Journal of Business and Psychology (2019) 34:125–137
https://doi.org/10.1007/s10869-018-09613-8

ORIGINAL RESEARCH

Do Not Cross Me: Optimizing the Use of Cross-Sectional Designs

Paul E. Spector[1]

Published online: 7 January 2019

Abstract

The cross-sectional research design, especially when used with self-report surveys, is held in low esteem despite its widespread use. It is generally accepted that the longitudinal design offers considerable advantages and should be preferred due to its ability to shed light on causal connections. In this paper, I will argue that the ability of the longitudinal design to reflect causality has been overstated and that it offers limited advantages over the cross-sectional design in most cases in which it is used. The nature of causal inference from a philosophy of science perspective is used to illustrate how cross-sectional designs can provide evidence for relationships among variables and can be used to rule out many potential alternative explanations for those relationships. Strategies for optimizing the use of cross-sectional designs are noted, including the inclusion of control variables to rule out spurious relationships, the addition of alternative sources of data, and the incorporation of experimental methods. Best practice advice is offered for the use of both cross-sectional and longitudinal designs, as well as for authors writing and for reviewers evaluating papers that report results of cross-sectional studies.

Keywords Causal inference · Causality · Cross-sectional design · Longitudinal design · Method variance · Philosophy of science · Research design · Research methodology

There is perhaps no research design that is more utilized yet more maligned than the cross-sectional design, especially when coupled with a single-source self-report survey. Authors of papers utilizing this design will typically note how their conclusions are limited, and it is not uncommon for them to discount the conclusions of their own papers in a limitations section, suggesting that one cannot really trust their results to address the paper's main expressed purpose, such as to test for the presence of mediation. There seems to be a universal condemnation of the cross-sectional design and at the same time acceptance of the superiority of the longitudinal design in allowing conclusions about temporal precedence and even causality. Often overlooked is that the cross-sectional design can tell us much that is of value and that the longitudinal design is not necessarily superior in providing evidence for causation. In this paper, I will discuss the sorts of inferences that are reasonable to make with the cross-sectional design, contrasting it with the more esteemed longitudinal design in which there is generally more faith placed than is warranted, and to explore design and analysis strategies to optimize the value of a cross-sectional study. I will conclude with advice for researchers about when and how to best use cross-sectional designs and for reviewers about how best to evaluate whether a cross-sectional design was a reasonable choice for addressing the expressed purpose of a paper.

There are two often expressed concerns with the cross-sectional design: common method variance and the inability to draw causal conclusions. Common method variance might arise due to occasion factors that bias different measures in a similar way. The inability to draw confident causal conclusions is due to the lack of temporal elements in the research design that could indicate temporal precedence as a necessary, although not sufficient, element in a causal case. The remedy for both of these concerns most often suggested is using a longitudinal (all variables are assessed at all time points) or prospective (different variables are assessed at different time points) design to introduce the element of time. These designs are often utilized to control for common method variance by separating in time the assessment of proposed independent and dependent variables (Podsakoff, MacKenzie, & Podsakoff, 2012). The separation in time between assessments

✉ Paul E. Spector
pspector@usf.edu

1 Department of Psychology, University of South Florida, PCD 4118, Tampa, FL 33620, USA

of the presumed cause and effect variables is assumed to allow more confident causal conclusions (for an overview of longitudinal design strategies, see Zapf, Dormann, & Frese, 1996).

To demonstrate the extent to which the longitudinal design is accepted as a cure for the limitations of a cross-sectional design, I did a content analysis of the most recent 2018 volume of *Journal of Business and Psychology* to see how many papers used cross-sectional survey methods and of those papers how many suggested the need for longitudinal designs. Of the 45 papers (not counting one erratum) published during this period, 17 (38%) utilized cross-sectional survey designs. In their limitations sections, 12 of 17 (71%) specifically noted the need for longitudinal designs, generally to allow for the determination of causality. Only two of the papers (including one of the 12 advocating longitudinal designs), mentioned the need for experimental and/or quasi-experimental studies. Clearly, researchers consider the longitudinal design as a superior design to be preferred over the cross-sectional, even in cases where multiple data sources were included, such as both employee and supervisor ratings.

Cross-sectional Designs and the Longitudinal Design Remedy

Common Method Variance

That observed, relationships among variables can be attributable to factors other than the intended constructs goes without saying. Spector and Brannick (2011) distinguish between factors that act upon constructs and produce spurious relationships and factors that affect measures of constructs but not constructs themselves, which constitute measurement biases. For example, suppose one conducts a study relating employee perceptions of supervisor consideration and employee job satisfaction, both of which are related to mood. If mood affects perceptions and job attitudes, it might cause a spurious correlation between them, that is, supervisor consideration does not cause job satisfaction; their intercorrelation is due to mood as a common cause of both. This would represent the unmeasured variables problem in models where mood is not included. If, however, mood has no impact on the underlying constructs, but merely affects their assessment, mood would be biasing those assessments and would serve as a source of common method variance, that is, an unintended influence on the assessment of the variables of interest.

A limitation of the cross-sectional design is that there might be transient occasion factors that bias measures and serve as sources of common method variance. For example, the mood of the individual completing a survey might affect responses to items across scales, and this might inflate correlations. Temporal separation can be an effective strategy to control for such method variance sources, but it should be kept in mind that it can only control a narrow range of potential sources, such as consistency biases within a single survey and occasion factors such as momentary mood. They do not control for potential sources of common method variance that are more enduring, such as characteristics of people or measurement methods. Furthermore, it is not entirely clear that occasion factors are a widespread problem with cross-sectional designs. Comparisons of corresponding cross-sectional versus longitudinal correlations in meta-analyses do not uniformly find larger correlations from cross-sectional designs (e.g., Nixon, Mazzola, Bauer, Krueger, & Spector, 2011; Pindek & Spector, 2016), and even when cross-sectional correlations are larger, it is not necessarily due to common method variance. Finally, there are methods that can control for these potential method variance sources within the cross-sectional survey, particularly if they can be identified (Spector, Rosen, Richardson, Williams, & Johnson, 2017), a point that I will return to later in the discussion of how best to utilize cross-sectional designs.

Causal Conclusions

The problem of inability to draw causal conclusions is certainly endemic to any method in which purported cause is not assessed prior to hypothesized effect. This limitation is obvious with cross-section designs where all variables are assessed contemporaneously. Merely separating measurements in time, however, is not in and of itself going to be any more conclusive, as most studies that utilize longitudinal or prospective designs fail to choose time points so that causes are assessed prior to effects. Rather, arbitrary points in time are chosen in most cases after the underlying causal process has been completed and the system has achieved a steady state or what Mitchell and James (2001) termed equilibration. Unfortunately, there is a tendency to consider the longitudinal design as a panacea to the problem of drawing causal conclusions, treating the design as what Meehl (1971) might have called an automatic inference machine.

