Work & Stress
An International Journal of Work, Health & Organisations

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/twst20

Strategies addressing the limitations of cross-sectional designs in occupational health psychology: What they are good for (and what not)

Toon W. Taris, Stacey R. Kessler & E. Kevin Kelloway

WORK & STRESS
2021, VOL. 35, NO. 1, 1–5
https://doi.org/10.1080/02678373.2021.1888561

Routledge
Taylor & Francis Group

EDITORIAL

Check for updates

Strategies addressing the limitations of cross-sectional designs in occupational health psychology: What they are good for (and what not)

Like many researchers in other areas of the social and behavioural sciences, scholars in occupational health psychology (OHP) often rely on cross-sectional survey designs when addressing the issues of interest in their field. Such designs usually take the form of online questionnaires in which the concepts to be examined – "antecedents" such as job characteristics, "consequences" such as burnout, and all sorts of "mediators" and "moderators" – are measured at a single point in time. This approach has several advantages compared to more elaborate designs, including its simplicity, cost-effectiveness, short data collection period, and participant burden. Overall, the cross-sectional design is well-suited for testing assumptions about the relationships of interest and provides a clear impression of the state of affairs in an organisation or among a group of workers at a given point in time.

However, the disadvantages of the cross-sectional approach are well-known. Chief among these is its inability to separate between a presumed cause and its possible effect: two concepts may well correlate significantly, but that does not mean that one *causes* the other. At the very least one needs to show that the "cause" precedes its "outcome" in time. In the past decades, this notion has led to the broad acceptance and implementation of several alternative approaches to cross-sectional designs that (are believed to) address some of its limitations. Below we discuss two such designs in which the same participants provide data at multiple time points. We argue that compared to the standard cross-sectional design, such designs are advantageous for improving the quality of research in OHP. However, these advantages do not necessarily coincide with the advantages researchers *believe* these designs have for assessing causality. Moreover, even if these approaches *could* address the causality issue, researchers do not always take full advantage of these possibilities.

What we talk about when we talk about longitudinal research

As yet there is no consensus about the defining features of a longitudinal study. A narrow definition was proposed by Taris (2000), stating that longitudinal data "are collected for the same set of research units … for … two or more occasions, in principle allowing for intra-individual comparison across time" (p. 1–2). In a similar vein, Ployhart and Vandenberg (2010) proposed that longitudinal research involves "the study of change and [contains] at minimum three [sic] repeated observations … on at least one of the substantive constructs of interest" (p. 97). These definitions thus focus on the study of intra-individual change, meaning that at least one of the study concepts should be measured at least twice. Such designs temporally separate a presumed outcome from its possible cause, and usually also allow for testing whether across-time change in this outcome is predicted by a presumed antecedent. Although this is not conclusive evidence for a causal relationship, showing that an antecedent precedes across-time change in an outcome certainly helps in arguing that this association can be interpreted causally.

However, more recently researchers have argued that this focus on change is overly restrictive, stating that longitudinal research "is simply research where data are collected over a … span of time" (Wang et al., 2017, p. 3). Thus, in this view longitudinal data may or may not involve *repeated* measurements of a specific concept, as long as the study consists of *temporally separated* parts. Such a separation may serve various goals, but studying intra-individual change and causality are not necessarily among these.

Although there is something to be said for both narrow and broad definitions of longitudinal research, the presence of such varying ideas about the nature and goals of longitudinal research does little to improve the clarity of this concept. Indeed, since the term *longitudinal research* may cover very different designs with correspondingly different advantages and purposes, readers may be confused about the strength of the conclusions and implications of a particular study. Is a particular "longitudinal" study indeed superior to standard cross-sectional designs in terms of its ability to show causal associations, or does it only *seem* to be that way due to the terminology used? The risk of misunderstandings is especially evident in two currently popular research designs: the time-separated study and the diary study.

Are time-separated studies useful in demonstrating causality? Participants in this type of study (that is sometimes also referred to as a *time-lagged* study) partake in at least two waves of data collection, but variables are measured only once. It is common for variables that are considered "antecedents" to be measured at an earlier wave than their presumed "outcomes," meaning that the criterion that cause must precede to show causality is satisfied, at least at first sight. However, since no variables are measured more than once, intra-individual change as implied in the narrow definition of longitudinal research cannot be established, meaning that there is no change in the presumed outcome that can be related to the "antecedent," which would have been a strong indicator of a causal relationship. Consequently, in terms of its ability to establish causality, this design is no better than a cross-sectional design. To be sure, time-separated designs *do* have advantages compared to single source, cross-sectional designs, but these mainly relate to reducing the potential risk of common method bias by separating the measurement of antecedents and outcomes (Podsakoff et al., 2003). Even this advantage relies on an assumption that common method variance is transient. However, at best, such designs address the self-report issue that is often present in survey studies, not the causality issue.

