a model is overidentified with one degree of freedom (e.g., a combination of Models B and C in Table 2). The addition of an occasion factor would lead to a just identified model that cannot be tested against the empirical data.

[2] There are many constraints, but most of them can be relaxed if additional waves of observations are available.

Table 2
Reanalysis of the Frese (1985) Data Using Structural Equations

Model and type	χ^2	*df*	Probability	GFI	AGFI	sRMR	AIC
	Overall fit index scales						
Causal models							
A. Null model	9.75	3	.021	.94	.81	.079	23.75
B. Stressor 1 → Complaints 2	4.12*	2	.130	.97	.87	.023	20.12
C. Complaints 1 → Stressor 2	9.49	2	.009	.95	.73	.084	25.49
D. Stressor 2 → Complaints 2	0.27**	2	.880	1.00	1.00	.013	16.27
E. Complaints 2 → Stressor 2	8.65	2	.013	.95	.95	.068	24.65
Third variable models							
Synchronous common factor models[a]							
1. Proportional; free[b]	4.44	3	.220	.97	.91	.054	18.44
2. Perfect; free[b]	5.30	4	.260	.97	.92	.058	17.30
3. Proportional; 1[b]	7.31	4	.120	.96	.89	.066	19.31
4. Perfect; 1[b]	7.70	5	.170	.95	.91	.060	17.70
5. Occasion factor model[c]	5.71	2	.058	.97	.83	.063	21.71

Note. GFI = goodness of fit; AGFI = adjusted goodness of fit; sRMR = standardized root mean squares residuals; AIC = Akaike information criterion.
[a] See Link and Shrout (1992).
[b] Indicates stationarity and stability of common factor.
[c] See Dwyer (1983).
* $p < .05$. ** $p < .01$. (Significantly better than the Null model.)

indicators of a latent construct that either changes over time or does not change. The latter model (without change), for example, corresponds to the critique of Watson and Clark (1984), Brief et al. (1988), and M. J. Burke et al. (1993) that a (stable) personality trait "negative affectivity" is responsible for stressor-strain correlations. In the particular case of the Frese (1985) data, none of the third variable models show comparable fit indexes as the causal Model D with its synchronous effect of stressors on complaints. Note, however, that the third variable models are more parsimonious. Most of their fit indexes are, for example, better than the parameters of the originally suggested causal Model B with a lagged effect of psychological stressors on complaints. If the correlational pattern is less clear than in the present case, third variable models usually produce similarly good fit indexes and clear decisions between the models are impossible. Finally, note that the occasion factor model (Model 5) is clearly inferior compared with the causal model of stressors on complaints (Model D). Thus, the hypothesis of mood or other occasion factors being responsible for the correlational pattern has to be rejected.

A Review of Longitudinal Studies

In the beginning of this article we mentioned the need for longitudinal studies in the area of stress at work. We wanted to know whether research had actually advanced and produced more longitudinal studies lately. With the help of database queries, review articles, and incidental knowledge of articles, we collected longitudinal studies using the following criteria: (a) they should be quantitative; (b) they should include more than one measurement point; (c) they should measure job-related variables such as work stressors, social stressors at work, job content variables, or work-related social support; (4) they should use variables of mental health as dependent variables. In the following discussion we concentrate on studies with a passive longitudinal design, that is, quasiexperimental designs are not discussed. We also excluded articles based on students' samples.

Furthermore, we excluded longitudinal studies on unemployment and health from our review. Space reasons and the fact that unemployment studies usually do not include more specific organizational stressors may justify why these articles were not considered. Moreover, it is our impression that the necessity for doing longitudinal research has been best understood in this field, and most of the reviews in the field of unemployment concentrate on the discussion of longitudinal research (cf. Feather, 1989; O'Brien, 1986; Warr, 1987). Studies that are based on repeated measurements but were not really interested to draw causal inferences were also excluded (e.g., Theorell, Orth-Gomér, & Eneroth, 1990). One publication was not considered because we were not able to reconstruct what the authors had really done. Finally, one article was excluded because the results of the longitudinal study were already published in another article by the same author. This literature search led to the consideration of 43 studies published in 45 articles (see Appendix A). An increase in the publication of longitudinal studies can be observed from 10 studies published until 1985, 16 studies published between 1986 and 1990, and 19 studies published from 1991 onward. However, given that the need for longitudinal designs has been emphasized repeatedly during the last two decades, and given the many advantages of using longitudinal research for the analysis of causal stressor-strain effects and the hundreds of publications in organizational stress in general, the number of longitudinal publications is less than one would expect.