The issue of drawing a causal conclusion is based on a preponderance of evidence, much like a court case. Much has been written by philosophers of science about the nature of causality and building the causal case. In our field, a causal case is often said to be built on three elements attributed to John Stuart Mill (Shadish, Cook, & Campbell, 2002):

1. Proposed cause and effect are related.
2. Proposed cause occurs prior to effect.
3. We can rule out feasible alternative explanations for observations of 1 and 2.

To that, a number of philosophers (e.g., Berofsky, 1966) discuss a fourth element of the need to articulate a mechanism through which the cause can lead to an effect.

4. Proposed cause works through an articulated mechanism.

A conceptualization of causality that addresses elements 2 and 3 and is particularly relevant to scientific research is the manipulationist approach that suggests causal conclusions arise from observing the results of interventions that precede effects in time (Illari & Russo, 2014; Woodward, 2003, 2017). As proponents of this approach, Hausman and Woodward (1999) describe causation as existing when the manipulation of a purported cause is followed by a change in the purported effect. They describe the process as "one can wiggle Y by wiggling X" (p. 533). In a research setting, one would test for the impact of X on Y by either creating or observing an "intervention" (something that acts upon and changes X) and then observing the subsequent change in Y. This approach would characterize an experiment or quasi-experiment (Shadish et al., 2002), but they note that experimentation is not required, as observations of a naturally occurring intervention would be sufficient to draw a reasonable causal conclusion. As straight-forward as this approach might seem in theory, there are complexities that make its implementation difficult in practice. First, the change (or wiggle) in X must precede the change in Y. Thus, it is necessary to nail down the temporal precedence of X and Y (Spector & Meier, 2014). Second, for a connection between X and Y to be causal, three conditions must be met (Illari & Russo, 2014):

1. The change in X is due only to the intervention. There is no additional factor that is also causing X, for example, that something associated with the intervention is a cause, or there is not some other process independent of the intervention occurring at the same time. This is clearly a potential problem in nonexperimental and quasi-experimental studies where it can be difficult to isolate the effects of an intervention independent of other things occurring in the environment at the same time (see Shadish et al., 2002 for discussion of threats to design validity.
2. The action of the intervention is only through X and not directly on Y. Intervention cannot affect Y independent of the effect on X, or through some other cause. This can occur in experimental, quasi-experimental, and nonexperimental studies, as the intervention intended to manipulate the theoretical construct represented as X is having an effect on Y for some other reason. For example, there can be unintended demand characteristics and experimenter effects that influence Y independent of X.
3. There are no additional causes of Y that are correlated with the intervention. In nonexperimental studies, there can be a host of factors that are related to the intervention that are the real cause of Y.

These three conditions address the third element of a causal case, that of ruling out alternative explanations. In distinguishing the intervention from the cause (X variable), the manipulationists are making explicit the distinction between a theoretical construct about which we wish to draw inferences and the operationalization used to manipulate it. The notion that the observed connection between X and Y might be attributable to additional causal variables has been discussed in other contexts such as endogeneity (Antonakis, Bendahan, Jacquart, & Lalive, 2010) and causal mediation (Imai, Keele, & Tingley, 2010; Pearl, 2014).

The manipulationist perspective can be illustrated with a thought experiment based on an example from Illari and Russo (2014) involving smoking, yellow fingers, and lung cancer. Imagine you are an epidemiologist in the early twentieth century and you observe that individuals who die of lung cancer often have yellow fingers. This leads to a hypothesis that yellow fingers cause lung cancer. To create an intervention that manipulates the hypothesize cause, you come up with the idea that smoking (the intervention) can be an effective way to create yellow fingers. You devise an experiment in a prison where you randomly assign long-term prisoners to a unit that allows smoking versus a unit that does not. Thirty years later, you end the experiment, analyze your data, and conclude that yellow fingers cause lung cancer. Of course, in this example, we know that the second requirement is not met, that is, the action of the intervention is not through yellow fingers, but rather is directly on cancer. We could test that possibility if we did a second experiment with a different intervention, say using bleach to remove the yellow stains from people whose fingers were stained by smoking. When this second experiment failed to find that removing the yellow stains reduced the likelihood of lung cancer, we would conclude that something else was the causal agent.

The idea of using a second experiment with a different intervention is key to the testing of theories and hypotheses in the sciences. The concept of converging operations suggests that scientists continue to attempt to disconfirm their findings and theories by devising new ways to test hypotheses. As new methods continue to produce consistent findings, confidence grows in the interpretation of results because alternative explanations are systematically tested and eliminated. This systematic approach to identifying and eliminating alternative explanations is used too infrequently in organizational research.

Contrasting Cross-sectional and Longitudinal Designs

To draw a confident causal conclusion, we need to link our purported cause to purported effect. We can start by just establishing that our X and Y variables are correlated. This can be done with a cross-sectional design as well as with each individual wave of a longitudinal design where X and Y are assessed at each wave. Of course, showing mere correlations at one point in time is hardly conclusive, and at best, such findings are merely consistent with our hypothesized causal relationship, but it is an important first step. The resources needed to conduct the typical cross-sectional design are far less than those needed to clearly establish temporal order, so it makes sense to begin the investigation of a potential causal link by establishing the existence of covariation.

The second element of the case is to show that X preceded Y in time. For the most part, this is not possible to demonstrate with a cross-sectional design (but I will discuss an exception later), but it is typically not possible with a longitudinal design either considering the way such designs are usually applied. In order to establish temporal order, you must assess X before Y happens, and assess Y after X occurs. Although a longitudinal design can provide a *measurement* of X before a *measurement* of Y, that is not the same thing as assessing X prior to Y happening and Y after X has occurred. In most cases, our studies are not assessing discrete events or states (a notable exception is the study of turnover), but rather, we are assessing levels of variables, such as perceptions of the work environment or attitudes about the job. Many such variables have both a stable trait and unstable state component, the relative magnitude of which depends on time (Cole, Martin, & Steiger, 2005). We do not generally know when the levels of our X and Y variables were achieved, and which might have occurred prior to the other. To use manipulationist thinking, if our X variable is the perception of job complexity, and our Y variable is job satisfaction, we would need to choose our timeframe in such a way that when complexity is wiggled, satisfaction will follow at some specifiable interval. We would need to know when the perception of complexity occurred relative to the experience of job satisfaction. Merely assessing complexity and satisfaction at different times does not guarantee that complexity itself preceded satisfaction, only that we measured complexity before we measured satisfaction. Temporal precedence would be more easily discernable if some intervention occurred, whether designed for the study or naturally occurring. So if on some specified date, jobs were redesigned to produce more complexity, one could do a before-after comparison to see that job satisfaction changed after complexity changed. But if we merely assess complexity and satisfaction at two arbitrary time points, it will not be possible to observe the effects of complexity because it could not be determined when the wiggles of complexity and satisfaction occurred.

