Scand J Work Environ Health 2003;29(1):1–4

Challenges in longitudinal designs in occupational health psychology

Many researchers in occupational health psychology assume that longitudinal research designs permit stronger conclusions concerning the causal relations among the study concepts than cross-sectional designs do. As the temporal order of the variables can be determined unambiguously, longitudinal designs are presumed to offer good opportunities to add further to our understanding of the causal processes generating the phenomena of interest. Consequently, the last three decades have witnessed an increasing popularity for longitudinal designs in both psychology and medicine (1, 2). Not less than 17 (35%) of the 40 original articles published in the 2001 volume of the *Scandinavian Journal of Work, Environment and Health* utilized a longitudinal design.

It is good to see that so many researchers acknowledge the disadvantages of cross-sectional research designs in examining causal processes: correlation is no causation. However, we should bear in mind that longitudinal study designs per se are no guarantee for drawing valid causal inferences and that, in addition to strong points, longitudinal study designs have drawbacks as well. Some of these are widely recognized, such as the issues of selective attrition and testing effects. Selective attrition may result in a restriction of the range of the variables of interest (eg, only healthy workers remain), meaning that the strength of associations among these is underestimated. Testing effects occur when respondents lose interest in participating in a study when they have to answer the same questions time and again or when respondents interpret questions differently because they are becoming more sensitive as regards their work circumstances and health (2).

On a more general level we believe that there are three common, important and less widely recognized misunderstandings with respect to longitudinal study designs in occupational health psychology. The first of these is that *we can prove causality by using longitudinal study designs.* The extent to which causal inferences can be made depends on the following four conditions: temporal ordering of the focal variables, the strength of the statistical association between them, theoretical plausibility of the presumed causal relationship, and exclusion of plausible rival hypotheses for this relationship (2, unpublished manuscript by de Lange et al). While the first three conditions are relatively easy to satisfy with a longitudinal design, it is impossible to exclude the possibility that particular associations are due to variables that were not measured in the study design. Thus we can never prove causal relationships; the best we can do is argue that it is plausible that certain statistical associations can be understood in causal terms.

The second common misunderstanding in longitudinal research is that *two observations are adequate for studying intraindividual processes.* Clearly, two observations provide information about change over time, but this information is usually insufficient for a thorough understanding of the process responsible for these changes. Here, we face the following two questions: (i) how do we choose the right time interval, and (ii) how do we know what happens within the time interval. First, it has often been noted that the measurement interval between the phases of a study should correspond with the time that a causal variable X needs to affect an outcome variable Y (the underlying causal interval). Although this issue is not new, it is our impression that, in occupational health psychology, the interval between the phases of a study is often chosen on pragmatic grounds, rather than on the basis of careful consideration of the possible underlying causal lags (unpublished manuscript by de Lange et al). The length of this causal lag is often unknown (2), but this lack of knowledge should not keep researchers from considering and addressing this issue. Researchers should indicate why they feel that the lags employed in their

study are appropriate for investigating the phenomenon under study, on the basis of theoretical arguments and empirical evidence. Such empirical evidence stems from a recent multiphase panel study by Dormann & Zapf (3) on the effects of conflict with supervisors and colleagues on depression. They showed that the strongest effects were found for a 2-year lag, rather than for a shorter or longer interval. Although this does not imply that a 2-year time lag is always optimal (that depends on the variables under study), Dormann & Zapf's research clearly demonstrates that multiphase studies are more useful for examining causal processes in occupational health psychology than the standard two-phase panel design.

Regarding our second question (what happens within the time interval), two observations per subject may permit an estimation of the amount of across-time change, but not of the individual growth curve. If individual scores are available for time 1 and time 2 of the study, we know very little about the scores in between these time points. Linear interpolation may result in a reasonable estimate of the individual score between time 1 and time 2, but it is just as likely that the phenomenon of interest develops in a non-linear fashion or that participants have changed "back and forth" during the observed interval. Two-phase panel studies thus provide little insight into the across-time development of the process of interest. Frese & Zapf (4) discussed five different models for the impact of stressors on outcome variables; none of these models can be distinguished from the competing models using a two-phase panel design. Therefore, in agreement with statistician David Rogosa, we conclude that "Two waves of data are better than one, but maybe not much better [p 174]" (5).