Some characteristics of longitudinal studies are presented in Table 3. They are applied to the longitudinal stress studies in Appendix B. In the following, we summarize the stress studies from a methodological standpoint.

Time Lag

We identified time lags of 1 month (Daniels & Guppy, 1994; Theorell et al., 1994), 3 months (6 studies), 6 months (11 studies), and 1 year (13 studies). A time lag of about 18 months occurred five times and a time lag of 2 years occurred four times. There were only two studies with a lag of 5 years, three studies with a lag of 6 years, one study with a lag of 8 years, and 2 studies with a lag of 10 years (one study included more than 10 years). Thus, most studies analyzed time lags up to 1 year. We got the impression from the reports that organizational reasons were much more important for choosing a particular time lag than theoretical considerations. There were only a few publications that discussed the time lag problem in detail.

Designs

Most studies (25) used a full two-wave panel design (i.e., with both independent and dependent variables measured at Time 1 and Time 2). However, there were also seven investigations that measured either the independent or the dependent variable only at one time point. Five studies used a three-wave design. Finally, there were four prospective studies that collected data at Time 1 and predicted a certain event, for example, the occurrence of coronary heart disease several years later. Most of the research with a long time lag fell into this category. Within this design, two groups (event yes vs. event no) were usually distinguished and compared with the help of logistic regression or hazard analysis methods.

Statistical Analysis

Six studies regressed strain Time 2 on stressors Time 1 (most of them also controlled for background, nonconstant variables, or prior strain. In 17 cases hierarchical multiple regression of y_2 on y_1 and x_1 or some similar method was used. One investigation also considered x_2 in the regression of y_2 on y_1 and x_1 (Article 17 in Appendix B). The CLPC approach was used seven times and either Kenny's (1975) or Pelz and Andrew's (1964) recommendations were followed (Articles 7, 9, 14, 30, 31, 44, and 45 in Appendix B). Finally, 10 studies used structural equations (Articles 3, 11, 12, 15, 26, 27, 28, 29, 32, and 40; Articles 28, 29, and 32 omitted measurement models), mostly applying the computer program LISREL (Jöreskog & Sörbom, 1989, 1993).

Test of Reverse Causation

As explained above, there are basically two reasons why researchers should use longitudinal designs: First, to make a decision between hypotheses of opposite causation, and, Second, to improve the possibilities to reject third variable explanations. We discuss these points in the following sections.

We were surprised to see that in many cases, the problem of reversed causal relationships between stressors and strains was not discussed. This confirms James and James (1989) who argued that reciprocal causation has not been a popular hypothesis in organizational research. It is obvious that the six studies of the type "regression of y_2 on x_1" can tell researchers little about causal effects of strains on stressors. But even studies that used a hierarchical multiple regression approach did not usually test for reverse causation hypothesis (exceptions are articles 2 and 13 in Appendix B). Even some LISREL analyses did not take reverse causation into account (e.g., Dignam & West, 1988), although they tested a series of models. The reverse effect was, of course, tested in the 7 publications using the CLPC approach,

Table 3

Causal Effects in Longitudinal Studies With Reference to Various Statistical Techniques

Technique	Causal effects	Reversed causation	Reciprocal causation	Background variables	Nonconstant variables	Occasion factors
Regression[a]	Calculate hierarchical regression, control for prior symptoms, and examine regression coefficients	Calculate reversed regressions and examine regression coefficients	Remains unclear	Measure and introduce into equation before stressors are entered	Measure and introduce into equation before stressors are entered	Partial out dependent variable at preceding time points
CLPC	Compare the cross-lagged correlation coefficients statistically; calculate both CLPC with and without stabilities partialled out		Remains unclear	Remain problematic if their influences change over time	If (proportional) stationarity is violated, influence cannot be reliably assessed	Can be rejected by technique if stabilities are equal or partialled out
LISREL	Calculate models with synchronous and lagged effects and examine overall fit indexes and structural coefficients	Calculate reversed LISREL model with synchronous and lagged effects and examine overall fit indexes and structural coefficients	Calculate any combination of synchronous and lagged reciprocal models and examine overall fit indexes and structural coefficients	Measure and introduce into the model	Measure and introduce into the model; unmeasured nonconstant variables can theoretically be modelled when number of waves exceeds two[b]	Allow within-time correlation of structural disturbances

Note. CLPC = cross-lagged panel correlations.

[a] If predictor variables are measured with error, the regression approach may lead to heavily distorted coefficients and is therefore questionable when causal inferences are drawn from such results (e.g., Dwyer, 1983; Kenny, 1979).