An important aspect to consider in establishing a causal connection has to do with the timeframe over which a phenomenon occurs, or the lag between cause and effect. To simplify, let us assume X and Y are discrete events, and that it takes time for X to affect Y. If we assess X and Y at the same time, we will fail to detect the effects of X on Y because the process has not been completed. To provide an adequate test of the effects of X on Y, we would need to know how long the lag is between X occurring and Y happening, say starting to smoke and then developing lung cancer.

It is tempting to assume that longitudinal designs yield better estimates of relationships, even in cases where causal conclusions cannot be reached, but that is not necessarily the case. Unfortunately, longitudinal designs can lead to erroneous inference when the timeframe chosen does not match the timeframe of the phenomenon in question. Suppose, for example, we are interested in the connection between perceived workload and emotional strain in accountants. We might design a two-wave longitudinal study to test the idea that workload is the cause of strain and take assessments on October 1 and the following April 1 (a typical 6-month lag). The workload for many US accountants is cyclical, however, with workloads and working hours being highly influenced by the April 15 federal tax filing deadline. From the beginning of the year through April, workloads become increasingly heavy and almost certainly lead to greater emotional strain due to work overload for many accountants. If we wish to study the connection between workload and strain, we need to look at their concurrent levels. If we conduct a cross-lag analysis with our two-wave study, we would be testing whether October workload predicts April strain, which in this case is nonsensical because the lag between workload and strain is not 6 months. So, we might find no lagged effect and conclude workload does not influence strain, when in fact it does. Alternately, suppose accountants who experience excessive workloads in the first third of the year are given lighter loads later in the year, so that the greater the April workload, the lighter the October workload. In that case, we might find that October workload is negatively related to strain because those with the lightest October workloads have the greatest April strains. Our conclusion would be backwards, and erroneously suggest that heavy workloads lead to less strain.

If we were serious about nailing down the workload-strain relationship, a better design would be a daily diary study that would take place over months before and after April. Each day, a sample of accountants would report their workload and strain. If we are lucky, the lag is at least a day so we can show that workload change precedes strain change, or that when the workload is wiggled, the next day's strain follows. If lags are very short, however, so that workload and strain occur almost simultaneously, it would be very hard to determine if perceptions of heavy workloads led to strain, or if strain (caused by something else) led to perceptions of heavy workloads. In other words, workload would need to be wiggled first to see that strain was subsequently wiggled. Noting that their wiggles occur at the same time is not helpful in addressing the second element of a causal case. To build a case for causality, we need to show that if X changes, Y follows in time.

The third element concerns ruling out alternative possibilities. The manipulationists argue that this can be done by using different interventions to manipulate X, so we can show that it is the manipulation of X and not some other thing that leads to a change in Y. In the earlier thought experiment, this would involve trying different interventions to manipulate yellow stains on fingers, in this example, smoking and bleach. However, there are a wide range of ways in which we can rule out alternative explanations for an observed relationship, some involving research designs to control potential alternative mechanisms and some involving the assessment and then statistical control of possible alternative mechanisms.

Both the cross-sectional and longitudinal design can be used to rule out some alternative possibilities, but in most cases, the longitudinal design fails to offer many advantages. As noted earlier, it can be used to rule out transitory occasion factors and consistency biases that might serve as sources of common method variance. However, as noted earlier, unless proper lags are chosen, this design might inadvertently fail to capture the true relationship among X and Y because the values of X changed between assessment waves, and so the longitudinal design is unable to accurately indicate the effects of X on future Y.

The final element has to do with articulating an explanatory mechanism through which X has an effect on Y, that is, presenting an argument about why the causal process occurs. This element is very much about theory and the development of explanations that go beyond merely describing a phenomenon. The testing and validation of proposed mechanisms serves an important role in directing future research and elucidating what variables must be taken into account when we test for our hypothesized effects (Antonakis et al., 2010). It also can provide additional confidence in the causal case that X leads to Y and not that there is some other variable associated with X that is the real cause. The richer the explanation, and the more the explanation can be empirically supported, the more confidence we can have in the causal claims.

Again, both cross-sectional and longitudinal designs can be helpful in testing for some explanatory mechanisms. This generally involves the assessment of additional variables in order to see if observed patterns of relationships are consistent with hypothesized relationships. If for example, we assume the effect of X on Y is through M (a mediation chain), we would expect to observe a given pattern of relationships. Of course, in order to adequately test this chain, we would need to assess X before M and Y happened and M after X occurred and before Y happened. This can be thought of with a billiard metaphor: the stick first hits the cue ball which then hits the 6 ball which then hits the 12 ball that rolls into the pocket. If all we had was a cross-sectional slice (e.g., a photograph) of the pool table, we would not know for certain which ball hit which other ball and in what order. To determine if the purported chain of events occurs, we must choose our lags in the longitudinal design so that we assess X after it occurs but before M and Y happen and then assess M after it occurs but before Y happens and then Y after M occurs. This means we know what lags to choose, which for most longitudinal designs we do not. Merely choosing some arbitrary time points does not provide more definitive evidence than does a cross-sectional design and might well lead to erroneous inference as in the accountant example. If we were to take a series of photographs of our pool table beginning at the time of the cue stick hitting the cue ball, the lag would determine whether those photographs would be helpful in indicating the sequence of events. Choose around 0.1 seconds, and likely, we would see clearly the causal chain. Choose a lag between photographs of 2 seconds, and most of our photographs would be taken after equilibration, that is, the series of collisions and final location of the 12 ball have already occurred.

Getting the Most from a Cross-sectional Design

Despite its limitations and the likely pushback from reviewers, the cross-sectional design remains the most popular one for many topics studied in organizational research and other fields that rely on survey methods. For example, Spector and Pindek (2016) content analyzed papers in the two most impactful occupational health psychology journals, finding that the single-source cross-sectional self-report study was by far the most popular (41% of articles). Note this is approximately the same as in my analysis of JBP (38%). There are good reasons to use the cross-sectional design, as it is efficient in the use of scarce researcher resources. It makes sense to start new areas of inquiry with the most efficient methods to provide initial evidence that a research question is deserving of attention. It also makes sense to use the cross-sectional design in more mature areas of inquiry to rule out alternative explanations where possible. However, to make the most of cross-sectional designs, researchers should consider what these designs can and cannot tell us and how best to implement such studies.

As noted, there are four elements to a causal case: establishing covariation, temporal precedence, ruling out alternatives, and explanatory mechanism. The cross-sectional design is able to provide evidence for all four if properly utilized. The evidence for the first is undoubtedly the strongest, as the cross-sectional design can certainly indicate whether two variables are related. Temporal precedence can be addressed for some cases if the proper approaches are used that incorporate an element of time. Alternatives and explanatory mechanisms can be tested by including variables in a study that can be added to analyses to rule them in or out. These are certainly not as conclusive as experimental or quasi-experimental studies, or those that can track phenomena over time to nail down the order in which things occurred, but they can be quite helpful.