Are diary studies useful in demonstrating causality? It depends … Unlike time-separated designs, diary studies – in which the same participants repeatedly provide information on their day-to-day activities, attitudes, et cetera – usually comply with the narrow definition of longitudinal research. That is, such studies can yield information about day-to-day change in the concepts of interest and researchers could thus examine intra-individual change. Examples include: how mood at the end of day t spills over to the morning of day t+1, whether fatigue increases across the days of a work week, whether and how events happening in the beginning of the work week affect well-being at the end of that week, and whether such effects depend on personality characteristics like coping or support from one's spouse (Bolger et al., 2003).

Yet, in practice studies relying on diary designs are often unable to draw causal inferences. That is, in diary studies researchers often collect multiple waves of data concerning the same concepts for a relatively small number of participants (typically ranging from fifty to as many as two hundred participants, cf. Gabriel et al., 2019). Whereas such data sets could be used to study intra-individual change, in practice researchers often do not model across-time change. Rather, the data obtained for each day of the diary study are usually pooled and analysed as if they were separate observations, using multi-level analysis to solve the issue that the data are not statistically independent (as daily observations are nested within participants). In this vein, a five-day diary study with fifty participants is basically treated as a cross-sectional sample of 250 observations, discarding the information about intra-individual change from one day to the other (and, thus, about causality) that was present in the initial data set.

Thus, while the diary design is longitudinal by nature, often the way the data set is *set up and analysed* is not, meaning that the potential of this design to deal with the causality issue is not fully consummated (cf. Gabriel et al., 2019). We feel that in these cases it is misleading and thus incorrect to emphasise the longitudinal nature of the design. Indeed, for generalisation purposes, one might well prefer a cross-sectional study over a diary study with the same number of daily observations.

However, we see this as an unfortunate situation that can be readily addressed by existing methods of analysis. The relatively small number of observations in a typical diary study makes it difficult to estimate cross-lagged relationships in a regression or structural equation modelling framework. However, moving the data from the conventional "wide" format (with each row containing the information of a single participant for all measurement occasions of the study) to the "long" format (with a row containing the information of a particular participant on one specific measurement occasion, allowing participants to contribute more than a single row of data to the data set) allows extension of the multilevel analysis to the estimation of growth curves (see for example Shek & Ma, 2011) focused on the extent and nature of change in the variables over time. One can also incorporate predictors of change into such models. Recent structural equation modelling techniques have been developed specifically for the analysis of intensive data – such as that resulting from diary studies (see Geiser, 2020; McNeish & Hamaker, 2020).

There is future for the cross-sectional design yet, and other implications. The reasoning above suggests that longitudinal studies that do not involve repeated measurements or in which no advantage is taken of the temporal information in the original data set are hardly, if at all, useful in examining causal research questions, and no more so than regular cross-sectional designs. In practice, these alternatives to both cross-sectional and repeated-measures longitudinal studies may be little more than shiny toys that researchers in a saturated field, where the competition for journal space is fierce, use to distinguish their research from that of others, without adding real value to their research. By extension, the growing tendency among reviewers and journals to routinely discard cross-sectional studies in favour of more sophisticated (but not necessarily stronger) designs is unfortunate. At *Work & Stress* we believe that content should come first: novel, interesting, and inspiring work will always take precedence over other research, even if the latter employs a stronger (or just a fancier) design than the first. To be sure, innovative work that is based on a strong design is even better, and since the number of submissions to *Work & Stress* has increased strongly over the last few years, we may be forced to emphasise design issues more heavily than before. Yet, we feel that cross-sectional work should not be discarded just for the sake of its design.