[b] There seems to be no practical experience with this approach; see Dwyer (1983) for more details.

and in 6 of the 10 LISREL studies. Two publications systematically tested the reverse hypothesis using hierarchical regressions (Billings & Moos, 1982; Fisher, 1985).

All in all, there were 24 studies that did not test the reverse hypothesis. Of these, 7 were not able to test this hypothesis because of design reasons. Fifteen of the 39 studies (prospective studies excluded) tested the reverse hypothesis, 8 of them rejected it, and 7 found some effects. In 2 studies there was only a weak effect, and in 1 study (Mäkinen & Kinnunen, 1986) the authors did not carry out statistical tests but compared time-lagged stressor-strain and strain-stressor-correlations by appearance. Thus, about one half of the studies that tested reversed causal relationships found at least some indications for a reverse causation effect (Number 2, 9, 15, 26, 31, 32, and 44 in Appendix B). This result should clearly underline the importance of routinely analyzing reversed causal effects in longitudinal studies.

CLPC was obviously not as easy to apply as one would think. Some researchers simply compared cross-lagged correlations without considering significance tests. Many authors attempted to comply with Kenny (1975) to reduce the stationarity problem. Some only analyzed those variables that fulfilled the stationarity condition (e.g., Carayon, 1993; Corrigan et al., 1994), thus leaving the others unanalyzed (or at least unpublished). This leads to the problem that relevant information is simply not taken into consideration when interpreting the results. Others discussed the issue of stationarity, but nevertheless analyzed all available correlations even when substantial differences in cross-sectional correlations appeared. This problem can be solved by using a structural equations approach because this approach allows cross-sectional covariances to change over time.

The LISREL studies typically analyzed reciprocal models (i.e., causal effects of stressors on strain and the reverse). For example, Kohn and Schooler (1982) reported an effect of anxiety on time pressure, and Marcelissen, Winnubst, Buunk, and de Wolff (1988) found—contrary to their hypothesis—a causal effect of affective strain, regular health complaints, worry, and diastolic blood pressure on coworker support. However, there were hardly any reports that made use of the full potential of LISREL by testing a series of models. In particular, those suggestions that we gave above which included systematically testing various synchronous and lagged causal effects in both directions and a series of nested models were not often heeded.

The Treatment of Third Variables

As argued before, the third variable problem is not automatically solved by doing longitudinal research. We later discuss hierarchical regression, CLPC, and structural equations approaches to this problem.

In the traditional hierarchical regression approach, third variables have to be included to control for their effects; this is similar to cross-sectional research. This was done in less than half of all cases (the studies that controlled for third variables were Articles 1^{+}, 2^{+}, 4^{+}, 5^{+}, 6^{+}, 8^{+}, 17^{+}, 18^{+}, 19^{+}, 20^{+}, 33^{*+}, 36^{*+}, 38^{*+}, and 40^{*+} of Appendix A; studies marked with a plus sign controlled for third variables measured only once; studies marked with an asterisk controlled for third variables measured at different time points).

In CLPC designs, the effects of variables that can reasonably be considered to be stable (age, sex, and marital or social status for shorter time lags), and that affect stressors and strain equally at Time 1 and Time 2, can be ruled out by the CLPC technique (rejection of the synchronous common factor model, cf. Kenny, 1975). However, third variables that can change over time, such as social desirability, remain a problem. As already stated, it is also not possible to rule out models that assume stressors and strains to be measures of the same underlying construct, for example of negative affectivity (at least not as long as Time 1 health measures are not partialled out). Therefore, the application of CLPC remains limited in this respect. In comparison, structural equations make it possible to test a variety of third variable models and compare them with various causal models.

In most prospective research, it was not necessary to test the reverse causation hypothesis because, for example, an acute heart disease or mortality cannot itself affect the working conditions. However, third variable explanations are as plausible as in other designs. An example: One's reduced physical efficiency is related both to bad working conditions and to the occurrence of a heart disease. Thus, third variables have to be explicitly considered, and all prospective studies did so.

There were some reports (Dignam & West, 1988; Dormann, Zapf, & Speier, 1995; Marcelissen et al., 1988) that modeled the occasion factors described above. Dormann et al. (1995) added occasion factors that led to a more significant causal effect of social stressors on depression compared with a model without occasion factors. Although occasion factors are often an alternative explanation for true effects in cross-sectional research, it should be noted that the introduction of these occasion factors in longitudinal studies may, in some cases, reveal causal relationships that otherwise would remain undetected (cf. also Marcelissen et al., 1988, who obtained structural coefficients higher than the original correlation when occasion factors were introduced).