Establishing Covariation

Most research in organizations and many other domains is at least implicitly driven by what might be called the environment to perception to outcome (E-P-O) paradigm. The idea is that conditions and events occur in the person's environment that are appraised and perceived by people and that lead to a variety of outcomes that included attitudes (e.g., job satisfaction), behavior (e.g., performance), emotions, and health symptoms. This general framework assumes that the person is embedded in the environment and perceives the environment in ways that lead to outcomes. Embellishments include a role for individual differences as both additive and moderating influences on the E-P-O flow. This general paradigm describes research on many topics including job characteristics, job performance, leadership, stress, and teams.

The basic building blocks of model tests are covariation among measures of the different classes of variables. At the simplest, this means establishing relationships between some purported environmental, perceptual, and outcome variables. At early stages of research, to address these sorts, or other sorts of questions, the cross-sectional design is the most useful. Although the design is vulnerable to transitory biases and occasion factors, it provides a snapshot of the extent to which the X and Y variables of interest are related without introducing the complexities of temporal flows that might distort relationships, as in the earlier example of cyclical workload and strain in accountants. The cross-sectional design is of particular value when the underlying processes being studied have already occurred (i.e., equilibration has been achieved), and what is being studied is the final state of the system whereby individuals who are high on say perceptions will tend to be high (or low) on outcomes.

There is little advantage of utilizing a longitudinal design in such cases unless one can be certain that one has the correct lags, that is, one can assess the purported cause before the effect has happened, and one waits the appropriate time before assessing the purported effect. With phenomena that might be linked quickly, this would be quite difficult to achieve in practice. For example, daily diary studies allow for the assessment of purported causes and effects within a day. Say one wishes to determine the effects of some daily experience, such as being treated rudely by someone, on emotional state. Emotional state and rude experience can certainly be assessed at the end of the day, and one can assess emotional state at the beginning of the day to show that for individuals who experience rudeness, the emotional state changes over the work day. However, without knowing the precise time of the rude incident, and when the emotional state changed, it will be impossible to say with certainty which was the cause and which was the effect. It is certainly possible that some other factor was responsible for both the rudeness and the change in emotional state. For example, suppose the people being studied are customer service employees and that customer service failure is the actual cause of rudeness by a customer and the negative emotional state of the employee. From the manipulationist point of view, this violates the second condition for a causal case in that the naturally occurring intervention (service failure) is the cause of Y rather than X being the cause of Y.

The use of a cross-sectional design would be inappropriate in cases where equilibration has not yet occurred, for example, if we assess X and Y at the same time, but Y has not yet happened. For example, if we want to study the effects of smoking on heart disease, it would not be fruitful to conduct a cross-sectional study of 20-year-olds because there has not been sufficient time for young smokers to have developed the disease. There can be similar phenomena in organizations where some outcomes can take time to occur, so depending on the sample, it might or might not make sense to conduct a cross-sectional design. In the area of workplace health, for example, we might not expect that job conditions and attitudes would relate to serious musculoskeletal disorders, such as carpal tunnel syndrome, in newly hired young workers because they would not have had sufficient exposure. In such cases, long-term studies would need to be conducted. Of course, it would be possible to conduct a cross-sectional study with older workers to see if elapsed time of exposure relates to the disorder. We likely would find that tenure relates to our health outcome, but we would need to disentangle the effects of exposure duration from merely the effects of age, and of course, we would need to figure out which, from the myriad of things people are exposed to, is the causal agent.

Temporal Precedence

Many of the constructs we are interested in studying are characteristics of people and organizations that can vary in levels over time. Many are internal states (e.g., cognitions and emotions) that are difficult to assess outside of self-reports. Often, we want to know if changes in one construct will lead to changes in another, but in many cases, it is difficult to assess them at appropriate time points to see if X changed prior to Y and if Y changed following a change in X. This is far easier if X and/or Y is a discrete event, so we might ask, for example, if dissatisfied individuals are more likely to quit their jobs, or if individuals who score low on conscientiousness as applicants are more likely to be caught stealing after being hired.

Most studies assess variables with scales that do not ask about the order or time frame in which things occur. Such scales sometimes ask about current state (e.g., current mood), about some specified time span (e.g., helping behavior over the past 30 days), or about the phenomenon in general (e.g., measures of personality traits). When questions are asked in this way, it is not possible to tease apart when one occurred relative to another. It is not, however, impossible to incorporate time into survey questions using retrospective approaches.

The retrospective event history is a technique in which individuals report on specific events and their temporal order (Glick, Huber, Miller, Doty, & Sutcliffe, 1990; Tuma & Hannan, 1984). This technique can be used with both qualitative and quantitative methods. Qualitative methods might include focus groups, interviews, or surveys where people are asked to recall specific events and when they occurred. With interviews, in-depth discussions can include not only the order of events, but also informants' explanations and interpretations of events that can be helpful in articulating tentative mechanisms deserving of further testing. What distinguishes the event history from the typical quantitative survey study is the introduction of time in respondent reports. They can be asked specifically when events occurred (e.g., dates), or they can be asked the order of events. In the simple case of two variables, X and Y, respondents can be asked to indicate which happened first and even how long it was from the first one to the second one. Respondents can be given ratings and can be asked to indicate the level of a variable at two or more points in time.

There is certainly a potential issue of attribution and recall bias with the use of retrospective reports that render this method far from conclusive. Such biases represent sources of common method variance and alternative explanations for findings. There are two points to keep in mind, however. First, all methods are limited, and no one study using any method is conclusive. Even randomized experiments are plagued with potential demand characteristics and experimenter effects, and limitations to generalizability, not to mention uncertainty about how well the intervention "wiggled" the intended X variable. It is not that the retrospective event history by itself will provide sufficient evidence for a causal case, but it can provide compelling evidence that a certain causal flow is likely and is deserving of further study using other methods. Second, when retrospective reports involve specific events, it might be possible to verify the time frame by consulting records or other sources of information. It is also possible with multiple respondents to check for inter-rater agreement about the order of events and the timeframe. If multiple respondents provide the same order of events, convergence can provide some level of confidence that the order is correct.

The retrospective report has the potential to indicate the temporal order of events and the likely time lag. This can be an improvement over the typical cross-sectional design and, in many cases, over the typical longitudinal design that includes arbitrary time periods and fails to assess X prior to Y. Results of this sort of study might provide evidence concerning the lag between our X and Y variables that can inform the design of a longitudinal study that can provide an alternative means of confirming or disconfirming the temporal flow and spacing.