A third common misunderstanding is that *scientific progress first and foremost stems from applying methodological designs, standard statistical procedures, and variation on a research theme, rather than from clever thinking*. It seems that many researchers put the emphasis on a particular methodological context, on examining already existing questions, and on refining already existing findings from different angles, instead of taking a fresh look and generating and testing novel ideas. Our impression is that many longitudinal studies are primarily designed to provide longitudinal replications of long-standing insights. Although such work is often useful and important, as it may yield more insight into the presumed causal process, we feel that current longitudinal research does not fully realize the potential of the design in generating and testing *non*standard ideas. For example, consider the relationship between work characteristics and health. We do know that high job demands, low control, and low support have adverse effects on worker health, not only cross-sectionally but longitudinally as well (6, 7). It is not all that interesting to replicate earlier findings for yet another health-related outcome variable or for somewhat different explanatory variables (eg, physical instead of psychological job demands). Rather, we would like longitudinal studies to provide answers to innovative questions, that is, to focus on *new* insights that may be obtained using such designs. Three possibly interesting and interrelated notions that can be tested using such designs concern (i) the effects of prolonged exposure to particular job-related conditions, (ii) the issue of subgroup analysis, and (iii) the issue of reciprocal effects.

As regards the first point (i), longitudinal research in occupational health psychology typically aims to explain the time-2 score on a particular variable from the time-1 scores on that variable and other variables, without controlling the amount of exposure to the explanatory variables *before* time 1 (ie, the history preceding the time-1 scores). There seems little interest in work, stress, and health research in the effects of prolonged exposure to particular (job-related) conditions, despite the prominent role that chronic or day-to-day exposure plays in most psychosocial theories (8). Whereas in occupational medicine it is common to include an estimate of the cumulative amount of exposure to X in examining the effects of X on Y (eg, the number of exposure years), as yet such practices are rarely applied in psychosocial longitudinal research.

The second issue (ii) relates to studying the total group of employees versus theoretically specified subgroup analysis. Employed workers mostly constitute a relatively healthy subset of the total population.

While most employees do not face high levels of adverse work conditions (peak loads), some may face day-to-day work conditions that are tolerable for one day but not for a lifetime. It is probably the long term, everyday exposure, combined with insufficient coping possibilities and insufficient recovery, that causes disease. This assumption implies that neither the average work score nor the average health score of a large group of (intrinsically healthy) employees will change much over a particular time lag (eg, 1 year). Indeed, time-1 and time-2 overall sample scores may well be interpreted as test-retest scores in such a design, the test-retest correlation being dependent on the length of the time interval. However, this placid impression may conceal that, for at least *some* participants, *some* change may have occurred. At the aggregate level such change may remain undetected; changes occurring for one participant may compensate changes occurring for another. It would thus seem particularly rewarding to focus on specific *subgroups* of the total sample. Potentially interesting subgroups include participants whose psychosocial work environment changed and participants reporting a change in the dependent variable (eg, a change in health status) or independent variables (eg, a change from a low-demand job to a high-demand job or vice versa). Other interesting subgroups would include participants reporting long-term exposure to unfavorable job conditions or participants reporting high levels of strain. It would seem likely that closer examination of these (and other) groups would yield more insight into the processes that underlie change in the outcome variables.

The third issue (iii) is the current lack of interest in reciprocal effects. Much research in this area examines the effects of job characteristics (as explanatory variables) on health (the outcome variable). However, it is often conceivable that health status also affects the (perception of the) characteristics of one's job. Workers may be unable to retain good job conditions due to poor health [Frese & Zapf's drift-hypothesis (4)], or they may perceive their jobs differently (eg, people may judge their jobs to be more taxing because their resources have been depleted). There is some limited evidence for such effects (9), but more evidence for the possible effects of health on work characteristics is certainly needed.

What are the design requirements for longitudinal studies? It will be clear that we believe that longitudinal designs are no panacea for all weaknesses of the cross-sectional design. To be sure, we *do not* question the idea that longitudinal designs allow for stronger conclusions concerning possible causal relationships among variables than cross-sectional designs do. We do think, however, that much current longitudinal research in occupational health psychology does *not fully* profit from the opportunities implicit in this design—opportunities to extend and enhance our understanding of the development of causal processes, to examine nonstandard research questions, and thus to conduct *more innovative* research on the interface between work and worker health. In order to consummate this potential, longitudinal studies should meet several requirements regarding both focus and design. Most importantly, researchers should carefully consider the process of interest before actually designing and carrying out their studies. How long does it take for the causal variable to affect the outcome variable? How likely is it that the designated "outcome" variable affects the "causal" variable as well? Does the causal variable affect the study outcomes in a linear or nonlinear fashion? Can we reasonably expect that the amount of across-time change in the sample is large enough to be able to explain part of this change, or will there be hardly any change at all? Will change—*if* it occurs—be concentrated in particular subgroups, and, if so, in *which* subgroups? Does it make sense for us to invest scarce resources in increasing the number of participants in our study or should we increase the number of measurements of the study? And so on.