Summary of Four Model Studies

Methodological arguments are often quite abstract. To make them more concrete, we conclude this article by summarizing four reports that we find particularly convincing from a methodological point of view. Their results are of particular importance because one can be more confident of their conclusions than of results of studies with a lower degree of methodological rigor.

Leitner (1993) and Frese (1985) were the only authors who reported using observational interviews to measure the stressors. Because much has been said about the study of Frese (1985) already, we do not consider it here again. Through the use of observational measures, Leitner (1993) avoided common method variance between stressors and self-reported health. This makes it less important to consider most third variables. The stressor—additional effort—was analyzed with respect to nine health indicators using CLPC. Significant effects were found for psychosomatic complaints, irritation, depression, eye problems, allergy, and life satisfaction. Thus, six of nine analyses revealed effects in the hypothesized direction. No reversed causation was found: Controlling for age, job tenure, school education, and professional education did not change these findings but led to an insignificant maximum reduction of the lagged correlations of .04. LISREL analyses, although only mentioned but not reported in the Leitner article, came to the same results.

Kohn and Schooler's (1982) LISREL study is well-known and, therefore, can be treated very briefly. They analyzed a full set of third variables and calculated various models of reciprocal causation. Because they only had data from two waves, it was not possible to analyze all relevant models. However, the authors skillfully used structural equations to get the most information out of a two-wave longitudinal design. In their final analysis, a combined distress–well-being indicator (including trustfulness, self-deprecation, idea-conformity, self-confidence, and anxiety) was related to 14 different job conditions. Six of these job conditions could be assessed as a cause of well-being (closeness of supervision, position in hierarchy, dirtiness, hours of work, job protections, and job income); there were three reversed effects (time pressure, heaviness, and "held responsible").

Marcelissen et al. (1988) used a three-wave design to analyze the reciprocal effects between several strain variables and coworker or supervisor support. Their LISREL analyses did not include measurement models, and only synchronous but no lagged effects were analyzed. Third variables were also not included. However, occasion factors were considered and separate analyses for lower and higher occupational levels were conducted. Three out of 16 causal effects from supervisor support on strains were found: regular health complaints (lower occupational level), worry concerning one's job (lower and higher level), but no reversed effects. For coworker support the results were quite the opposite. There were no effects from coworker support on any strain variable, but 5 of 16 reversed causal coefficients reached significance: affective strain (low and high level), worry concerning one's job (low), diastolic blood pressure (low), and regular health complaints (high occupational level).

Schonfeld (1992) was the only one who controlled for preemployment symptoms. His LISREL analyses tested reciprocal models with different time lags but did not incorporate an occasion factor or other third variables. School environment stressors were synchronously related to depressive symptoms but there was no significant reversed effect.

In summary, 16 of the 50 tested causal hypotheses and 8 of 50 tested reversed causation models of these four studies were supported. On average, the causal effect was .12. While this coefficient is not high, it is in line with our arguments presented in the beginning of this article.

However, one can point even in this set of excellent articles to methodological considerations that are necessary but missing: testing lagged and synchronous effects, controlling for third variables, and testing reversed and reciprocal causation (the only exception was Kohn & Schooler, 1982, who did all of this).

Conclusions

In summary, the lack of common standard procedures to analyze longitudinal data is reflected in the empirical articles reviewed for this article. The concerns of Williams and Podsakoff (1989) that many researchers unjustifiably feel that applying a longitudinal design automatically solves many problems inherent in cross-sectional approaches can be supported by our review.

The following recommendations can be made with regard to methodological issues of longitudinal stress research:

1. All variables should be measured at all time points using the same measurement method for the respective variables. If this is not done, one cannot examine reverse or reciprocal causation hypotheses that consider strain Time 2 as the dependent variable. Additionally, certain third variable hypotheses such as occasion factor hypotheses require all variables measured at each time.

2. As in cross-sectional studies, researchers should carefully consider third variables as potential confounders of the stressor-strain relationship and include them into their design. As mentioned above, structural equation models possess the capability to take unmeasured third variables into account. However, such models usually require that certain constraints be put on the models (e.g., stationarity restrictions). Because there is still little practical experience with complex common factor models and because estimation problems occur even in simpler analyses, it is better to explicitly measure third variables whenever possible.

3. The time lag should be thoroughly planned (Kessler, 1987). Simulation studies show that a time lag that is too long is less of a problem than one that is too short (cf. Dwyer, 1983). Time lags that are too short may lead to the conclusion that no causal effects exist, whereas a time lag that is too long solely leads to an underestimation of the true causal impact. It would be best to conduct a multiwave design with equal time lags. If short-term effects are adequate then it should be possible to replicate them as synchronous paths within the same model. If there are long-term effects, the appropriate time interval can be assessed by comparisons of different time lag models (e.g., Dormann et al., 1995).