Ruling out Alternative Explanations

There are several ways in which cross-sectional designs can be used to rule out alternative explanations for results. Some involve features of the research design, and others involve statistical controls to rule out the possibility that a suspected alternative might be at play. It is also possible to test for the possibility of alternative explanations by embedding cross-sectional approaches into an experimental design in which individuals are randomly assigned to different conditions, which might involve different formats of a survey, or different conditions under which the study is conducted.

Alternative Sources of Data A commonly used approach to expand the cross-sectional design is to utilize multiple sources of data, such as employee, peer, and supervisor reports. Pindek and Spector (2016) found that 10% of OHP articles used this approach, although they did not report how many of those used a cross-sectional design. An advantage of using an alternative source is that it can serve as a control for some sources of method variance (Podsakoff et al., 2012). Certainly, factors that might bias self-report measures can be reduced when other sources of data are included. It should be noted, however, that different sources can share biases when those sources are in contact with one another, such as employees and their supervisors (Spector, 2006). Nevertheless, this approach can be helpful when one has in mind specific biases that might be ruled out by the use of the other source. For example, if one wishes to rule out the possibility that observed correlations are due to a personality variable, such as emotional stability, an alternative source might be helpful.

Although we sometimes rely on alternative sources of data to assess constructs, it is not always clear that other sources are able to provide as accurate a measure. For example, it has been shown that self-reports of a number of phenomena show better discriminant validity than other reports where measures of multiple constructs are taken, such as counterproductive work behavior versus organizational citizenship behavior (Dalal, 2005; Spector, Bauer, & Fox, 2010) and job characteristics (Glick, Jenkins, & Gupta, 1986; Spector, Fox, & Van Katwyk, 1999). Other reports might be considered global measures that lack the specificity to address many more precise questions.

Adding a second source to a self-report study can be helpful if used properly. Such sources can answer two main questions. First, how much convergence is there between sources? Some of the constructs we might wish to study are fairly concrete so that individuals might agree about them. Other constructs are more abstract and subjective, so that we should not be surprised that there is little convergence. Spector (1992) compared the convergence of self-reports with coworker and supervisor reports for several measures of the work environment, finding considerable variability ranging from a weighted mean of 0.03 for feedback to 0.45 for workload. Finding convergence can add confidence that self-reports reflect something beyond idiosyncratic subjective impressions and that there might be something that is in some ways objective, that is, that there is a consensus.

Second, one can look at patterns of correlations to rule out the possibility that the observed correlations within a source are attributable to some source-specific bias. If one finds, for example, that corresponding correlations are significant both within and between sources, one can have some confidence that the results go beyond bias. One should not, however, assume that if the between-source correlations are smaller than the within-source, the differences reflect the amount of common method variance. As noted, alternative sources are not necessarily accurate, so their correlations with an outcome are as likely, if not more likely, to be underestimations (Frese & Zapf, 1988).

Ideally, a pattern of results with a multi-source study will show some reasonable level of convergence in the assessment of the same constructs across sources and a similar pattern of between-source versus within-source correlations among purported causes and effects. Such patterns help rule out that the pattern of relationships is due to common biases. This design does not, however, shed direct light on the nature of what shared biases might be, and if results are different between within-source and between-source correlations, one cannot be certain of the reason. Is it that the within-source correlations are due to a common bias or some third variable, or is it that the phenomenon is something that only the respondent can accurately report?

The main issue in deciding to include an alternative source is whether the interest is in only the subjective view of the individual, or in a more objective feature of the person or workplace. Some phenomena concern internal states of people that would be difficult for an alternative source to assess. Other phenomena might exist at an environmental versus perceptual level. For characteristics of jobs and many features of the work environment, it makes sense that coworkers, supervisors, or other alternative sources would be able to provide reasonably converging reports. It should be kept in mind, however, that those reports are in fact subjective and are only objective in the sense that there is some level of consensus, for example, that workloads are heavy versus light. Behaviors are also something that a peer or significant other might be able to report, but it should be considered that it is only public and not private behaviors that would be observable and that the other person might not spend a great deal of time observing the person of interest. Furthermore, people might behave differently in front of others, particularly if those others are in a superior power position. For example, employees might engage in impression management tactics that give supervisors a distorted view of their behavior. In the domain of organizational citizenship behavior, some employees might engage in certain behaviors and not others, at certain times and not others, directed toward certain people and not others, and in front of certain people rather than others (Bolino, 1999). Thus, the peer or supervisor might get an exaggerated impression of a given employee's level of citizenship behavior relative to other employees who are not engaging in impression management.

Adding "Control Variables" to Rule out Alternative Explanations The multi-source study is in a sense an unmeasured bias approach in which the design can control for the action of a class of factors that are neither identified nor measured. A more specific way to rule out alternative explanations for observed correlations is to identify factors that reflect the explanation and assess them. This can be done within the context of a single-source or multi-source design. In the context of common method variance, Podsakoff et al. (2012) identified several such sources, such as mood or neuroticism. However, alternative explanations do not have to involve common method variance or biased measurement. Rather, an alternative mechanism might be something that is the real cause of our X and Y variables, rendering their correlation as spurious.

The addition of control variables to rule out alternative explanations requires the identification of what those alternatives might be. This could be based on empirical observation, intuition, or theory. To return to the earlier smoking example, once the connection between yellow fingers and lung cancer was established, the next step in the research program would be to see if there are alternative explanations to rule in or out. In this case, it might be ruled out that exposures to yellow ink or yellow paint are likely causes. A series of studies that search for a common factor that when controlled eliminated both yellow stains and lung cancer would provide evidence for the underlying process. Although the original thought experiment used a prospective design, the same logic could apply with a cross-sectional design. Suppose you have a sample of people, some of whom have cancer and some not, some of whom have yellow fingers and some not, some of whom are painters and some not, and so on. Analyses could be conducted to see if the correlation between yellow fingers and lung cancer disappears after controlling for these potential explanations. As they are eliminated one by one, confidence would increase that our X might be a cause of Y, but in the absence of some intervention that affects X and then Y, it will be difficult to reach a final conclusion just by ruling out alternatives alone.

The appropriate use of control variables is very much based on the development of theoretical explanations that propose alternative mechanisms that explain the relationship between X and Y. If we know nothing at all about lung cancer, there is little upon which to propose alternative mechanisms. As we do more research, and note what does and does not relate to lung cancer, we can begin to piece together possible scenarios. We might note that lung cancer is associated with smoking, yellow fingers, and other characteristics of people. As we collect and analyze data, we will eventually come to the realization that the common denominator in lung cancer is smoking and that the other things that are associated with lung cancer, like yellow fingers, are a by-product of smoking. Of course, in reality, smoking is not the only cause of lung cancer, so the picture is far more complex. The same is true of organizational and nonorganizational phenomena. There is not necessarily a one-to-one correspondence between our X and Y variables, so there can be a number of factors that contribute to Y other than X. Nevertheless, the process of using control variables to rule out potential alternative explanations is an effective way to progress our understanding of the connection between a given X and a given Y variable.