Our reasoning also has several other implications and suggestions. First, researchers (and perhaps readers as well) should be aware that a longitudinal design involving repeated measurement of at least some variables, does *not automatically* allow for causal inferences. The choices researchers make when preparing and analysing their data set can severely limit possibilities for drawing causal inferences, as is evident in many diary-based studies. Second, since the label "longitudinal" tends to be equated to "allows for causal inferences," researchers who do not take advantage of the temporal information in the data or who do not study intra-individual change should make this very clear from the start. Indeed, we would advise them to *refrain from using this term altogether* to prevent any misunderstandings. Thirdly, given the apparent confusion surrounding the use of the term "longitudinal" as well as the availability of clear labels for the designs that are commonly grouped under this heading, one may ask whether *we should continue using this term*. Rather, perhaps we should just focus on labels like "diary design," "time-separated design," "time series design" or "repeated-measures design." Our final recommendation applies to the time-separated design. If researchers have the opportunity to survey participants more than once, we would encourage them to *turn this design into a true repeated-measures design* by collecting antecedent and outcome data at more than one time point. True, this would increase participant burden – but only slightly, while doing so would greatly increase a study's ability to draw causal inferences.

The present issue

This issue presents five papers that combine novelty with a design that goes beyond the regular cross-sectional study. Balducci and colleagues (2021) present both a diary study and a one-year, two-wave longitudinal study to examine the impact of workaholism on workload, exhaustion, and job performance. Tóth-Király et al. (2021) also focus on workaholism, relating it to work engagement and passion in a two-wave, three-year panel study. Nielsen and colleagues (2021) employ a pre-posttest design to study the outcomes of organisational change. Clauss and colleagues (2021) examine psychological detachment as a mediator between workload and exhaustion in a three-wave, time-separated design. Finally, Lavaysse and Probst (2021) present a three-wave growth curve analysis to examine the effects of coping with stereotype threat on performance and workplace safety. Together, these studies show how researchers can use narrow and broad longitudinal designs to their advantage, without making overly bold claims regarding the causal nature of the relationships under study.

References

Balducci, C., Alessandrini, G., Zaniboni, S., Avanzi, L., Borgogni, L., & Fraccaroli, L. (2021). The impact of workaholism on day-level workload, emotional exhaustion, and on longer-term job performance. *Work & Stress*, *35*(1), 6–26. https://doi.org/10.1080/02678373.2020.1735569

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, *54*(1), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Clauss, E., Hoppe, A., Schachler, V., & O'Shea, D. (2021). Occupational self-efficacy and work engagement as moderators in the stressor-detachment model. *Work & Stress*, *35*(1), 74–92. https://doi.org/10.1080/02678373.2020.1743790

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2019). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, *22*(4), 969–1006. https://doi.org/10.1177/1094428118802626

Geiser, C. (2020). *Longitudinal structural equation modeling with Mplus: A latent state-trait perspective*. Guilford Publications.

Lavaysse, L., & Probst, T. (2021). Pregnancy and workplace accidents: The impact of stereotype threat. *Work & Stress*, *35*(1), 93–109. https://doi.org/10.1080/02678373.2020.1774937

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, *25*(5), 610–635. https://doi.org/10.1037/met0000250

Nielsen, K. M., Dawson, J., Von Thiele Schwartz, U., & Hasson, H. (2021). What about me? The impact of employee change agents' person-role fit on their job satisfaction during organisational change. *Work & Stress*, *35*(1), 57–73. https://doi.org/10.1080/02678373.2020.1730481

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94–120. https://doi.org/10.1177/0149206309352110

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88* (5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Shek, D. T., & Ma, C. (2011). Longitudinal data analyses using linear mixed models in SPSS: Concepts, procedures and illustrations. *The Scientific World Journal*, *11*, 42–76. https://doi.org/10.1100/tsw2011.2

Taris, T. (2000). *A primer in longitudinal data analysis*. Sage Publications.

Tóth-Király, I., Morin, A. J. S., & Salmela-Aro, K. (2021). A longitudinal perspective on the associations between work engagement and workaholism. *Work & Stress*, *35*(1), 27–56. https://doi.org/10.1080/02678373.2020.1801888

Wang, M., Beal, D. J., Chan, D., Newman, D. A., Vancouver, J. B., & Vandenberg, R. J. (2017). Longitudinal research: A panel discussion on conceptual issues, research design, and statistical techniques. *Work, Aging, and Retirement*, *3*(1), 1–24. https://doi.org/10.1093/workar/waw033

Toon W. Taris
Department of Social, Health and Organizational Psychology, Utrecht University, Utrecht, Netherlands
a.w.taris@uu.nl http://orcid.org/0000-0003-1946-3307

Stacey R. Kessler
Leven School of Management, Entrepreneurship and Hospitality, Coles College of Business, Kennesaw State University, Kennesaw, GA, USA

E. Kevin Kelloway
Psychology Department, Saint Mary's University, Halifax, Canada