4. Assumptions about the time course (Frese & Zapf, 1988) of the variables under study should be made. For example, if strong adaptation processes are expected, then it should be wise to study participants beginning their jobs, that is, before they have the opportunity to adapt to the working conditions (cf. Schonfeld, 1992).

5. A linear structural equations approach is recommended to analyze the data (cf. Dwyer, 1983; James & James, 1989; Link & Shrout, 1992; Williams & Podsakoff, 1989).

6. Measurement models should be included in the models. Errors in measurement that may attenuate relationships across variables can be accounted for by the introduction of measurement models.

7. Multiple competing models should be tested (James, Mulaik, & Brett, 1982). The rationale for this recommendation is that models that survived competing tests with a series of alternative models can be trusted more. The following recommendations are useful: (a) To reduce the complexity, one should test measurement models separately before testing the structural model (two-step approach, cf. Anderson & Gerbing, 1988). This is especially true for longitudinal models because several measurement models that express different "behaviors" of construct-indicator relationships over time can be tested. This task becomes too complicated if done together with procedures to test causality assumptions. Although this approach has been criticized (e.g., Fornell & Yi, 1992) because measurement models can change with variations in the structural part of the equation system, our practical experiences suggest that in most cases these changes are minor. (b) Causal effects including reversed effects should be systematically introduced into the models. (c) Effects of third variables should be systematically tested. (d) Occasion factors should be tested. Occasion factors are easy to model, they are always identified, and they can be combined with several kinds of causal models. This allows to test whether or not a certain causal model holds in spite of this special kind of third variables.

When we started reviewing the literature on longitudinal studies it was our hope to find a clear trend to methodologically sounder studies in recent years. This trend is not as clear as we expected. We hope that this article contributes to more systematic longitudinal designs of organizational stress studies that make it possible to test reverse causal hypotheses and a series of third variable explanations. If we have then a better understanding of causal stressor-strain relationships, then time and effort to write this article was well invested.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411–423.

Bamberg, E. (1991). Arbeit, Freizeit und Familie [Work, leisure time and family]. In S. Greif, E. Bamberg, & N. Semmer (Eds.), *Psychischer Streß am Arbeitsplatz* (pp. 201–221). Göttingen, Germany: Hogrefe.

Bamberg, E. (1992). Stressoren in der Erwerbsarbeit und in der Freizeit: Zusammenhänge mit psychischen Befindensbeeinträchtigungen [Stressors at work and in leisure time: Their relationship with psychological dysfunctioning]. *Zeitschrift für Arbeits- und Organisationspsychologie, 36,* 84–91.

Beck, A. T. (1972). *Depression: Causes and treatment.* Philadelphia: University of Pennsylvania Press.

Begley, T. M., & Czajka, J. M. (1993). Panel analysis of the moderating effects of commitment on job satisfaction, intent to quit, and health following organizational change. *Journal of Applied Psychology, 78,* 552–556.

Billings, A. G., & Moos, R. H. (1982). Social support and functioning among community and clinical groups: A panel model. *Journal of Behavioral Medicine, 5,* 295–311.

Birdi, K., Warr, P. B., & Oswald, A. (1995). Age differences in three components of employee well-being. *Applied Psychology: An International Review, 44,* 345–373.

Blalock, H. M. (1961). *Causal inferences in non-experimental research.* Chapel Hill, NC: University of North Carolina Press.

Blalock, R. (1962). Four variable causal models and partial correlations. *American Journal of Sociology, 68,* 182–194.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Brannick, M. T. (1995). Critical comments on applying covariance structure modeling. *Journal of Organizational Behaviour, 16,* 201–213.

Brief, A. P., Burke, M. J., George, J. M., Robinson, B., & Webster, J. (1988). Should negative affectivity remain an unmeasured variable in the study of job stress? *Journal of Applied Psychology, 73,* 193–198.

Burke, M. J., Brief, A. P., & George, J. M. (1993). The role of negative affectivity in understanding relations between self-reports of stressors and strains: A comment on the applied psychology literature. *Journal of Applied Psychology, 78,* 402–412.

Burke, R. J., & Greenglass, E. (1995). A longitudinal study of psychological burnout in teachers. *Human Relations, 48,* 187–202.

Campbell, D. T. (1963). From description to experimentation: Interpreting trends in quasi experimentation. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 212–242). Madison, WI: University of Wisconsin Press.