Experimental Approaches Cross-sectional studies can be experimental in that X and Y are assessed under different conditions to rule out the effects of potential third variables that could be manipulated. For example, if one suspects that mood might serve as a common cause of both X and Y, one could measure X and Y under varying conditions of mood. Thus, in one condition, you measure X, administer a mood induction manipulation, and then measure Y. Respondents could be randomly assigned to a positive, negative, or no mood induction condition. If mood is in fact the common cause, changing the mood between measuring X and Y should reduce their correlation.

Measurement characteristics are another area that can be studied with an experimental approach. The design of your scales can be manipulated to see if there is an effect on the X-Y relationship. This approach was taken by Spector et al. (2010) to test hypotheses suggested by (Dalal, 2005) that the strong negative correlation between counterproductive work behavior and organizational citizenship behavior might be due to measurement artifacts. Spector et al. experimentally manipulated the artifacts, including response formats (agreement versus frequency), the content of the scales (overlapping or nonoverlapping items), and the source of ratings (self versus supervisor) to show support for Dalal's suggestions.

Best Practices Advice for Cross-sectional Designs

Cross-sectional designs should be considered a basic tool for conducting research. They are relatively cheap to conduct, can be highly efficient in researcher and participant time, and can adequately address many questions. They can be an important starting point for a programmatic approach to addressing a research question that begins with simple designs and builds design complexity as more information becomes available that can inform how subsequent study designs should be formulated. There are a number of purposes for which cross-sectional designs are optimal, as well as purposes for which longitudinal designs are better suited. These should be kept in mind at the time of planning a study. Once the study is complete, it can be an uphill battle to convince reviewers that a paper describing a cross-sectional study is worthy of publication. Authors can enhance their publication chances by making a case for the design they chose. On the other side, reviewers should be open to the possibility that a paper using a cross-sectional design can have significant value, although it is up to authors to make that case.

When to Use a Cross-sectional Design

Cross-sectional designs should be the method of choice in the following situations:

- You do not know if X and Y covary. This is the case when you get into a new domain where little is known, or when you are investigating a new variable in an old domain. For example, as I write this, there is growing interest in cyber behaviors, such as in the cybersecurity domain. Many such behaviors represent new contexts in which people perform their tasks, but we do not know if variables that relate to general job performance might relate to performance in the virtual world. Furthermore, many of our studies investigate boundary conditions where established relationships are attenuated. These are clear cases where cross-sectional designs can indicate whether pairs of variables are related and whether moderators might be at play.
- You are conducting exploratory research. Almost by definition, exploratory research concerns situations where you do not yet know what patterns of relationships to expect and what the timeframe might be. Often, the goal is to investigate whether a large set of potential causes might relate to some outcome variable/s of interest. The first step is to collect large sets of data in order to search for meaningful patterns. This is typically best done first with a cross-sectional design that allows for the efficient collection of many variables from large samples.
- You do not know the timeframe. It is difficult to design a longitudinal study if you do not know how long to expect

for the X variable to cause the Y variable. As noted earlier, utilizing longitudinal designs in such situations risks reaching erroneous conclusions about how strongly and if variables relate to one another. In such cases, the cross-sectional design is a safer bet to indicate covariation. Retrospective reports can shed light on temporal order and timeframe and can indicate potential effects of X on Y (for an example of this approach using critical incidents, see Schwarzmüller, Brosi, & Welpe, 2018)

- You wish to examine the effects of a naturally occurring X. Sometimes, an investigator wishes to examine the effects of some experience or condition that has occurred prior to the study. For example, one might wish to know the effects of surviving a corporate merger. Although the ideal design would be to assess people before and after the merger, such opportunities are not easy to come by. An alternative would be to identify a sample of employees who work in similar jobs, some of whom experienced mergers in the past. A cross-sectional design could be utilized where the X occurred prior to the survey and is assessed with retrospective questions in the survey. Finding differences on Y for people who experienced versus did not experience a merger has its limitations, but it can provide hints that mergers might have long-lasting effects and that such effects are worthy for further study.
- You are interested in ruling out alternative explanations for covariation. This is the case where there are feasible alternative explanations for a phenomenon that can be ruled out by the use of statistical control. For example, Watson, Pennebaker, and Folger (1986) proposed that the relationships between workplace stressors and strains could be attributed to the personality trait of negative affectivity. A number of cross-sectional studies have addressed this possibility by comparing correlations with and without NA as a statistical control (e.g., Brief, Burke, George, Robinson, & Webster, 1988; Chen & Spector, 1991).

When to Use Longitudinal or Prospective Designs

The longitudinal (all variables assessed at each time point) and prospective (different variables are assessed at different times) design can be especially useful when it is possible to determine when X and Y occur relative to one another. In such cases, it is possible to identify a variable as a proxy risk factor (variable predicts another) or a causal risk factor (some intervention occurred prior to Y; Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). They are preferred when:

- You wish to test the effects of an intervention. In cases where you are conducting the intervention, it is desirable to assess your outcomes prior to the intervention to get a baseline measure and then to assess the outcome one or ideally multiple times after the intervention to provide insights into how the effects of the intervention unfold over time. This is particularly useful when there are multiple outcomes and the potential for different patterns of effects. For example, Griffin (1991) utilized a four-wave study (before intervention and 6, 24, and 48 months after) to assess the impact of job redesign on employees. The use of the pretest measure allowed him to show that job satisfaction increased immediately after intervention but returned to baseline, job performance improved but after a delay, and job condition perceptions changed immediately and were maintained over time.
- You wish to test the effects of an experience that occurs between waves. Discrete events can be good candidates for longitudinal investigation, particularly if you can compare individuals who have and have not experienced the events prior to the study. The study of young workers or newly hired employees can be good populations to study. For example, Spector, Yang, and Zhou (2015) investigated the potential impact of climate on physical violence exposure in a sample of new nursing graduates during the first year of their careers. Using a two-wave design, they assessed climate at time 1 in a subsample who had not yet experienced violence and showed that it predicted violence exposure at time 2. This design established climate as a proxy risk factor, although whether it is a causal risk factor could not be determined without manipulating climate to see its effects.
- You know how long the time lag will be between the X and Y variables. In this case, you can assess X at the first wave, and then assess Y at the appropriate time frame during which Y should occur. This is a common strategy in turnover research where job attitudes are assessed at time 1, and turnover is generally assessed 6 to 12 months later.