Clearly, the answers to these questions have implications for the design of the study. It is unlikely that simple pre-fab designs (eg, the two-phase design in which an essentially cross-sectional study is extended with a repeated measure of the outcome variable) will be able to provide satisfying answers to the preceding questions. Rather, we expect that study designs will become more varied and complicated. Given a particular minimum of participants in order to reach sufficient statistical power, studies will include more waves rather than more participants in order to capture more fully the developmental

aspects of the process of interest. Simple two-phase designs are insufficient to tell us much about the rate of change or about the shape of presumed causal relationships; thus multiphase designs should be preferred to these simple designs. Furthermore, studies will more often employ a full panel design (ie, the causal and effect variables should be measured in all waves of the study) rather than an incomplete panel design in order to allow for the examination of reciprocal effects. Thus we expect that researchers will collect *more data* on *fewer subjects* than is currently common in occupational health psychology. This situation may also imply that researchers must invest resources in acquiring new skills in order to analyze such data properly. Structural equation modeling may be needed to model the reciprocal relationships between pairs of variables across time, while multilevel modeling can be useful in analyzing the shape of change relationships. If researchers in occupational health psychology are to enhance their understanding of causal processes, they will have to invest in better and more extensive longitudinal data collection, in methods for analyzing such data, and in the ideas underlying their studies. One thing seems certain: occupational health research will become even more challenging in the near future!

The present issue of the *Scandinavian Journal of Work, Environment & Health* includes a longitudinal paper on the causal relationship between unemployment and health, authored by F Curtis Breslin & Cam Mustard. Not only do we believe this to be an excellent paper, but we also believe it presents a fine example of a study that goes beyond the longitudinal replication of long-standing insights, and it thus enhances our understanding of the effects of unemployment on health and vice versa.

References

1. Austin JT, Scherbaum CA, Mahlman RA. History of research methods in industrial and organizational psychology: measurement, design, analysis. In: Rogelberg SG, editor. Handbook of research methods in industrial and organizational psychology. Malden: Blackwell; 2002. p 3–33.
2. Taris TW. A primer in longitudinal data analysis. London: Sage; 2000.
3. Dormann C, Zapf D. Social stressors at work, irritation, and depressive symptoms: accounting for unmeasured third variables in a multi-wave study. J Occup Organ Psychol 2002;75:33–58.
4. Frese M, Zapf D. Methodological issues in the study of work stress: objective vs subjective measurement of work stress and the question of longitudinal studies. In: Cooper CL, Payne R, editors. Causes, coping and consequences of stress at work. Chichester: Wiley; 1988. p 375–411.
5. Rogosa D. Myths about longitudinal research. In Schaie KW, Campbell RT, Meredith E, Rawlings SC, editors. Methodological issues in aging research. New York (NY): Springer; 1988. p 171–209.
6. Karasek RA, Theorell T. Healthy work: stress, productivity, and the reconstruction of working life. New York (NY): Basic Books; 1990.
7. de Lange AH, Taris TW, Kompier MAJ, Houtman ILD, Bongers PM. Effects of stable and changing demands-control histories on worker health. Scand J Work Environ Health 2002;28:94–108.
8. Kompier M. Job design and well-being. In: Schabracq MJ, Winnubst JAM, Cooper CL, editors. The handbook of work and health psychology. Chichester: Wiley; 2002. p 429–54.
9. De Jonge J, Dormann C, Janssen PPM, Dollard MF, Landeweerd JA, Nijhuis FJN. Testing reciprocal relationships between job characteristics and psychological well-being: a cross-lagged structural equation analysis. J Occup Organ Psychol 2001;74:29–45.

Toon W Taris, PhD, Michiel Kompier, PhD
Department of Work and Organizational Psychology
University of Nijmegen
PO Box 9104
NL–6500 Nijmegen, The Netherlands
[E-mail: t.taris@psych.kun.nl]