Campbell, D. T., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Caplan, R. D., Cobb, S., French, J. R. P., Harrison, R. van, & Pinneau, S. R. (1975). *Job demands and worker health.* Washington, DC: National Institute for Occupational Safety and Health.

Caplan, R. D., & Jones, K. W. (1975). Effects of work load, role ambiguity, and Type A personality on anxiety, depression, and heart rate. *Journal of Applied Psychology, 60,* 713–719.

Carayon, P. (1993). A longitudinal test of Karasek's job strain model among office workers. *Work and Stress, 7,* 299–314.

Cohen, S., & Edwards, J. R. (1989). Personality characteristics as moderators of the relationship between stress and disorder. In R. W. J. Neufeld (Ed.), *Advances in the investigation of psychological stress* (pp. 235–283). New York: Wiley.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357.

Contrada, R. J., & Krantz, D. S. (1987). Measurement bias in health psychology research designs. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 57–78). New York: Wiley.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis for field settings.* Boston: Houghton Mifflin.

Corrigan, P. W., Holmes, E. P., Luchins, D., Buican, B., Basit, A., & Parks, J. J. (1994). Staff burnout in a psychiatric hospital: A cross-lagged panel design. *Journal of Organizational Behaviour, 15,* 65–74.

Daniels, K., & Guppy, A. (1994). Occupational stress, social support, job control, and psychological well-being. *Human Relations, 47,* 1523–1544.

Depue, R. A., & Monroe, S. M. (1986). Conceptualization and measurement of human disorder in life stress research: The problem of chronic disturbance. *Psychological Bulletin, 99,* 36–51.

Dignam, J. T., & West, S. G. (1988). Social support in the workplace: Tests of six theoretical models. *American Journal of Community Psychology, 16,* 701–724.

Dohrenwend, B. S., & Dohrenwend, B. P. (Eds.). (1974). *Stressful life events: Their nature and effects.* New York: Wiley.

Dormann, C., Zapf, D., & Speier, C. (1995). *Social support, social stressors at work and depression: Testing for main and moderating effects with structural equations in a three-wave longitudinal study.* Manuscript submitted for publication.

Dwyer, J. E. (1983). *Statistical models for the social and behavioral sciences.* New York: Oxford University Press.

Feather, N. T. (1989). *The psychological impact of unemployment.* New York: Springer.

Feldman, J. (1975). Considerations in the use of causal-correlational techniques in applied psychology. *Journal of Applied Psychology, 60,* 663–670.

Fisher, C. D. (1985). Social support and adjustment to work: A longitudinal study. *Journal of Management, 11,* 39–53.

Fornell, C., & Yi, Y. (1992). Assumptions of the two-step approach to latent variable modeling. *Sociological Methods & Research, 20,* 291–320.

Frankenhaeuser, M. (1991). The psychophysiology of sex differences as related to occupational status. In M. Frankenhaeuser, U. Lundberg, & M. Chesney (Eds.), *Women, work, and health* (pp. 39–61). New York: Plenum Press.

Frese, M. (1982). Occupational socialization and psychological development: An underemphasized research perspective in industrial psychology. *Journal of Occupational Psychology, 55,* 209–224.

Frese, M. (1985). Stress at work and psychosomatic complaints: A causal interpretation. *Journal of Applied Psychology, 70,* 314–328.

Frese, M. (1989). Theoretical models of control and health. In S. L. Sauter, J. J. Hurrel, & C. L. Cooper (Eds.), *Job control and worker health* (pp. 108–128). New York: Wiley.

Frese, M. (1995). Social support as a moderator of the relationship between work stressors and psychological dysfunctioning: A longitudinal study with objective measures. Manuscript submitted for publication.

Frese, M., & Okonek, K. (1984). Reasons to leave shiftwork and psychological and psychosomatic complaints of former shiftworkers. *Journal of Applied Psychology, 69,* 509–514.

Frese, M., & Semmer, N. (1991). Streßfolgen in Abhängigkeit von Moderatorvariablen: Der Einfluß von Kontrolle und Sozialer Unterstützung [Stress consequences depending on moderator variables: The impact of control and social support]. In S. Greif, E. Bamberg, & N. Semmer (Eds.), *Psychischer Streß am Arbeitsplatz* (pp. 135–153). Göttingen, Germany: Hogrefe.

Frese, M., & Zapf, D. (1988). Methodological issues in the study of work stress: Objective vs. subjective measurement of work stress and the question of longitudinal studies. In C. L. Cooper & R. Payne (Eds.), *Causes, coping, and consequences of stress at work* (pp. 375–411). Chichester, England: Wiley.