How to Present Cross-sectional Designs

Reviewers, editors, and readers are rightly skeptical of studies that use cross-sectional designs, so it is incumbent upon authors to present a compelling case for the design approach used and the conclusions reached. Every step of the study design needs to be justified in the context of the paper's expressed purpose. Of course, this process begins at the time the study is planned. A cross-sectional design should be utilized because it was a correct choice given the researcher's purpose, and not just because it was the easy choice. The following should be kept in mind when writing a paper that used a cross-sectional design:

- The expressed purpose should match the design. Not only should the purpose of a study be clearly explained, but a compelling case should also be made for why a cross-sectional design would be appropriate to address that purpose. This means explaining what precisely the study is designed to test, and how that will address the purpose. For example, a paper that is studying a new phenomenon should first establish that the phenomenon is important (merely that it is understudied is insufficient), and then note that the first step is to ascertain some foundational relationships that are not currently understood. Avoided should be cases where the design is clearly not up to the challenge of meeting the purpose, such as testing mediation with a cross-sectional design. Such papers do get published, but it can be a struggle to convince reviewers and editors not to reject out of hand.

- Use cross-sectional design features to rule out alternative explanations. This can include additional measures that allow statistical control or the use of alternative measures that can be helpful when there is interest in more objective features of organizations. Merely establishing relationships among variables is insufficient unless the study is investigating a new phenomenon and reviewers can be convinced it is important enough to justify a study that merely shows X and Y (or Xs and Ys) are related.
- Present a systematic analysis strategy that tells a compelling story. This might mean first establishing a clear relationship between X and Y, and then ruling out feasible alternative explanations, and/or illustrating boundary conditions through the use of moderator tests. Again, the reviewer must be convinced of the importance of what you have done and why it advances our understanding of your phenomenon.
- If possible, incorporate a time element into the design. This could involve a retrospective event history, or it could introduce time in some other way. As noted earlier, it is possible to compare people who had versus did not have some experience in the past to see if it is related to an important variable at the present time. Even though this is a concurrent measure of all variables, it links what happens in the past (assuming people can accurately report) with something in the present.
- Clearly link the data to the conclusion. The basis for inference should be articulated in a logical manner, with feasible alternative explanations noted. Ideally, alternative explanations would be addressed with data, but that requires prior planning, and sometimes, authors get new insights after studies are conducted.
- Limitations sections should be thoughtful. The limitations section should demonstrate some in-depth analysis of the limitations of the study, including arguments and evidence both for and against limitations. It is fine to suggest future directions that might address limitations, but throw-away suggestions such as merely using experimental or longitudinal designs are not helpful. Such statements invite reviewers to wonder why authors did not use the design they felt was really needed. What is more helpful is suggestions for very specific design strategies that can help elucidate the time frame of the phenomenon, or control for potential confounding variables. Even better are design suggestions that were informed by the current study's results, thus demonstrating the paper's contribution to knowledge.

How to Review Cross-sectional Papers

Many reviewers and editors have a knee-jerk reaction to cross-sectional designs, and some journals even discourage submission of papers that report results of such studies. This is unfortunate because the value of a study is not in its design, but rather in the importance of its question and whether the design can adequately address it. Reviewers should approach papers that report cross-sectional designs, as they would approach all papers, open-mindedly, but with a healthy dose of scientific skepticism. As much as possible, reviewers should be objective judges willing to give the paper a fair hearing. The main question is whether the paper makes a compelling case for what was done and why it is important. Consider the following points:

- Does the paper present a clear purpose of the study?
- Is the question being raised important and relevant?
- Does the design make sense in relationship to the purpose? If the purpose is to test for something that the author disowns in the limitations section (e.g., causal conclusions), the design did not match the question.
- Is an element of time incorporated into the study in some way that makes sense?
- Do the authors tell a compelling and logical story that links their purpose to the design and the data to conclusions?
- Are feasible alternative explanations and potential boundary conditions addressed? This is not always necessary, but can provide an additional contribution that enhances the paper's value.
- Does the study add to our understanding of an important phenomenon? No single study, no matter what the design, is in itself conclusive, but rather, it is a body of research across many researchers using a variety of methods that allow us to have confidence in conclusions. Cross-sectional studies have contributed and will continue to contribute to our knowledge base.

Conclusions

It is unfortunate that the cross-sectional design is held in such low esteem given how much it has contributed to our knowledge of organizational (and nonorganizational) phenomena. The lowly cross-sectional design has served us well throughout the history of organizational (and nonorganizational) research to show connections among variables that can serve as the basis for understanding and theorizing about many phenomena. The knowledge that many pairs of variables are associated, even without knowing the causal connections, is extremely valuable as a basis for theory and the target of intervention. It is not that mere covariation tells us that if we wiggle X, we will get the desired effect on Y, but it tells us that X might be a good place to focus attention if we wish to create an intervention to improve Y. It represents the first and likely easiest step in figuring out whether X might be a cause of Y. In the context of medicine, Kraemer et al. (2001) distinguish a hierarchy of three types of risk factors. A correlate is a risk factor that merely relates to illness. A proxy risk factor is something that predicts disease in the future, but we do not know if it is the cause or is merely associated with the cause. To be a proxy risk factor, we need to show that exposure over time will lead to disease, such as smoking. A causal risk factor, consistent with the manipulationist view, is something that, when manipulated, has a reliable effect on illness. If we get people to stop smoking, their likelihood of cancer declines.

Conducting research studies is not without significant cost, even for studies that are not grant funded. There are monetary costs, as the research time for most organizational researchers is being paid by their employers, and there are opportunity costs as everyone working on a given project could be doing something else. There is the cost of university infrastructures that support research efforts. Our research strategies should maximize efficiencies by programmatically addressing research questions in a logical manner, which means first establishing covariation using relatively inexpensive methods before investing considerable resources into conducting studies that will identify temporal precedence and ultimately causality, at least in a causal risk sense. This means that the cross-sectional method should be the method of choice in cases where we need to establish that variables are related and in cases where this efficient design can achieve our purposes.

At least part of our dissatisfaction with cross-sectional designs is that they are unable to adequately test many of the hypotheses we throw at them, such as mediator effects. Unfortunately, without designing studies so that X occurs before Y happens and Y is assessed after X occurs, we cannot provide evidence for causality that goes beyond what the cross-sectional design can do. Our usual approaches, even using longitudinal designs, are not up to the challenge of addressing mediation (Stone-Romero & Rosopa, 2008), let alone more complex causal connections. We are fooling ourselves if we think that our typical research designs can really tell us about the underlying causal processes that we wish to understand (Spector & Meier, 2014). To provide convincing evidence that X is the cause of Y will require an approach that focuses more on understanding the timing and order of events and less on applying complex statistics to data from designs that cannot tell us the timing and order of events.

The cross-sectional design is well positioned to tell us if variables in which we are interested are related, and it can help us rule out a host of potential alternative explanations for why X and Y are related. Event histories can be used cross-sectionally to provide insights into the likely order of events, and surveys can be designed to ask people for their judgments about the causes of events. Those judgments can be checked for consensus and compared to other forms of evidence to build a causal case. The cross-sectional design is not as anemic as many would believe, nor are the more esteemed longitudinal designs as valuable as is generally assumed. Each has its place in our arsenal of research design tools, with the cross-sectional design being an efficient and invaluable go-to tool for investigating important organizational phenomena.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010.

Berofsky, B. (1966). Causality and general laws. *The Journal of Philosophy, 63*(6), 148–157. https://doi.org/10.2307/2024170.

Bolino, M. C. (1999). Citizenship and impression management: Good soldiers or good actors? *The Academy of Management Review, 24*(1), 82–98. https://doi.org/10.2307/259038.

Brief, A. P., Burke, M. J., George, J. M., Robinson, B. S., & Webster, J. (1988). Should negative affectivity remain an unmeasured variable in the study of job stress? *Journal of Applied Psychology, 73*(2), 193–198.

Chen, P. Y., & Spector, P. E. (1991). Negative affectivity as the underlying cause of correlations between stressors and strains. *Journal of Applied Psychology, 76*(3), 398–407. https://doi.org/10.1037/0021-9010.76.3.398.

Cole, D. A., Martin, N. C., & Steiger, J. H. (2005). Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model. *Psychological Methods, 10*(1), 3–20. https://doi.org/10.1037/1082-989X.10.1.3.

Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology, 90*(6), 1241–1255. https://doi.org/10.1037/0021-9010.90.6.1241.

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). Oxford, England: John Wiley & Sons.

Glick, W. H., Huber, G. P., Miller, C. C., Doty, D. H., & Sutcliffe, K. M. (1990). Studying changes in organizational design and effectiveness: Retrospective event histories and periodic assessments.

Organization Science, 1(3), 293–312. https://doi.org/10.2307/2635007.
Glick, W. H., Jenkins, G., & Gupta, N. (1986). Method versus substance: How strong are underlying relationships between job characteristics and attitudinal outcomes? *Academy of Management Journal, 29*(3), 441–464.
Griffin, R. W. (1991). Effects of work redesign on employee perceptions, attitudes, and behaviors: A long-term investigation. *Academy of Management Journal, 34*(2), 425–435. https://doi.org/10.2307/256449.
Hausman, D. M., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *The British Journal for the Philosophy of Science, 50*(4), 521–583. https://doi.org/10.1093/bjps/50.4.521.
Illari, P., & Russo, F. (2014). *Causality: Philosophical theory meets scientific practice*. Oxford, UK: Oxford University Press.
Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods, 15*(4), 309–334. https://doi.org/10.1037/a0020761.
Kraemer, H. C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *The American Journal of Psychiatry, 158*(6), 848–856. https://doi.org/10.1176/appi.ajp.158.6.848.
Meehl, P. E. (1971). High school yearbooks: A reply to Schwarz. *Journal of Abnormal Psychology, 77*(2), 143–148.
Mitchell, T. R., & James, L. R. (2001). Building better theory: Time and the specification of when things happen. *The Academy of Management Review, 26*(4), 530–547. https://doi.org/10.2307/3560240.
Nixon, A. E., Mazzola, J. J., Bauer, J., Krueger, J. R., & Spector, P. E. (2011). Can work make you sick? A meta-analysis of the relationships between job stressors and physical symptoms. *Work & Stress, 25*(1), 1–22. https://doi.org/10.1080/02678373.2011.569175.
Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods, 19*(4), 459–481. https://doi.org/10.1037/a0036434.
Pindek, S., & Spector, P. E. (2016). Organizational constraints: A meta-analysis of a major stressor. *Work & Stress, 30*(1), 7–25. https://doi.org/10.1080/02678373.2015.1137376.
Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452.
Schwarzmüller, T., Brosi, P., & Welpe, I. M. (2018). Sparking anger and anxiety: Why intense leader anger displays trigger both more deviance and higher work effort in followers. *Journal of Business and Psychology, 33*(6), 761–777. https://doi.org/10.1007/s10869-017-9523-8.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
Spector, P. E. (1992). A consideration of the validity and meaning of self-report measures of job conditions. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 123–151). West Sussex, UK: John Wiley.
Spector, P. E. (2006). Method variance in organizational research: Truth or urban legend? *Organizational Research Methods, 9*(2), 221–232. https://doi.org/10.1177/1094428105284955.
Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology, 95*(4), 781–790. https://doi.org/10.1037/a0019477.
Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14*(2), 287–305. https://doi.org/10.1177/1094428110369842.
Spector, P. E., Fox, S., & Van Katwyk, P. T. (1999). The role of negative affectivity in employee reactions to job characteristics: Bias effect or substantive effect? *Journal of Occupational and Organizational Psychology, 72*(2), 205–218. https://doi.org/10.1348/096317999166608.
Spector, P. E., & Meier, L. L. (2014). Methodologies for the study of organizational behavior processes: How to find your keys in the dark. *Journal of Organizational Behavior, 35*(8), 1109–1119. https://doi.org/10.1002/job.1966.
Spector, P. E., & Pindek, S. (2016). The future of research methods in work and occupational health psychology. *Applied Psychology: An International Review, 65*(2), 412–431. https://doi.org/10.1111/apps.12056.
Spector, P. E., Rosen, C. C., Richardson, H. A., Williams, L. J., & Johnson, R. E. (2017). A new perspective on method variance: A measure-centric approach. *Journal of Management, 0*(0), 0149206316687295. https://doi.org/10.1177/0149206316687295.
Spector, P. E., Yang, L.-Q., & Zhou, Z. E. (2015). A longitudinal investigation of the role of violence prevention climate in exposure to workplace physical violence and verbal abuse. *Work & Stress, 29*(4), 325–340. https://doi.org/10.1080/02678373.2015.1076537.
Stone-Romero, E. F., & Rosopa, P. J. (2008). The relative validity of inferences about mediation as a function of research design characteristics. *Organizational Research Methods, 11*(2), 326–352. https://doi.org/10.1177/1094428107300342.
Tuma, N. B., & Hannan, M. T. (1984). *Social dynamics models and methods*. Saint Louis, US: Elsevier.
Watson, D., Pennebaker, J. W., & Folger, R. (1986). Beyond negative affectivity: Measuring stress and satisfaction in the workplace. *Journal of Organizational Behavior Management, 8*(2), 141–157. https://doi.org/10.1300/J075v08n02_09.
Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: New York City.
Woodward, J. (2017). Scientific explanation. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Retrieved from https://plato.stanford.edu/archives/fall2017/entries/scientific-explanation. Accessed 28 Dec 2018.
Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies in organizational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology, 1*(2), 145–169. https://doi.org/10.1037/1076-8998.1.2.145.