Fusilier, M. R., Ganster, D. C., & Mayes, B. T. (1987). Effects of social support, role stress, and locus of control on health. *Journal of Management, 13,* 517–528.

Gutek, B. A., Repetti, R. L., & Silver, D. L. (1988). Nonwork roles and stress at work. In C. L. Cooper & R. Payne (Eds.), *Causes, coping, and consequences of stress at work* (pp. 141–174). Chichester, England: Wiley.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81,* 945–960.

James, L. R., & James, L. A. (1989). Causal modeling in organizational research. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (vol. 4, pp. 371–404). New York: Wiley.

James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Assumptions, models, and data.* Beverly Hills, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd ed.). Chicago: SPSS.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8.* Chicago: Scientific Software.

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). Chichester, England: Wiley.

Kasl, S. V. (1983). Pursuing the link between stressful life experiences and disease: A time of reappraisal. In C. L. Cooper (Ed.), *Stress research* (pp. 79–102). New York: Wiley.

Kasl, S. V. (1986). Stress and disease in the workplace: A methodological commentary on the accumulated evidence. In M. F. Cataldo & T. J. Coates (Eds.), *Health and industry: A behavioral medicine perspective* (pp. 54–85). 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. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin, 82,* 887–903.

Kenny, D. A. (1979). *Correlation and causality.* New York: Wiley.

Kessler, R. C. (1987). The interplay of research design strategies and data analysis procedures in evaluating the effects of stress on health. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 113–140). New York: Wiley.

Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis: Models of quantitative change.* New York: Academic Press.

Kohn, M. L., & Schooler, C. (1982). Job conditions and personality: A longitudinal assessment of their reciprocal effects. *American Journal of Sociology, 87,* 1257–1286.

Kohn, M. L., & Schooler, C. (1983). *Work and personality. An inquiry into the impact of social stratification.* Norwood, NJ: Ablex.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

Leitner, K. (1993). Auswirkungen von Arbeitsbedingungen auf die psychosociale Gesundheit [The effects of working conditions on psychosocial health]. *Zeitschrift für Arbeitswissenschaft, 47,* 98–107.

Leventhal, H., & Tomarken, A. (1987). Stress and illness: Perspectives from health psychology. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 3–26). New York: Wiley.

Leymann, H., & Gustafsson, A. (in press). Mobbing and the development of post-traumatic stress disorders. *European Journal of Work and Organizational Psychology.*

Link, B. G., & Shrout, P. E. (1992). Spurious associations in longitudinal research. *Research In Community And Mental Health, 7,* 301–321.

Lipowski, Z. J. (1975). Psychiatry of somatic diseases: Epidemiology, pathogenesis, classification. *Comprehensive Psychiatry, 16,* 105–124.

Locascio, J. J. (1982). The cross-lagged correlation technique: Reconsideration in terms of exploratory utility, assumption specification, and robustness. *Educational and Psychological Measurement, 42,* 1023–1036.

Mäkinen, R., & Kinnunen, U. (1986). Teacher stress over a school year. *Scandinavian Journal of Educational Research, 30,* 55–70.

Marcelissen, F. H. G., Winnubst, J. A. M., Buunk, B., & de Wolff, C. J. (1988). Social support and occupational stress: A causal analysis. Special Issue: Stress and coping in relation to health and disease. *Social Science and Medicine, 26,* 365–373.

Marmot, M. G., & Madge, N. (1987). An epidemiological perspective on stress and health. In S. V. Kasl & C. L. Cooper (Eds.), *Stress and health: Issues in research methodology* (pp. 3–26). New York: Wiley.

Nelson, D. L., & Sutton, C. (1990). Chronic work stress and coping: A longitudinal study and suggested new directions. *Academy of Management Journal, 33,* 859–869.

O'Brien, G. E. (1986). *Psychology of work and unemployment.* New York: Wiley.

Orpen, C. (1982). Type A personality as a moderator for the effects of role conflict, role ambiguity, and role overload on individual strain. *Journal of Human Stress, 8,* 8–14.

Parkes, K. R. (1990). Coping, negative affectivity, and the work environment: Additive and interactive predictors of mental health. *Journal of Applied Psychology, 75,* 399–409.

Pelz, D. C., & Andrews, F. M. (1964). Detecting causal priorities in panel study data. *American Sociological Review, 29,* 836–848.

Poulin, J., & Walter, C. (1993). Social worker burnout: A longitudinal study. *Social Work Research and Abstracts, 29*(4), 5–11.

Rabkin, J. G., & Struening, E. L. (1976). Life events, stress, and illness. *Science, 194,* 1013–1020.

Richter, P. G., Schirmer, F., & Dettmar, P. (1989). Zum Zusammenhang zwischen aktuellen und langfristigen Beanspruchungsfolgen bei geistiger Arbeit [On the relationship between actual and long-term stress consequences in mental work]. *Psychologie für die Praxis, 7,* 135–150.

Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin, 88,* 245–258.

Schonfeld, I. S. (1992). A longitudinal study of occupational stressors and depressive symptoms in first-year female teachers. *Teaching and Teacher Education, 8,* 151–158.

Seeber, A., & Iregren, A. (Eds.). (1992). Behavioural effects of contaminated air: Applying psychology in neurotoxicology [Special issue]. *Applied Psychology: An International Review, 41(3).*

Semmer, N. K. (in press). Individual differences, work stress, and health. In M. J. Schabracq, J. A. Winnubst, & C. L. Cooper (Eds.), *Handbook of work and health psychology.* Chichester, England: Wiley.

Semmer, N. K., Zapf, D., & Greif, S. (in press). "Shared job strain": A new approach for assessing the validity of job stress measurements. *Journal of Occupational and Organizational Psychology.*

Spector, P. E. (1992). A consideration of the validity and meaning of self-report measures of job conditions. In C. L. Cooper & I. T. Roberston (Eds.), *International review of industrial and organizational psychology 1992* (vol. 7, pp. 123–151). Chichester, England: Wiley.

Spector, P. E., & Brannick, M. T. (1995). The nature and effects of method variance in organizational research. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (vol. 10, pp. 249–274). New York: Wiley.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.

Steptoe, A., & Wardle, J. (1995). Health attitudes and behaviour: An international perspective. In K. Pawlik (Ed.), *Bericht über den 39. Kongress der Deutschen Gesellschaft für Psychologie in Hamburg 1994* (pp. 156–172). Göttingen, Germany: Hogrefe.

Syme, S. L., & Berkman, L. F. (1976). Social class, susceptibility, and sickness. *The American Journal of Epidemiology, 104,* 1–8.

Tetrick, L. E., & LaRocco, J. M. (1987). Understanding, prediction, and control as moderators of the relationship between perceived stress, satisfaction, and psychological well-being. *Journal of Applied Psychology, 72,* 538–543.

Theorell, T., Leymann, H., Jodko, M., Konarski, K., & Norbeck, H. E. (1994). "Person under train" incidents from the subway driver's point of view: A prospective 1-year follow-up study. The design, and medical and psychiatric data. Special Issue: Suicide on railways. *Social Science and Medicine, 38,* 471–475.

Theorell, T., Orth-Gomér, K., & Eneroth, P. (1990). Slow-reacting immunoglobulin in relation to social support and changes in job strain: A preliminary note. *Psychosomatic Medicine, 52,* 511–516.

Vossel, G. (1987). Stress conceptions in live event research: Towards a person-centered perspective. *European Journal of Personality, 1,* 123–140.

Waldron, I., Herold, J., Dunn, D., & Staum, R. (1982). Reciprocal effects of health and labor force participation among woman: Evidence from two longitudinal studies. *Journal of Occupational Medicine, 24,* 295–310.

Wall, T. D., Jackson, P. R., Mullarkey, S., & Parker, S. K. (in press). The demand-control model of job strain: A more specific test. *Journal of Occupational and Organizational Psychology.*

Warr, P. B. (1987). *Work, unemployment, and mental health.* Oxford, England: Oxford University Press.

Warr, P. B. (1992). Age and occupational well-being. *Psychology and Aging, 7,* 37–45.

Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465–490.

Williams, L. J., & Podsakoff, P. M. (1989). Longitudinal field methods for studying reciprocal relationships in organizational behavior research: Toward improved causal analysis. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (vol. 11, pp. 247–292). Greenwich, CT: JAI Press.

Zapf, D. (1989). *Selbst- und Fremdbeobachtung in der psychologischen Arbeitsanalyse. Methodische Probleme bei der Erfassung von Stress am Arbeitsplatz* [Self and expert observation in psychological job analysis. Methodological problems in the measurement of stressors at work]. Göttingen, Germany: Hogrefe.

Zapf, D., & Frese, M. (1991). Soziale Stressoren am Arbeitsplatz und psychische Gesundheit [Social stressors at work and psychological health]. In S. Greif, E. Bamberg, & N. Semmer (Eds.), *Psychischer Streß am Arbeitsplatz* (pp. 168–184). Göttingen, Germany: Hogrefe.

(Appendixes follow on next page)